

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

**The Effect of Stress on Memory: Eyewitness Performance in Juveniles and Young Adults**

**Marianne Ball (BLI.MAR028)**

A minor dissertation submitted in partial fulfillment of the requirements for the award of the degree of Master of Arts in Psychological Research

**Supervisor: Professor Colin G. Tredoux**

**Faculty of the Humanities**

**University of Cape Town**

**2009**

The Effect of Stress on Memory: Eyewitness Performance in Juveniles and Young Adults

Marianne Ball (BLLMAR028)

A minor dissertation submitted in partial fulfillment of the requirements for the award of the degree of Master of Arts in Psychological Research

Supervisor: Professor Colin G. Tredoux

Faculty of the Humanities

University of Cape Town

2009

---



**COMPULSORY DECLARATION**

This work has not been previously submitted in whole, or in part, for the award of any degree. It is my own work. Each significant contribution to, and quotation in, this dissertation from the work, or works, of other people has been attributed, and has been cited and referenced.

Signature:  Date: 09 / 04 / 2009

---



## ACKNOWLEDGEMENTS

I would like to thank Professor Colin G. Tredoux for his exceptional supervision. His affability, patience, and invaluable feedback at every stage of this project have made all the difference.

Special thanks go to Anthony Hodge (fellow Masters student) for his encouragement in times of tribulation and his unstinting willingness to help in any capacity.

I would like to acknowledge the contribution of my research assistants, Nicola Rohland and Candice Lawrie. They were eternally good natured and flexible throughout the data collection process, and their commitment was exemplary.

Acknowledgement goes to the A W Mellon Foundation and the National Research Foundation for their generous financial support.

---



## ABSTRACT

An experimental design was used to investigate the relationship between stress and eyewitness memory in adolescents (Experiment 1) and young adults (Experiment 2). Psychosocial stress was induced using a public speaking task. As eyewitness outcome measures, participants were asked to provide a physical description and to make a lineup identification of a) a research assistant with whom they had interacted under stressful circumstances (prior to public speaking), and b) a research assistant with whom they had interacted under neutral circumstances. Participants also responded to questions that tested their memory for central and peripheral details of the interactions. As validity checks on the stress manipulation, participants' heart rate and skin conductance levels were measured using an Ambulatory Monitoring System and anxiety ratings were obtained using the State-Trait Anxiety Inventory. In the adolescent study ( $n = 26$ ), inferior performance on target descriptions, detail recollection and lineup identification was generally associated with the stress condition, but this finding was not strongly supported. In the young adult study ( $n = 22$ ), arousal had no effect on the quality of target descriptions or the accuracy of lineup identifications, but seemed to enhance memory for both central and peripheral details. Future research is called for, particularly with regard to juvenile witnesses. Understanding whether stress impairs eyewitness memory, and subsequently compromises reliability, can assist in preventing misidentifications—the single greatest threat to the delivery of justice.

---



**TABLE OF CONTENTS**

<b>TITLE PAGE</b>	<b>1</b>
<b>COMPULSORY DECLARATION</b>	<b>2</b>
<b>ACKNOWLEDGEMENTS</b>	<b>3</b>
<b>ABSTRACT</b>	<b>4</b>
<b>LIST OF FIGURES</b>	<b>9</b>
<b>LIST OF TABLES</b>	<b>10</b>
<b>CHAPTER ONE: INTRODUCTION</b>	<b>12</b>
<b>CHAPTER TWO: LITERATURE REVIEW</b>	<b>17</b>
Stress and Eyewitness Memory: Background	17
Neurological Research	17
Theoretical Framework	18
Stress and Eyewitness Memory: Empirical Findings	23
Adult Eyewitnesses	23
Child Eyewitnesses	27
Methodological Issues in the Study of Stress	29
Measures of Eyewitness Memory	31
Person Descriptions	31
Interrogative Recall	35
Lineup Identifications	38
Summary and Conclusion	42

---



	6
<b>CHAPTER THREE: EXPERIMENT 1</b>	<b>44</b>
<i>Method</i>	44
Rationale, Aim and Hypotheses	44
Design	45
Participants	46
Measures, Instruments and Materials	47
Physiological Measures	47
Self-reported Anxiety	47
Psychosocial Stressor	47
Research Assistants	48
Evaluators	49
Stimulus Materials	49
Eyewitness Tasks	50
Scoring Materials	52
Procedure	53
Data Analysis	57
Scoring	57
Statistical Analysis	59
<i>Results</i>	60
Measures of Stress	60
Heart Rate	60
Skin Conductance Level	60
Autonomic Reactivity	61
Self-reported Anxiety	61
Overview	63

---



	7
Measures of Memory	63
Target Descriptions	63
Memory for Detail: Free Recall Questions	68
Memory for Detail: Multiple Choice Questions	73
Lineup Identifications	78
<b>CHAPTER FOUR: EXPERIMENT 2</b>	<b>85</b>
<i>Method</i>	85
Rationale, Aim and Hypotheses	85
Design	86
Participants	87
Measures, Instruments and Materials	87
Physiological Measures	87
Self-reported Anxiety	87
Psychosocial Stressor	87
Research Assistants	88
Evaluators	88
Stimulus Materials	88
Eyewitness Tasks	88
Scoring Materials	89
Procedure	89
Data Analysis	92
Scoring	92
Statistical Analysis	94

---



	8
<i>Results</i>	95
Measures of Stress	95
Heart Rate	95
Skin Conductance Level	95
Autonomic Reactivity	96
Self-reported Anxiety	96
Overview	97
Measures of Memory	98
Target Descriptions	98
Memory for Detail: Free Recall	101
Lineup Identifications	103
<b>CHAPTER FIVE: DISCUSSION</b>	<b>108</b>
Statistics: Summary and Discussion	108
Elicitation of a Stress Response	108
Target Descriptions	110
Interrogative Recall	111
Lineup Identifications	113
Summary	115
Limitations and Directions for Future Research	116
Conclusion	118
<b>REFERENCES</b>	<b>121</b>
<b>APPENDICES</b>	<b>132</b>

---



**LIST OF FIGURES**

Figure 1. Graphic representation of the Yerkes-Dodson Law	19
Figure 2. Graphic representation of the empirical evidence for the Yerkes-Dodson Law	20
Figure 3. Fazy and Hardy's (1988) cusp catastrophe model of anxiety and performance	22
Figure 4. Plot of effects for the relationship between stress and eyewitness memory	29
Figure 5. Percentage differences across conditions for free recall questions (Experiment 1)	69
Figure 6. Graphic representation of the regression of the change in total score for detail recollection (Experiment 1)	72
Figure 7. Free recall vs. multiple choice performance for details of the event (Experiment 1)	75
Figure 8. Free recall vs. multiple choice performance for details of the environment (Experiment 1)	76
Figure 9. Free recall vs. multiple choice performance for total score (Experiment 1)	77

---



## LIST OF TABLES

### EXPERIMENT 1

Table 1. Counterbalanced Variables in the Design of Experiment 1	54
Table 2. Intercorrelations Between Independent Raters of Description Precision	58
Table 3. Intercorrelations Between Independent Raters of Description Informativeness	58
Table 4. State- and Trait-Anxiety: Descriptive Statistics	62
Table 5. Target Description Scores: Descriptive Statistics	64
Table 6. Results of the Dependent-samples <i>t</i> Test for Description Dimensions	65
Table 7. Regression Results for Target Description Dimensions	67
Table 8. Free Recall Question Scores: Descriptive Statistics	68
Table 9. Regression Results for Total Detail Recollection	72
Table 10. Multiple Choice Question Scores for Detail Recollection: Descriptive Statistics	73
Table 11. Free Recall vs. Multiple Choice Question Scores for Detail Recollection: Descriptive Statistics	74
Table 12. Lineup Identifications in the Target-Present Condition: Descriptive Statistics	79
Table 13. Lineup Identifications in the Target-Absent Condition: Descriptive Statistics	79
Table 14. Lineup Accuracy	81
Table 15. Calculation of Differential Accuracy	82
Table 16. Gravity of Error for Lineup Selections	84

### EXPERIMENT 2

Table 17. Counterbalanced Variables in the Design of Experiment 2	89
Table 18. Intercorrelations Between Independent Raters of Description Precision	93
Table 19. Intercorrelations Between Independent Raters of Description Informativeness	93

---



	11
Table 20. State- and Trait-Anxiety: Descriptive Statistics	97
Table 21. Target Description Scores: Descriptive Statistics	98
Table 22. Regression Results for Target Description Dimensions	100
Table 23. Free Recall Scores for Detail Recollection: Descriptive Statistics	101
Table 24. Regression Equation for Total Detail Recollection	103
Table 25. Lineup Identifications in the Target-Present Condition: Descriptive Statistics	104
Table 26. Lineup Identifications in the Target-Absent Condition: Descriptive Statistics	104
Table 27. Lineup Accuracy	105
Table 28. Gravity of Error for Lineup Selections	107

University of Cape Town



## CHAPTER ONE: INTRODUCTION

Eyewitnesses play a pivotal role in the administration of justice. Their testimony can be profoundly constructive when it leads to the conviction of a guilty culprit, but it can also be catastrophic when unreliable. Eyewitness testimony is accorded substantial weight in legal proceedings: Jurors have been known to credit witnesses' accounts of a suspect's guilt when all other evidence points to the contrary (Loftus, 1979). Unreliable eyewitness testimony (specifically erroneous identification) is the most prevalent source of wrongful conviction (Huff, 2004). In more than 75% of DNA exoneration cases in recent years, the false convictions have hinged on the identifications of mistaken eyewitnesses (The Innocence Project, 2008). The social sciences can play a significant role in preventing such miscarriages of justice (McMurtrie, 2005). Research into factors that compromise eyewitness reliability can ultimately help to conserve the resources of the criminal justice system and to protect the freedom of innocent citizens. 'Stress' is one such factor, and is the focus of this thesis.

Though eyewitness testimony carries such significance in legal proceedings, it is inherently unreliable (Wells, Seelau, Rydell, & Luus, 1994). The task of accurately perceiving and recalling a fleeting sequence of events to which the witness may have paid little attention in the first place is highly demanding. This task is thought to be particularly challenging for younger witnesses. For most of legal history, the judiciary has regarded child witnesses with skepticism (Ceci & Bruck, 1993). This wariness has generally been extended to adolescents (a term referring to juveniles "between the onset of puberty and maturity" [WorldNet-Online]). Witnesses younger than eighteen in South Africa are considered children and are subject to the cautionary rule – their testimony "must be scrutinised with great care" by the court (Robinson, 1997, p. 182).

Over the last couple of decades, increasing numbers of adolescents have come into contact with the justice system, whether as perpetrators, victims or witnesses of crime (U.S. Department of Justice, 2006). Youth involvement in crime in America peaked in the mid-1990s, stimulating

---

research on the treatment of juveniles in the legal system (e.g. Grisso, 1997; Scott, Repucci, & Woolard, 1995). This research, for the most part, focused on the competence of adolescents as trial defendants, but not as eyewitnesses. Pozzulo and Lindsay (1998) observe that “if courts are to accept child identification evidence, it is imperative to know at what age children's performance approximates that of adults” (p. 567). A handful of researchers took on this challenge in the 1970s and 1980s, but the literature from that time presents a fragmented picture of adolescent eyewitness performance.

What has been established through empirical research is the fact that adolescents are amply capable of being reliable witnesses in some capacities, but are less trustworthy in others. In the sphere of verbal memory, adolescents are equivalent to adults: Marin, Holmes, Guth and Kovac (1979) found that seventh- and eighth-graders gave reports as accurate as those provided by their college-aged counterparts. Similarly, Cohen and Harnick (1980) found that 12-year-olds could perceive, interpret and describe an event as well as college students. Goodman and Reed (1986) did not include an adolescent sample in their study, but found that children as young as six years could answer objective questions about an event as accurately as adults.

Findings on adolescents' visual memory performance are less convergent, however. Chance and Goldstein (1984) found that face recognition performance improves steadily throughout childhood and levels off around early adolescence (at which point an adult-like hit-rate<sup>1</sup> of 70 – 80% is achieved). Similarly, Marin et al. (1979) found that the seventh- and eighth-graders were as accurate in their lineup identification as college-aged witnesses. In an earlier seminal study, however, Ellis, Shepherd, & Bruce (1973) had asserted that face recognition development continues throughout the adolescent years: Their 17-year-old participants displayed 79% face recognition accuracy, while the 12-year-olds displayed only 72% accuracy. Still other researchers produced evidence of a more complex developmental trend. Carey, Diamond and Woods (1980) found that facial recognition improves until ten years of age, but stabilises or even

---

<sup>1</sup> *Hit rate* refers to the frequency of accurate positive identifications. In a face recognition paradigm, *true hits* occur when participants have actually seen a face before and correctly identify it as familiar. In an eyewitness paradigm, *true hits* occur when the target face is actually present in the lineup and participants correctly identify it.

---

deteriorates in the early adolescent years; it then begins to improve again after 16 years of age. In explanation, Carey et al. (1980) posited that the hormonal changes of puberty affect the functioning of the brain regions associated with face processing. Flin (1980) published a similar finding, except that she found the onset of improved performance to be around 13 years of age (following a temporary decline at 11 and 12 years).

The most recent meta-analysis on witness age and identification accuracy has brought some clarity to the field: Pozzulo and Lindsay (1998) concluded that the onset of adult-like accuracy occurs in early adolescence, confirming the findings of Chance and Goldstein (1984) almost a decade and a half earlier. No such finding emerged for correct rejection rates, however: Adults outperformed adolescents when it came to rejecting culprit-absent lineups (Pozzulo & Lindsay, 1998). In sum, adolescent witnesses are as competent as adults in their powers of observation and recollection, and their ability to provide a reliable narrative. They are also as accurate as adults at making correct identifications in target-present lineups, but they are inferior to adults at making correct rejections from target-absent lineups.

Adolescents have been chosen as the focus of the first experiment in this thesis because “they have been neglected by eyewitness researchers” (Pozzulo & Lindsay, 1998, p. 568). Juveniles in the legal system are treated with equivocation: Their competence to testify is generally not questioned, but their judgment and capacity for rational thought is pervasively regarded as less mature than adults’ (Scott, Reppucci, & Woolard, 1995). Research has shown them to be as successful as adults at observing, remembering and reporting events accurately, and at recognizing faces, but they are still classified as ‘children’ in the South African legal system and are subject to the cautionary rule. With reference to the legal context, Pozzulo and Lindsay (1998) pose the question: “When are teenagers effectively adults?” (p. 568). Rather than tackling this theoretical behemoth in its entirety, this thesis will instead investigate the effect of one critical variable: Does stress affect adolescent eyewitness performance in the same way that it affects adult eyewitness performance?

---

In modern society, *stress* has become something of a blanket construct, encompassing a wide range of responses to aversive situations. In day-to-day usage, we commonly draw on this term to express a state of emotional upheaval, chronic tension (such as professional burnout), or acute pressure (such as an imminent deadline). All of these usages are a far cry from Cannon's original description of *stress* as a biologically adaptive fight-or-flight response, characterised by the activation of the sympathetic nervous system (Miley, 1999). The nebulosity surrounding this term has seeped into the eyewitness literature. In Christianson's (1992) review, for example, *stress* refers to emotional reactions provoked by negative events, ranging from moderate distress to trauma. By contrast, Morgan et al. (2004) use *stress* to refer to the threat of physical assault and the attendant "alterations of neurobiological processes and psychological experiences" (p.266). In this thesis, *stress* is conceptualised as the natural, defensive set of responses of all human beings to an acute threatening situation, characterised by autonomic activation and psychological anxiety. A stressor is any agent which upsets one's 'homeostasis' (Kim & Yoon, 1998), whether through a threat to bodily integrity or a threat to ego (Deffenbacher, 1994). The stress manipulation in the present study—a public speaking task—hinged on the latter. I will be using the term *stress* to refer to an acute state of physiological disequilibrium (as represented by autonomic changes) and psychological disequilibrium (as represented by heightened anxiety). Many eyewitnesses are presumably in such a state when exposed to a perpetrator, because of the sense of threat which pervades most criminal events.

Particularly with regard to adolescent witnesses, whose evidence is already accepted warily in South African courts, it would be useful to know the extent to which heightened stress compromises eyewitness accuracy. In South African courts, the danger inherent in accepting eyewitness testimony is widely acknowledged (Tredoux & Chiroro, 2005). The commonly cited judgment from *R. v. Shekelele and Another* (1953) emphasises the need for careful testing of identification evidence: "An acquaintance with the history of criminal trials reveals that gross injustices are not infrequently done through honest but mistaken identifications. People often resemble each other. Strangers are sometimes mistaken for old acquaintances" (1953 (1) (SA) 636 (T), p. 638). Even if a witness appears "*bona fide* and honest", the court must satisfy itself that the witness is trustworthy and reliable (*S. v. Mehlape*, 1963 (2) (SA) 29 (A), p. 635). In determining the validity of an eyewitness identification, the court considers both characteristics

---

of the witness (such as age, eyesight and possible bias [*S. v. Mputing*, 1960 (1) (SA) 785 (T)] and characteristics of the event (such as visibility, vantage point and proximity to the perpetrator [*S. v. Mputing*], as well as duration of exposure, and prior opportunities for observation [*S. v. Mehlope*]). Certain variables that may have inhibited the witness's ability to form an accurate memory representation in the first place are carefully weighed by the court, particularly where the witness had limited opportunity to view the culprit (*S. v. Molelekwa and Others*, 2007 (JOL) 19040 (T)). Other variables are frequently overlooked, however: Assessments of witness reliability are "wholly unaffected by considerations of shock, panic and the like" (*R. v. T.*, 1958 (2) (SA) 676 (A), p. 397). In *S. v. Mehlope*, for instance, the complainant was awakened to find that he was being pinned down by three men, who assaulted him and made off with his wristwatch and a sum of money. The court acknowledged that the complainant "was awakened to find himself in a most unpleasant and possibly even frightening predicament" (p. 636) but did not take this into consideration when assessing the validity of his identification. Questions of the complainant's degree of acquaintanceship with the appellant and the nature of the viewing conditions were raised, but the effect of stress on the complainant's memory was not considered. Similarly, in *R. v. T.*, much attention was given to the lighting conditions at the crime (the rape of the teenage complainant by a stranger). However, the court did not consider the extreme stress and trauma the complainant must have experienced at the time of her exposure to the perpetrator, despite her statement that "she became panic-stricken and really did not notice very much" (p. 396). It is clear that the impact of heightened stress on eyewitness reliability is not traditionally considered by South African courts. This thesis raises the question of whether it should be. This review will proceed with an exploration of the relationship between stress and memory, considering it from both a neurological and a theoretical angle. I will then present a synthesis of the findings concerning stress and eyewitness memory in adults and children, as well as a discussion of the challenges endemic to the study of stress with a young sample. I will conclude the review with a discussion of various measures of eyewitness performance (person descriptions, interrogative recall and lineup identifications) and the issues pertinent to each measure.

---

## CHAPTER TWO: LITERATURE REVIEW

### Stress and Eyewitness Memory: Background

Until recently, it has been unclear whether stress enhances, impairs, or negligibly affects eyewitness memory (Deffenbacher, Bornstein, Penrod, & McGorty, 2004). An early idea was that stressful events become indelibly engraved in one's mind, and are preserved as particularly vivid and enduring memories (Marshall, 1966, as cited in Brigham, Maass, Martinez, & Whittenberger, 1983). This notion received little empirical support, and was ultimately refuted by the findings of neurological research from the past two decades.

#### *Neurological Research*

The brain secretes high levels of the hormone, cortisol, in response to stress (Lovallo & Thomas, 2000). Implicated in the regulation of cortisol secretion is the hippocampus, which contains more corticosteroid receptors than any other brain region (Lovallo & Thomas, 2000). Though memory function is a distributed process with many neural correlates, the hippocampus is the crucial brain structure associated with memory (Kim & Yoon, 1998; Squire et al., 1992; Wiggs, Weisberg & Martin, 1997). The hippocampal region is involved particularly in the encoding, storage and retrieval of semantic and episodic memory (explicit memory for facts and events respectively) (Squire et al., 1992). Episodic memory is of particular interest in the study of eyewitness recollection, as it refers to an individual's subjective recollection of a specific context or event (Wiggs et al., 1997).

The presence of numerous cortisol receptors in the hippocampus suggests that elevated levels of cortisol secretion could affect memory function (McEwen & Magarinos, 1997). This has been

---

shown empirically with pharmacological manipulation of cortisol levels in both rats and human subjects. Oitzl and de Kloet (1992), for example, found that the administration of corticosteroids to rats interfered with their memory encoding and consolidation when it came to navigating a water maze. In human beings, free cortisol levels seem to be strongly associated with decrements in declarative memory<sup>2</sup>. Kirschbaum, Wolf, May, Wippich and Hellhammer (1996) induced cortisol increases in their subjects either through the 'Trier Social Stress Test' or through oral administration of hydrocortisone. Elevations in cortisol were linked with impaired performance on free- and cued-recall of memorised word lists, spatial navigation and mental rotation tasks. Similarly, Wolkowitz, Reus, and Weingartner (1990) found cognitive deficits in their subjects following the administration of high doses of prednisone. In particular, recognition memory was compromised. When presented with word lists, subjects struggled to distinguish target words they had seen before the drug administration from distractor words. In sum, the chief brain structure associated with memory is clearly vulnerable to the effects of stress, as indicated by findings that "cortisol interacts with hippocampal neurons to induce cognitive deficits" (Lupien & McEwen, 1997, p. 16) It follows that heightened stress could be expected to negatively impact the fidelity of eyewitness memory.

### *Theoretical Framework*

Researchers investigating the relationship between stress and memory performance have traditionally drawn on the Yerkes-Dodson law (1908) as a theoretical explanation of this phenomenon (Deffenbacher et al., 2004). See Figure 1.

---

<sup>2</sup> Declarative memory is the "conscious or voluntary recollection of previous information" (Lupien & McEwen, 1997, p. 16). Together, semantic and episodic memory constitute declarative memory.

---

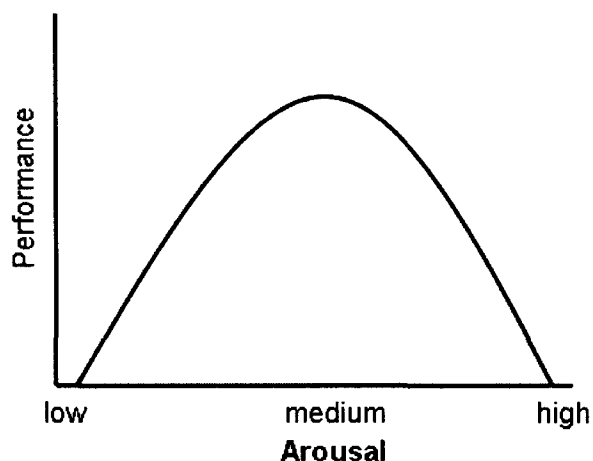


Figure 1. Graphic representation of the Yerkes-Dodson Law

The Yerkes-Dodson Law posits that stress affects performance according to an inverted-U function, with performance improving up to a certain optimal level of stress and then declining with subsequent stress increases. The curve has not been conclusively demonstrated with human beings. The initial rise in the curve (as represented by the light gray portion in Figure 2) appears to have some validity (for example, memory improves with alertness). However, the decline of the curve (as represented by the dark gray portion in Figure 2) has not been empirically proven (Christiansen, 1992). This is particularly difficult to demonstrate, as researchers cannot specify absolute levels of arousal, thus “findings that seem to refute the inverted-U hypothesis can be explained away by claiming that subjects were insufficiently aroused, or too aroused...” (Christiansen, 1992, p. 297).

---

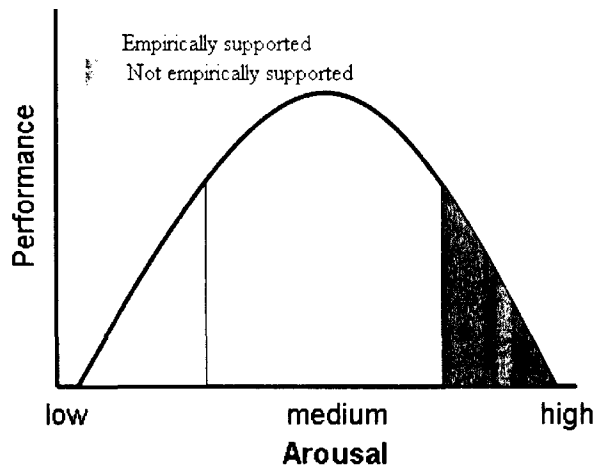


Figure 2. Graphic representation of the empirical evidence for the Yerkes-Dodson Law

Deffenbacher (1994) concludes that the Yerkes-Dodson law is not an appropriate model to explain the relationship between stress and eyewitness memory. Rather than being a simple, unidimensional continuum, eyewitness memory for stressful emotional events is more likely a complex set of interactions among an assortment of variables. This approach is captured in Deffenbacher's (1994) multidimensional theory, in which he integrates previous theorists' explanatory constructs of the relationship between arousal and performance.

Tucker and Williamson's theory of asymmetric control systems (1984) underpins Deffenbacher's integrative model. Tucker and Williamson (1984) proposed that our responses to the demands of the environment are regulated by two distinct neural control systems: the *arousal system* and the *activation system*. The *arousal system* (characterized by right hemisphere dominance and decreased heart rate, blood pressure and muscle tone) provokes an orienting response. *Arousal* allows us to direct our attention towards the most interesting or informative aspect of the stimulus display and thus facilitates memory. The *activation system* (characterized by left hemisphere dominance and increased heart rate, blood pressure and muscle tone) prepares the body for action. *Activation* manifests as a fight-or-flight response (featuring increased cortisol levels). When activation exceeds moderate levels, decrements in memory may occur.

Deffenbacher's integration of this theory is as follows: If the situation is stimulation-based, a state of *arousal-mode dominance* will arise and memory will be enhanced; if the situation is threat-based, a state of *activation-mode dominance* will arise and memory will either show moderate enhancement or dramatic deterioration, depending on the attendant levels of physiological activation and cognitive anxiety.

Deffenbacher's integrative theory also draws on Fazy and Hardy's (1988) catastrophe model (Figure 3) (as cited in Deffenbacher, 1994). This model relies on the interaction between cognitive anxiety and physiological activation to explain patterns of performance. Cognitive anxiety (which manifests as apprehension, worry or general psychological distress) is a splitting factor, modulating the relationship between physiological arousal and performance (Hardy & Parfitt, 1991). When cognitive anxiety is held constant at a low level, the relationship between anxiety and performance conforms to the inverted-U hypothesis (as shown by the back face of Figure 3). When cognitive anxiety is elevated, however, performance may improve moderately with small increases in physiological arousal or plummet dramatically with more extreme physiological activation. Since performance declines catastrophically (rather than deteriorating gradually, as in the Yerkes-Dodson hypothesis) it cannot be restored to optimum levels with small reductions in anxiety (Deffenbacher, 1994). From the catastrophe model, Deffenbacher (1994) borrowed the idea that anxiety has a physiological and a cognitive component, with variations in the cognitive component determining the effect of physiological activation on performance.

---

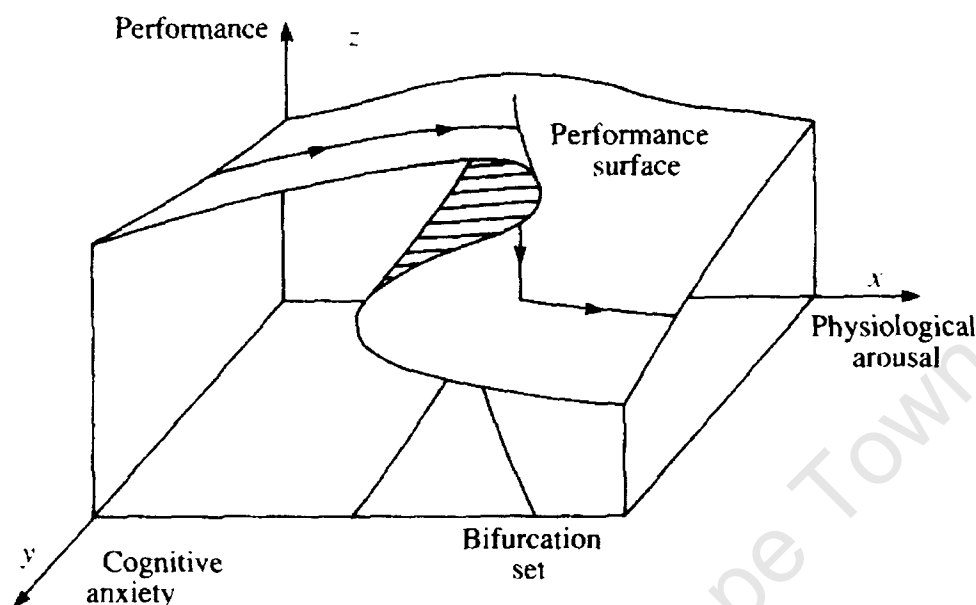


Figure 3. Fazez and Hardy's (1988) cusp catastrophe model of anxiety and performance (from Hardy and Parfitt, 1991)

Deffenbacher (1994) suggests that the memory performance of an individual in a state of activation mode dominance can be predicted based on the balance between his levels of cognitive anxiety and physical activation. When cognitive anxiety is low (as in most laboratory studies), we can expect memory performance to follow the traditional inverted-U function as physiological activation increases. When cognitive anxiety is high, however, we can expect memory performance to improve with small increases in physiological activation, but to plummet catastrophically with moderate increases in physiological activation (Hardy & Parfitt, 1991).

Finally, Deffenbacher (1994) suggests (with reference to Hockey's state control model of 1986) that the stressed individual's memory performance is affected by appraisal of environmental demands and by coping resources. In sum, Deffenbacher's integrative theory posits that memory performance in response to stress is an intricate interplay among a number of factors: The

individual's mode of attentional regulation (arousal vs. activation), the prevailing state of cognitive anxiety, degree of physiological activation, and coping style. Christiansen (1992) suggests the contribution of still other variables, such as the nature of the event (emotional vs. neutral), the type of information the individual is required to recall (central details vs. peripheral details), time of retrieval (immediate vs. delayed) and conditions of retrieval (mood-states, for instance). In the present study, I set out to test the stress-performance relationship in eyewitnesses. This was accomplished using a public speaking task as a psychosocial stressor. I tested participants' memory performance for an eyewitness event in which they took part while in a state of anticipatory stress (immediately before delivering a speech in Experiment 1, or before being interviewed by a panel of evaluators in Experiment 2). Based on Deffenbacher's theoretical formulation, I expected participants to exhibit memory decrements in response to the stress manipulation: They would presumably be in a state of activation mode dominance with heightened cognitive anxiety and moderate physiological activation (induced by the threat to ego inherent in the public speaking task).

### Stress and Eyewitness Memory: Empirical Findings

The effect of stress on eyewitness memory has been the topic of wide research with adults and limited research with children, but adolescent eyewitnesses have received no attention. In the absence of research on juveniles, I will discuss the findings from adult studies followed by the findings from child studies. This section will conclude with a discussion of the methodological issues and ethical hurdles confronting researchers of stress effects in a juvenile sample.

#### *Adult Eyewitnesses*

Findings from adult studies have unanimously indicated that stress degrades eyewitness memory. Though effect sizes are not typically provided, they are of great utility in a review of the

---

literature, as they give an indication of whether the difference between two groups or conditions was meaningful (Howell, 2004). I have therefore calculated and reported the effects for the studies discussed below as *Cohen's d*. According to Cohen's rule of thumb, 0.2 can be considered a small effect, 0.5 a medium effect, and 0.8 a large effect (Cohen, 1988).

In an early investigation of the effect of 'type of incident' on adult eyewitness memory, Clifford and Hollin (1981) manipulated violence as a variable which renders an event more or less stressful. Half of their participants watched a non-violent event (a group of men approach a woman for directions), while the other half watched the same scene unfold in a violent way (the same woman is mugged by the same group of men). The authors found poorer eyewitness memory performance for the violent than for the non-violent conditions, suggesting a negative relationship between stress and memory ( $d = -0.27$ ). In a study yielding similar results, Brigham et al. (1983) used electric shocks as a stimulus to induce stress in their sample of undergraduates. Participants in the high-arousal condition were subjected to a series of electric shocks at unpredictable intervals, while subjects in the moderate-arousal condition received only one electric shock with the knowledge that there were no more to come. Participants were exposed to 25 slides of target faces following the single shock in the moderate-arousal condition, and interspersed with random shocks in the high-arousal condition. After the shock phase was completed, participants from both groups were required to distinguish the target faces from unknown distractor faces. The authors found that recognition accuracy was poorer for the high-arousal subjects than for the moderate-arousal subjects ( $d = -0.68$ ). Using a very different approach to induce arousal, MacLin, MacLin and Malpass (2001) paired random unknown faces with images from the International Affective Picture System (IAPS). Half of the participants were exposed to faces paired with neutral images, while the other half saw faces paired with negatively arousing images which have been shown to provoke physiological symptoms of arousal, such as heart rate and skin conductance changes (Ito, Cacioppo, & Lang, 1988, as cited in MacLin et al., 2001). Participants from the high arousal condition showed diminished face recognition accuracy ( $d = -0.96$ ). Though Clifford and Hollin (1981) used an eyewitness research paradigm where Brigham et al. (1983) and MacLin et al. (2001) favoured a face recognition

---

research paradigm, all three studies produced congruent findings: Negatively arousing stimuli impair memory.

Two studies using live-event methodologies have lent support to Deffenbacher's integrated theory (1994). In the first of these, Bothwell, Brigham and Pigott (1987, as cited in Deffenbacher, 1994) assigned participants to a high, moderate or low-arousal condition and used the conditioned aversive stimulus of a syringe to induce anxiety. In the high-arousal condition, participants were ushered into a room where a syringe lay on a table in an unopened packet; in the moderate-arousal condition, the syringe packet on the table was empty; and in the low-arousal condition, there was nothing on the table. Participants were later required to identify the confederate who had led them into the room. In addition to using multiple physiological measures of arousal, Bothwell et al. (1987) took into account participants' levels of cognitive anxiety. This was measured with the Eysenck Personality Inventory (where the authors equated the presence of neuroticism with high levels of cognitive anxiety). High cognitive anxiety exerted a strong effect on participants' identification performance. All subjects in the moderate- and high-arousal conditions were in a state of activation mode dominance (as indicated by their increased heart rates), but only those with high cognitive anxiety showed a catastrophic decline in identification accuracy ( $d = -0.74$ ). Consonant with Deffenbacher's hypothesis, this study showed that when individuals are experiencing heightened physiological arousal, increased cognitive anxiety can lead to dramatic decrements in eyewitness memory.

In the second study using a live-event methodology, Peters (1988) examined the effect of naturally occurring stress on adult eyewitness accuracy. He compared subjects' memory for a nurse, from whom they received an inoculation, with their memory for the researcher, to whom they were exposed after the inoculation. Subjects provided less accurate descriptions and demonstrated poorer recognition of the nurse, whom they encountered in a state of physiological arousal, than for the researcher (41% vs. 66%,  $d = -0.64$ ). Participants were presumably in a state of activation mode dominance and cognitive anxiety when they met the nurse, seeing as they were facing the prospect of physical pain (the injection). In support of Deffenbacher's theory,

---

heightened physiological arousal (average heart rate elevation of 17.1 beats/minute) was associated with substantial deterioration in eyewitness performance (a 23% decline in accuracy between the stress and non-stress conditions).

The most recent study of adult eyewitness performance following heightened stress at event-time was conducted by Morgan et al. (2004). The participants in this study were military personnel in the experiential phase of their survival school training. The subjects were confined in a Prisoner of War Camp, where they were placed in isolation and then interrogated, both in a high- and low- intensity environment. All subjects participated in a high-intensity interrogation, which involved real physical confrontation by the interrogators, and a low-intensity interrogation, which involved verbal trickery, but no threat of physical violence. When called upon to select their interrogators from a lineup, the majority of subjects performed more accurately in identifying their low-stress than their high-stress interrogators. For simultaneous photo-spread lineups (the mode of presentation used in the current study), a large negative effect was found ( $d = -1.13$ ).

Some studies have revealed the opposite effect of stress on eyewitness memory, such as an investigation into the eyewitness performance of police trainees reported by Yuille, Davies, Gibling, Marxsen & Porter (1994). This study involved a role-playing event, in which one third of the trainees participated while the other two thirds observed. The non-stressful condition involved the questioning of a polite and cooperative suspect, while the stressful condition involved the questioning of a belligerent and uncooperative suspect. Both the participants and the observers recalled details of the stressful role-playing event with more accuracy than they recalled details of the non-stressful one ( $d = 0.75$ ). Deffenbacher et al. (2004) attribute such findings to the elicitation of an orienting response in participants rather than a defensive one. Indeed, the element of personal threat necessary to provoke activation mode dominance was notably absent from Yuille et al.'s role-playing manipulation (1994), even for those who participated in it. From those studies which *have* successfully elicited a defensive response from participants (confirmed by physiological validity checks in most instances), a clear picture

---

emerges: Heightened stress at the time of exposure to an unknown individual impairs subsequent memory for that individual. This finding was confirmed in the most recent meta-analysis of the field (Deffenbacher et al., 2004,  $d = -0.31$ ).

### *Child Eyewitnesses*

While the findings summarised above have eliminated much of the dissent concerning stress and eyewitness performance in adults, relatively little is known about the effects of stress on children's eyewitness accuracy. A couple of studies have mirrored the results of research on adults, revealing a negative relationship between stress and memory in children, but others have found no difference in memory performance for children across stressful and non-stressful events.

Since researchers cannot ethically induce stress in their child participants, they have turned to naturally occurring, anxiety-provoking events in which to study child witnesses. The seminal study of naturally occurring stress and child witness accuracy was conducted by Peters (1987). Children's visits to the dentist—naturally-occurring, meaningful events where anxiety was likely to be high—constituted the stressful event, while follow-up visits by the researcher to the children's homes constituted the neutral event. As in the study by Morgan et al. (2004) and Peters' adult study (1988), a repeated measures design was used. For the stressful event, the dentist and oral hygienist served as target individuals whom the subjects were later required to identify; the researcher served the same purpose for the neutral event. In target present lineups, identification of the researcher from the less stressful setting was superior to recognition of the dentist from the more stressful setting ( $d = -0.75$ ).

In more recent studies, Merritt, Ornstein and Spicker (1994), and Shrimpton, Oates and Hayes (1998) both constructed their experimental samples from children undergoing medical

---

procedures (voiding cystourethrogram and venipuncture respectively). Merritt et al. (1994) found that stress inhibits eyewitness memory in children, whereas Shrimpton et al. (1998) found no relationship between stress and memory. Using similar methodology, Lindberg, Jones, Collard and Thomas (2001) compared the eyewitness memory of children who experienced an inoculation with the memory of matched controls who watched another child being inoculated on video. The children who experienced stress directly identified the inoculation nurse with more accuracy after a month's delay than those who experienced stress vicariously ( $d = 0.54$ ). The association between heightened stress and improved memory in this study, however, can be convincingly explained by the fact that the event was personally experienced and therefore more salient for those children who received inoculations. In sum, the few studies done on the effect of stress on eyewitness memory in children have yielded a disparate set of results, and no solid conclusions can be drawn concerning this age-group.

To conclude this synthesis of the literature on stress and eyewitness memory, the effect sizes from the three studies most similar to the present one have been depicted graphically in Figure 4. Peters (1987), Peters (1988) and Morgan et al. (2004) employed repeated-measures designs and live-event methodologies and produced remarkably congruent results. All of the studies depicted in Figure 4 show a negative effect, indicating that stress impairs eyewitness accuracy in target-present lineups<sup>3</sup>. The mean effect size for these three studies is also depicted in Figure 4 ( $d = -0.77$ ,  $CI_{.95} = -1.09, -0.46$ ). Only the confidence interval for Peters' child study (1987) contains the null value of zero, which can be explained by the fact that this study had a relatively small sample size ( $N = 35$ ) and thus lacked power. Given the methodological resemblance of the present study to the three represented in Figure 4, it seems reasonable to expect that it will produce similar findings.

---

<sup>3</sup> Effect sizes were calculated using the Probit Method for dichotomous proportions.

---

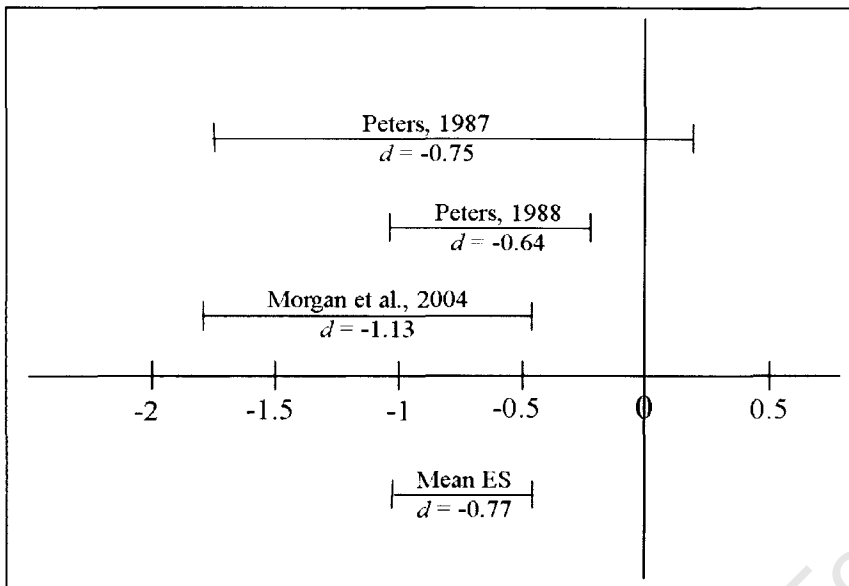


Figure 4. Plot of effects for the relationship between stress and eyewitness memory. Horizontal bars denote 95% confidence intervals.

#### *Methodological Issues in the Study of Stress*

The paucity of research concerning adolescent eyewitnesses can be largely attributed to the ethical constraints that hamper all investigators in the field of stress research: Methods which may result in distress or psychological harm to participants are prohibited (American Psychological Association, 1990, as cited in Christiansen, 1992). Needless to say, the ethical restrictions which apply to adult participants are amplified when it comes to juvenile populations, and subjecting young participants to some of the methods employed with adults (such as violent films or electric shocks) would be objectionable on ethical grounds.

As discussed previously, the construct of ‘stress’ in this thesis refers to a state in which “physiological and psychological homeostasis” is perturbed (Kim and Yoon, 1998, p. 505). To create heightened stress, a situation must involve an element of personal threat, be it to bodily integrity or ego. Tasks involving vigilance, escape, avoidance or pressure are also known to provoke stress responses (Deffenbacher, 1994).

Researchers have used a diverse array of methods to induce stress in their subjects. Indeed, the equivocal findings in the literature on stress and child eyewitness memory may be the result of varying operationalisations of the stress construct (Peters, 1987). A widespread method of eliciting stress involves the exposure of subjects to a staged crime, most often either as a slideshow (e.g. Dekle, Beale, Elliot & Huneycutt, 1996; Parker & Ryan, 1993) or as a videotaped incident (e.g. Dysart & Lindsay, 2001; Beal et al., 1995). Perhaps the chief limitation of such studies is the absence of the element of personal 'threat' or 'alarm' which would naturally be present in real life-threatening situations. Because they cannot ethically emulate the stress levels of extra-laboratory crime scenes, many studies manage only to provoke in their participants an orienting response (represented by arousal mode dominance) rather than a defensive stress response (represented by activation mode dominance). Thus, certain studies which aim to determine the effects of stress on eyewitness memory do not elicit the fight-or-flight response which would most likely be experienced by witnesses to crime. This diminishes the forensic validity of their findings, because it is precisely this state of heightened physiological activation that is of ecological interest to eyewitness researchers.

To enhance forensic validity, the presence of a stress response must be verified by evidence of arousal. Researchers employ behavioural indices (Shrimpton et al., 1998), and subjective rating scales (Morgan et al., 2004; Peters, 1987) in studying the effects of stress on memory, but the usefulness of such measures is limited in the absence of physiological data. The individual in a state of activation mode dominance, or heightened stress, will exhibit elevated stress hormones, accelerated heart rate, increased muscle tone, raised blood pressure and increased skin conductance (Cacioppo, Tassinary & Berntson, 2000). Previous studies with adults have used measures of skin conductance response (Clifford and Hollin 1981; Brigham et al., 1983), finger pulse volume response (Brigham et al., 1983), and heart rate (Brigham et al., 1983; Peters, 1988) to verify that participants are in a state of stress. With the exception of measures of salivary cortisol levels in the study by Merritt et al. (1994), however, physiological measures of stress have been conspicuously absent in the research on juvenile witnesses. Physiological measures were employed as validity checks in this study, to verify that participants were truly in a state of activation mode dominance.

---

## Measures of Eyewitness Memory

Witnesses to a crime are confronted with multiple tasks when they become embroiled in legal proceedings: They must provide the police with a factual account of the event, they must describe the perpetrator, they may be required to identify the perpetrator, and if the case goes to trial, they may be called to testify in court. Each of these aspects of eyewitness testimony has received attention from social science researchers, but the three particular aspects addressed in this thesis are: person descriptions, interrogative recall of event details and lineup identifications.

### *Person Descriptions*

Following a criminal event, obtaining the witness's description of the culprit is one of the primary steps in the justice process: Police will begin their investigation by searching for a suspect who matches the description. Providing the police with this description is perhaps the most difficult task facing the eyewitness. Making a lineup identification requires the witness to match a visual stimulus with a memory representation of that visual stimulus – a relatively simple cognitive task. By contrast, providing a target description requires the witness to switch modalities, using the verbal modality to represent a stimulus that was perceived visually (Sporer, 1996). A further reason that providing a detailed and accurate description is challenging is that the eyewitness encoded the face holistically, but is required to describe it in a fragmentary fashion, characterising the features individually (Sporer, 1996).

Presumably resulting from the cognitive complexity of the task, witnesses generally provide person descriptions that are vague and incomplete (Meissner, Sporer, & Schooler, 2007; Sporer, 1996; Van Koppen & Lochun, 1997). In one of the earliest archival studies of person descriptions, Kuehn (1974) found that witnesses can relay a general impression of the culprit, but are hard-pressed to provide specific details (such as eye colour). In reviewing 100 person descriptions from police archives, Kuehn (1974) observed that witnesses typically address only obvious features such as gender, age, height, build, race, weight, complexion and hair colour (in

---

this order). Witnesses in Kuehn's study provided a rather meager average of 7.2 descriptors. In a more recent archival study, Van Koppen and Lochun (1997) observed similar performance for witnesses describing robbers across 431 cases: The median number of descriptors reported was eight. Consistent with findings on the saliency of various facial features, witnesses typically provide more details for features in the upper half of the face. Hair and eyes are mentioned with greater frequency than mouth, nose or chin (Van Koppen & Lochun, 1997).

Though witnesses tend to provide fewer descriptors than one might expect, they are fairly accurate in those they do provide. Van Koppen and Lochun (1997) found that the median value for accuracy was 72% in the 1650 verifiable descriptions they reviewed. In a field study of great forensic value, Yuille and Cutshall (1989) found the accuracy of 13 actual witnesses to a single shooting incident to be remarkably high, even after a long retention interval (approximately five months between the event and the research interview). Yuille and Cutshall (1989) suggest that their witnesses performed with such a high degree of accuracy because of the spectacular nature of the incident, the presence of life-or-death consequences and the personal involvement in the event for some of the witnesses. Wagstaff et al. (2003) observe that field and archival studies generally yield higher rates of accuracy than laboratory studies. They attribute this to the grave consequences associated with eyewitness inaccuracies in real-life situations, leading actual eyewitnesses to be more tentative than mock eyewitnesses.

Yuille and Cutshall (1989) found enhanced accuracy in the recollections of witnesses who reported heightened stress at the time of the event. The five witnesses whose stress ratings were at the upper end of the stress rating scale were almost entirely accurate in their reporting of the event. The remaining witnesses reported significantly lower stress levels during the event and demonstrated inferior accuracy. Sporer (1992) found a similar result in his archival analysis of person descriptions. He categorized witnesses into three groups, each of which was associated with more or less stress: victims, bystanders (who had a role in the event but were not victims), and incidental witnesses (who were able to provide the police with information concerning the culprit, but were not involved directly in the event). The incidental witnesses (who were presumably unaroused at the time of encountering the culprit) were less precise than the victims or bystanders in their descriptions of the culprit. This result suggests that heightened stress

---

improves eyewitness memory, at least when it comes to describing the perpetrator. Finally, in a more recent archival analysis, Wagstaff et al. (2003) found that stress—operationalised as high levels of violence—had no effect on description performance (except to enhance the accuracy concerning the offender’s hair colour). These findings contradict the general consensus that stress impairs eyewitness memory (Deffenbacher et al., 2004). For Yuille and Cutshall’s study (1989) and Sporer’s analysis (1996), this can perhaps be explained by the presence of a notable confound: Higher levels of stress were associated with participation in the event. It is therefore possible that direct involvement made the incident more salient for these witnesses, and promoted enhanced retention (Yuille & Cutshall, 1989). In any event, we cannot validate the degree of stress experienced by witnesses in these studies, as they were reported after the fact by the witnesses themselves or coded retroactively by the researchers. Furthermore, authors of archival studies cannot establish the accuracy of witnesses’ descriptions with any degree of certainty, so conclusions from archival data should be drawn with caution (Meissner et al., 2007).

One archival study *has* found evidence of inferior eyewitness performance following stress at encoding: Van Koppen and Lochun (1997) used a threat index to assess the impact of stress in their review of cases involving robberies of commercial targets (such as banks and shops). The authors coded for the level of threat that witnesses might have experienced, assigning three points when the robbery was witnessed directly, two points when the escape of the robber was witnessed, and one point when the preparations for the robbery were witnessed. Extra points were added when the criminal event involved death or injury, weapons, violence, hostage-taking and a number of other stressful elements. The authors found that greater threat was associated with less complete descriptions (i.e. lower numbers of descriptors provided). Similarly, the findings from one laboratory and one field study have indicated that heightened stress at encoding leads to impoverished person descriptions. Clifford and Hollin (1981) found that participants who watched the non-violent video provided more accurate descriptions of the principal character than those who watched the violent version. Peters (1988) found that subjects provided more accurate descriptions of the researcher whom they had encountered in a state of non-arousal than of the nurse, with whom they had interacted under stressful circumstances (receiving an inoculation). Given the primary hypothesis of this thesis—that stress adversely

---

affects eyewitness memory—and the findings from the study most similar to mine in design (Peters, 1988), I predict that participants will provide inferior descriptions of the target they encounter in a state of heightened stress.

Worth final mention is the impact of age on the quality of witnesses' culprit descriptions. Barring a recent study by Pozzulo and Warren (2003), very little research has been directed at the performance of juvenile witnesses on person description measures. Children tend to provide statements that are briefer and less detailed than those of adults (Marin et al., 1979) – this may be a function of limited language abilities or limited perception and encoding of appearance (Pozzulo & Warren, 2003). Until recently it was unknown whether this finding applied to adolescents as well. Pozzulo And Warren (2003) observed the completeness and accuracy of juveniles' (10- to 14-year olds) target descriptions relative to those of young adults (18- and 25-year olds). Both age groups were exposed to a videotaped target giving a presentation in Experiment 1, and to a live target giving a presentation in Experiment 2. In accord with the findings on young children, the adolescents provided fewer descriptors than the young adults, but the age groups were equivalent in terms of overall accuracy. The younger witnesses struggled to provide accurate estimates of height and weight, and had difficulty characterising facial features (perhaps because they lacked the more sophisticated vocabularies available to the adults) (Pozzulo & Warren, 2003). Of interest was the detection of an in-group age bias: The 18- to 25-year olds were superior to the adolescents in describing the first target, who was also a young adult; by contrast, the 18- to 25-year olds were as poor as the adolescents in describing the second target, who was their elder by half a generation.

In the present study, both the adolescent participants in Experiment 1 and the young adult participants in Experiment 2 were required to provide descriptions of a target whom they had met in a state of anticipatory stress. These descriptions were compared with descriptions they provided of a different target whom they had met under non-stressful circumstances. Based on the findings from field and laboratory studies, I expected the stress manipulation to adversely affect the quality of the target descriptions provided by both samples. In light of the above findings, I expected that the adolescent witnesses in this study would perform generally poorly

---

on the target description measure. The university sample was expected to perform better, both because of their maturity and the fact that the targets were also young adults.

### *Interrogative Recall*

While police will require a narrative account of the criminal event from any eyewitness, it is also likely that they will question the witness to extract more information. Child studies typically employ interrogative recall as a measure of eyewitness memory (Merritt et al. 1994; Shrimpton et al., 1998) as it is known to elicit more information than narrative recall, particularly with young witnesses<sup>4</sup> (Goodman & Reed, 1986; Marin et al., 1979; Cassel & Bjorklund, 1995). In an experimental situation, interrogative recall can be investigated using two types of memory tests: a test of free recall (such as providing a blank space in which to fill in the answer) or a test of recognition (such as providing multiple choice options). Like previous researchers studying interrogative recall (e.g. Cassel & Bjorklund, 1995; Heuer & Reisberg, 1990; Libkuman, Nichols-Whitehead, Griffith, & Thomas, 1999; Loftus & Burns, 1982), I chose to employ both types of memory test in this study. The reason for this is two-fold: Firstly, the inclusion of a forced-choice measure (e.g. multiple choice questions) assesses memory in a more 'fine-grained way' than free recall (Heuer & Reisberg, 1990), eliminating the potential for confabulation; secondly, the inclusion of a recognition measure as well as a recall measure permits inferences about whether poor memory performance is the result of faulty encoding or faulty retrieval.

Police will question witnesses concerning the central details of the event, but also the peripheral details. In rare cases, witnesses' memories for peripheral details may be the key to solving the crime. It is more likely, however, that administrators of justice will question witnesses about peripheral details in order to assess their credibility (Cassel & Bjorklund, 1995). A negative correlation seems to exist between identification accuracy and memory for peripheral details

---

<sup>4</sup> Juvenile witnesses have generally proved more suggestible than adult witnesses when confronted with leading questions (Bruck & Ceci, 1999), but this is not one of the topics under investigation in this thesis. The questions posed to participants in the present study were closed-ended (requiring participants to provide or select a specific simple piece of information) and designed to be objective, so that questioning techniques could not be assumed to compromise participants' accuracy.

---

(Wells & Leippe, 1981; Cutler, Penrod & Martens, 1987). Wells and Leippe (1981) found that mock jurors tended to discredit witnesses who recalled fewer peripheral details about the event. Cutler et al.'s findings (1987) indicated that jurors should in fact put *more* faith in the identifications of witnesses who recall fewer peripheral details.

There is little research on how stress affects eyewitness memory for central and peripheral details of an event, but the effect of *negative emotionality* has been thoroughly investigated. The phenomenon of 'flashbulb memories' (where the trivial aspects of a highly emotional event are recalled in exquisite detail) suggests that emotion might facilitate memory for all types of detail. This notion has received some support (e.g. Heuer & Reisberg, 1990; Libkuman et al., 1999), but the bulk of the research literature indicates that heightened emotional content facilitates memory only for central details, while memory for peripheral details is inhibited<sup>5</sup>.

Easterbrook (1959) first proposed that emotional arousal narrows the focus of one's attention, limiting the number of cues to which one attends. Following from this is the idea that one's memory for central details of an emotional event should be enhanced (because central details receive the most attention), while memories for peripheral details should be impoverished. This hypothesis has been well supported (e.g. Burke, Heuer, & Reisberg, 1992; Christianson, Loftus, Hoffman, & Loftus, 1991; Clifford & Hollin, 1981; Loftus & Burns, 1982; Safer, Christianson, Autry, & Osterlund, 1998). Loftus and Burns (1982), for example, exposed two groups of participants to a film clip of a bank robbery; for one group, the robbery ended non-violently, with the bank manager reassuring the bank patrons; for the other group, the film culminated in a violent incident, with a young boy being shot in the face. Participants who watched the violent, 'mentally shocking' version showed substantial memory impairment for peripheral details of the film (such as the number on the football jersey of one of the young bystanders in the video). Burke, Heuer and Reisberg (1992) produced a similar finding: They borrowed their experimental protocol from Heuer and Reisberg (1990) and exposed participants to a series of slides depicting

---

<sup>5</sup> Different findings may arise from different classifications of to-be-remembered material (Burke, Heuer & Reisberg, 1992). Heuer and Reisberg (1990) drew a conceptual distinction between plot-relevant and plot-irrelevant material, identifying the former as 'central details' and the latter as 'peripheral details'. Other authors have made a perceptual/spatial distinction, classifying any detail associated with the main characters as 'central', including, for example, the target figure's clothing (Christianson & Loftus, 1991). In the studies presented in this thesis, the live event did not involve a prominent narrative and therefore did not lend itself to the use of plot-related distinctions. In light of this, Christianson and Loftus' (1991) classification system was favoured.

either a neutral story or an arousing one. In the former scenario, a boy and his mother go to visit his father who is a mechanic; in the latter version, they go to visit his father who is a surgeon. In the arousing version, the slides of the father working on an engine are replaced by graphic images of him carrying out an operation. Participants who viewed the arousing version demonstrated enhanced memory for details associated with the event's center, but diminished retention of all other details. Easterbrook's claim also gained support from studies on the 'weapon focus' effect, where witnesses focus on the perpetrator's weapon to the exclusion of other details and subsequently perform poorly on eyewitness tasks (such as perpetrator descriptions and lineup identifications)(Stebly, 1992).

The findings described above are out of keeping with the dominant findings concerning the effect of stress on memory. Stress certainly seems to impair adult eyewitness memory (Deffenbacher et al., 2004), so why does negative emotionality *facilitate* memory for central details? With the knowledge that stress has a negative impact on eyewitness memory, we would expect memory for both central and peripheral details to be inhibited. Deffenbacher et al. (2004) postulated that the above mentioned studies found enhanced memory for central details because they provoked orienting responses in their participants, rather than defensive responses. The experiments discussed above comprised of tasks that required "simple perceptual intake" (Deffenbacher et al., 2004, p. 688), such as watching a film or slide show. Thus the element of threat (to bodily integrity or ego) was absent, and the negative emotionality generated by these manipulations cannot be equated with a state of true stress (Deffenbacher et al., 2004). The studies for this thesis were specifically designed to manipulate participants' stress levels and to generate a defensive response. It would thus be inappropriate to make predictions based solely on the findings of studies which failed to do just this. Instead, I applied the overarching hypothesis for the study – that stress will impair eyewitness memory performance – to this aspect of eyewitness memory, but with an adjustment informed by Easterbrook's claim: I predicted that memory for both central and peripheral details would be impaired in the stress condition, but that impairment for peripheral details would be more extreme.

---

### *Lineup Identifications*

In South Africa, the lineup (identification parade) is the most accepted safeguard against eyewitness misidentification (Tredoux, 1996), but identification errors still occur. Prior to trial, the witness will identify the accused from a lineup of eight members. The seven foils should bear enough of a physical resemblance to the suspect so that he does not stand out, but the similarity should not be so great that the suspect blends in (Malpass, Tredoux, and McQuiston-Surrett, 2007). Though the requirement for physical similarity is somewhat nebulous, it typically extends to height, weight, clothing and general appearance (Malpass et al., 2007). When the suspect stands out, witnesses might make errors of commission (whereby an innocent suspect is positively misidentified and incriminated, allowing a guilty suspect to go free). When the suspect blends in, witnesses are more likely to commit errors of omission (the identification of a foil) (Dysart & Lindsay, 2001). The latter are 'known' errors (in that foils are known to be innocent), and are therefore not as legally grave as the former. Foil identifications do, however, have adverse consequences, as they undermine the credibility of the witness (Goodman & Reed, 1986).

Culprit-presence versus culprit-absence has been shown to have drastic effects on recognition accuracy, particularly where young witnesses are concerned (Dekle et al., 1996; Parker & Ryan, 1993; Goodman & Reed, 1986; Dysart & Lindsay, 2001; Pozzulo & Lindsay, 1999; Steblay, Dysart, Fulero & Lindsay, 2001; Gross & Hayne, 1996; Beal, Schmitt & Dekle, 1995; Lindsay, Pozzulo, Craig, Lee & Corber, 1997). Children may be almost as competent as adults at identifying an individual from a culprit-present lineup, but their performance plummets when they are presented with a target-absent lineup. Children are much less likely to correctly reject a target-absent lineup than adults, thus committing significantly more false positive errors (Pozzulo & Lindsay, 1998). This finding seems true also of adolescents, who are as reliable as adults when the target is present, but much less so when the target is absent (Pozzulo & Lindsay, 1998). Given that police are generally not aware when they have an innocent suspect, the presence or absence of the culprit in the lineup is not entirely within their jurisdiction. What is

---

malleable, however, is the presentation of the lineup, such that it minimises the probability of an innocent suspect being identified if the lineup does not contain the true culprit.

Lineup presentation is an example of a system variable – a factor affecting eyewitness accuracy that is within the control of the criminal justice system (Wells & Olson, 2003). System variable research is critical in the investigation of eyewitness testimony, because it uncovers systematic weaknesses in the criminal justice system which can be changed to prevent misidentification. Though system variables are not the focus of this thesis, it is necessary to be familiar with the current state of knowledge concerning them for methodological reasons. Lineup presentation and lineup construction will be discussed as the two cardinal system variables to which the criminal justice system (and eyewitness researchers) must attend.

The lineup presentation method most widely used in Anglo-American legal systems and in the eyewitness research paradigm is the simultaneous lineup (where the lineup members are viewed at the same time) (McQuiston-Surrett, Malpass & Tredoux, 2006). Simultaneous lineup presentations have been criticised for encouraging witnesses to make ‘relative judgment’ decisions (Wells, 1984, as cited in Wells & Olson, 2003). According to the Relative Judgment Processes Theory (Wells, 1984, as cited in Wells & Olson, 2003), witnesses identify the lineup member who, relative to the other members of the lineup, looks most like the witness’s memory representation of the culprit. This strategy is not problematic when the culprit is present in the lineup, but it leads to more false positive identifications in culprit-absent lineups, because there will always be someone who looks more like the culprit than the others. Considerable attention has been directed at investigating the effectiveness of sequential lineups as an alternative (where the lineup members are viewed consecutively). In sequential lineups, the presentation of one member at a time encourages witnesses to make absolute judgments (‘is this the culprit or not?’) (Wells & Olson, 2003). A handful of studies (e.g. Parker & Ryan, 1993; Lindsay et al., 1997; Dysart & Lindsay, 2001; Steblay et al., 2001) have converged on the point that sequential lineups reduce the number of false positives and are therefore preferable. However, McQuiston-Surrett et al. (2006) have suggested that the sequential method superiority emerging from recent

---

research is a function of variations in study methodology and that advocacy of the sequential lineup presentation may be premature. Thus, neither lineup presentation has unequivocally emerged as superior.

In addition to manipulating lineup presentation, the criminal justice system can strive to reduce misidentifications by paying close attention to lineup construction. First of all, *lineup size* is of critical importance. This refers to how well the fillers serve as plausible alternatives to the suspect (Malpass et al., 2007). A filler who is not a viable alternative to the suspect serves no function in the lineup. If only some of the fillers are plausible alternatives, then the lineup is, in effect, smaller than its nominal size (total number of lineup members) (Malpass et al., 2007). This diminishes lineup fairness, because the likelihood that the suspect could be chosen by chance increases substantially. Secondly, *lineup bias* is a system variable which significantly impacts on the fairness of an eyewitness identification. Lineup bias is the degree to which the suspect stands out relative to the other members of the lineup (Malpass et al., 2007). A lineup can be biased towards the suspect, in which case errors of commission are likely to occur, or biased away from the suspect, in which case errors of omission are likely to occur. In a well-constructed lineup, each member should have an equal chance of being selected based on the witness's description (Dekle et al., 1996). The fairness of a lineup (as indicated by lineup size and lineup bias) can be evaluated using a mock witness procedure (Malpass et al., 2007). If mock witnesses who have never encountered the suspect are able to select him from a lineup at a rate above or below chance (either based on a verbal description or no information at all), then the lineup is clearly unfair. Lineup size and lineup bias are interrelated but distinct concepts which should be evaluated and documented when conducting eyewitness research (Malpass et al., 2007).

In contrast to system variables, estimator variables refer to elements beyond the control of the criminal justice system which can affect a witnesses' identification accuracy, such as the circumstances surrounding a crime (Wells & Olson, 2003). Certain factors may have inhibited the witness's ability to form a mental representation of the culprit in the first place, in which

---

case, system variable manipulation will not necessarily produce an accurate identification. Estimator variables can impact significantly on a witness's ability to make an accurate identification judgment. Knowing the effects of estimator variables can inform the court of how much weight to accord an eyewitness identification. The estimator variable of *stress at event time* constitutes the primary focus of this thesis. Knowing whether stress at the time of viewing the perpetrator inhibits later recognition gives administrators of justice a better idea of whether a witness is likely to commit a misidentification when presented with a lineup.

Of the three eyewitness measures discussed here, researchers have paid the most attention to lineup identifications in relation to stress. The findings on this topic are largely consistent. For example, Clifford and Hollin (1981), Peters (1988) and Morgan et al. (2004) all found decreased identification accuracy following stress at encoding. Participants in Clifford and Hollin's (1981) violent-film condition selected the target 30% of the time, whereas participants in the non-violent-film condition selected the target 40% of the time. Using a more provocative stressor (an injection) and a live event methodology, Peters (1988) found an even more dramatic contrast: Subjects demonstrated 41% accuracy for the nurse (who administered the inoculation) versus 66% accuracy for the researcher (a neutral target). Morgan et al. (2004) used the most realistic and extreme stress manipulation and found the greatest disparity. When viewing a photo-spread, subjects identified their interrogators from the high-stress condition with 34% accuracy, whereas they identified their interrogators from the low-stress condition with 76% accuracy. This difference was slightly lessened by the use of a live line-up method, but was still highly significant. In sum, the general finding that stress impairs memory, as discussed earlier, clearly holds true for the specific outcome measure of lineup identifications: Stress at the time of exposure diminishes the accuracy with which one can identify a target from a lineup.

The act of identification is the crux of most cases involving eyewitness testimony. A positive eyewitness identification is taken as weighty evidence against the suspect, and lineup identification has therefore been included as one of the three dependent variables in this thesis. Culprit-presence versus culprit-absence constitutes one of the independent variables in this thesis

---

because eyewitness reliability may differ with each set of circumstances. It is of interest to know the extent to which stress at encoding impairs witnesses' performance on target absent lineups relative to target-present lineups. Both lineup presentation and lineup construction have been carefully considered in the design of this study, as these system variables could compromise eyewitness accuracy and confound the results if not properly addressed.

### Summary and Conclusion

Heightened stress at the time of a witness's exposure to a perpetrator could compromise the witness's reliability, indirectly leading to a miscarriage of justice. Stress has been convincingly shown to impair eyewitness memory in adults. This pervasive empirical finding is supported by neurological research, which has revealed that the chief episodic memory center in the brain (the hippocampus) is vulnerable to the effects of stress. Deffenbacher's theoretical model (1994) seems to explain this phenomenon most aptly: When an individual in a state of heightened physiological activation experiences additional cognitive anxiety, memory decrements ensue. Such decrements in eyewitness memory can be measured by obtaining person descriptions and lineup identifications from witnesses, as well as testing their memory for central and peripheral details of the witnessed event.

Increasing attention is being paid to adolescent witnesses in the criminal justice system, but our knowledge of how stress affects their reliability remains limited. No experimental stress manipulations can possibly simulate the trauma experienced by witnesses of violent crimes within ethical limits. However, extra-laboratory investigations of naturally occurring stress in juveniles, employing a live-event methodology, may provide some insight into the effects of stress on their eyewitness accuracy. The forensic relevance of such research can be enhanced by obtaining validity checks, specifically physiological correlates, of the subjects' stress levels. In order to prevent miscarriages of justice, courts need to know whether the conditions surrounding a crime render eyewitnesses less capable of correctly identifying a culprit. Given that heightened

---

stress is present in the majority of eyewitness cases, particularly those involving juveniles, research into this variable is profoundly warranted.

University of Cape Town

---

## CHAPTER THREE: EXPERIMENT 1

### *Method*

#### Rationale, Aim, and Hypotheses

*Rationale.* If stress at the time of viewing a culprit contributes to the chances of a witness providing unreliable testimony or making an identification error, then it is critical that the impact of this variable be better understood. There is a dearth of research concerning the eyewitness performance of adolescents following stress, making it clear that this topic merits investigation.

*Aim.* I aimed to establish the effect of stress at encoding on eyewitness memory performance in adolescents.

*Hypotheses.* With regard to the effect of the stress manipulation, I expected that participants would show evidence of activation mode dominance. This was expected to manifest as elevated heart rate, skin conductance and state-anxiety.

With regard to eyewitness memory performance, I expected that participants would perform worse on all eyewitness memory measures following the presence of stress at encoding. Specifically, I predicted that:

- 1) Participants' target descriptions of a research assistant met at a stressful exposure session would be less correct, accurate, precise and informative than their descriptions of a research assistant met at a non-stressful exposure session.
-

- 2a) Participants would show poorer recall for details of the stressful exposure session than for details of the non-stressful exposure session.
- b) Recall for central details of the stressful exposure session would be less impaired than recall for the peripheral details of the stressful exposure session.
- 3a) Participants would be less accurate in their lineup identifications of the research assistant from the stressful exposure session than in their lineup identifications of the research assistant from the non-stressful exposure session.
- b) Participants would be less accurate in their identifications when the lineup type was target-absent than target-present.
- c) Participants would perform particularly poorly when making identifications in the target absent condition of the research assistant for the stressful exposure session.

### Design

This study is experimental in design. Specifically, a 2 (Event type) X 2 (Lineup type) repeated measures design has been used. The levels of the first independent variable, 'event type', are *stressful* and *non-stressful*; this refers to the circumstances under which the participant was exposed to an unknown research assistant. The levels of the second independent variable, 'lineup type', are *target-present* and *target-absent*; this refers to whether or not the target research assistant appeared in the lineup.

There are three major classes of dependent variable in this study: description performance, detail recollection, and identification accuracy. All three dependent variables are measures of eyewitness memory. Description performance refers to the participant's written description of the target individual and is operationalised in four ways: total number of correct descriptors, accuracy, precision and informativeness. Detail recollection refers to the participant's memory for details of the exposure sessions. Three different categories of detail recollection are investigated: details of the research assistant, details of the event, and details of the environment.

---

Details of the research assistant and details of the event are considered central details, while details of the environment are considered peripheral details<sup>6</sup>. Identification accuracy is the frequency with which the target individual is correctly identified or rejected from a lineup. All three dependent variables are examined within subjects, comparing their memory performance for the stress- and non-stress-conditions.

### Participants

Twenty six white male high school students participated in this study (mean age = 14.3 years). Due to limited resources, I used only white targets and participants (a substantial other-race effect has been found for face-recognition [Meissner & Brigham, 2001], and it would be undesirable to have this variable confound the results). Females were excluded from the sample because their cortisol levels are subject to fluctuation in accord with their menstrual cycles. As discussed in the literature review, cortisol levels exert an influence on memory performance. Females exhibit stress-induced free cortisol levels similar to those of males only during the late luteal phase of their cycle (typically less than 1 week per month) (Kirschbaum, Kudielka, Gaab, Schommer, & Hellhammer, 1999). Given the design of the study, it would have been unfeasible to accommodate this element, and females were thus excluded from the sample.

The participants were recruited by their teacher. Compensation was offered in the form of a talk on 'Careers in Psychology' to be delivered to all students in the participating grades.

---

<sup>6</sup> For examples of the central and peripheral details about which participants were questioned, please refer to Task 2 in Appendix D.

---

## Measures, Instruments and Materials

### *Physiological Measures*

An Ambulatory Monitoring System (AMS) was used for the measurement of bio-signals (Groot, de Geus & de Vries, 1998). The AMS is a lightweight, portable device that participants can wear relatively comfortably and unobtrusively beneath their clothing. In this experiment, the AMS was used to take electrocardiogram readings and to measure skin conductance levels (SCL). Three electrodes were attached to each participant's torso in order to measure heart rate, and two electrodes were attached to the middle and index fingers to obtain SCL readings.

### *Self-reported Anxiety*

The Spielberger State-Trait Anxiety Inventory (STAI) (Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983) was used to measure participants' subjective levels of anxiety (see Appendix A). The STAI is a self-report measure containing two 20-item scales. The *State* scale gauges participants' present state of anxiety with questions pertaining to how they feel 'right now... at this moment'. The *Trait* scale reflects the extent to which anxiety is a strong character attribute, with questions pertaining to how they 'generally feel'. The scale has good psychometric properties, demonstrating a high degree of internal consistency ( $\alpha = 0.91$ ) and validity (correlation between State- and Trait-anxiety scales = .65; correlations with other measures of trait anxiety [IPAT, TMAS, AACL] range from .73 to .85) (Spielberger et al., 1983). The STAI has been widely used with high school students (Spielberger et al., 1983) and was thus an appropriate choice of instrument for this study.

### *Psychosocial Stressor*

A public speaking task was used to induce psychosocial stress in the participants. This choice of methodology was based on research using the 'Trier Social Stress Test' (Kirschbaum, Pirke, & Hellhammer, 1993), which involves a public speaking task in front of a panel of evaluators, as well as solving mental arithmetic problems presented by the panel. The TSST has been shown to

---

elicit significant and reproducible psychobiological stress responses, specifically, elevations in heart rate and increases in the concentration of certain hormones (including cortisol).

Kirschbaum et al. (1993) suggest that the substantial element of ego involvement and the anticipation of negative consequences associated with public speaking are largely responsible for the physiological changes experienced by participants. The constraints of working in a school environment dictated that the methodology for the present experiment deviate somewhat from the 'Trier Social Stress Test' protocol. Participants were required to deliver a five minute speech in front of the teacher and approximately 20 classmates on a pre-assigned topic. The speech constituted the public speaking requirement of the class syllabus, and was therefore graded by the teacher. It was felt that this evaluative element would provoke the anticipation of negative consequences in participants. This, together with the substantial ego involvement elicited by the relatively large audience, was deemed sufficient to induce stress-related physiological changes.

#### *Research Assistants*

Two white, female undergraduate research assistants were employed to take participants' physiological measurements and to serve as the target individuals whom participants would have to describe, answer questions about, and identify from lineups. Every participant encountered one of the research assistants during the stress session, and the other research assistant during the non-stress session. The research assistants were counterbalanced so that neither was exclusively associated with the stress or the non-stress condition. This ensured that any within-subject differences were attributable to the presence or absence of stress, rather than such confounding variables as distinctiveness or attractiveness, which might make one target face more memorable than another (McQuiston-Surrett et al., 2006). To minimise the chance of participants committing source-monitoring errors, I selected two research assistants bearing little physical resemblance to one another. The chosen targets differed with regard to hair colour, hairstyle, eye colour, weight and build. My subjective judgment of their physical dissimilarity was confirmed by the physical descriptions gathered from a class of undergraduate psychology students<sup>7</sup>: The

---

<sup>7</sup> The purpose of obtaining these descriptions was to create a standardised version that could be used for scoring the target-descriptions provided by actual participants. The procedure for developing the standardised description for each target will be discussed in greater depth at a later point.

---

overlap between the targets extended only to descriptions of height ('average'), lips ('thin') and eyebrows ('thin').

### *Evaluators*

A third research assistant was employed as a confederate in the present experiment. This research assistant posed as an external evaluator during the participants' speeches. Participants were told that she had been sent from the university to evaluate the level of public speaking at their high school. It was felt that this 'authoritative' presence might intensify the anxiety experienced by the participants, thus generating a more reliable stress response. Participants were debriefed concerning this deception at the end of the experiment.

### *Stimulus Materials*

Lineups were constructed for each of the target research assistants. From a database of high quality photographs of student faces from the University of Cape Town, I selected possible foils for each target. Standardised descriptions<sup>8</sup> of the targets were used as a rough basis for selecting foils. Within the limits of the UCT database, I was unable to find foils that matched the target descriptions on every particular, but all the potential foils bore a reasonable perceptual similarity to the targets. As this was a subjective judgment, I asked five independent viewers for their opinions. For each research assistant, the eight foils that were unanimously thought to be most similar were included in the lineup. Seven of the foils appeared in both target-present and target-absent lineups, while the eighth appeared only in target-absent lineups in place of the missing target face. The lineups for the two target research assistants did not contain any foils in common.

Using Adobe Photoshop CS3, all faces were standardised to have a height of 5cm. Face images were edited to remove clothing below the neckline, and were presented in colour from a frontal

---

<sup>8</sup> The compilation of the standardised descriptions will be discussed under *Scoring Materials*.

viewpoint. When presented with a simultaneous lineup, witnesses tend to distinguish among faces on the basis of systematic difference which may ultimately be irrelevant (such as variations in lighting and contrast, for example). To counteract witnesses' propensity to search for trivial difference, all faces were pasted against different coloured backgrounds.

A mock-witness evaluation was conducted to assess lineup bias and fairness (Doob and Kirshenbaum, 1973, as cited in Malpass et al., 2007). I enlisted the help of UCT undergraduates to serve as mock witnesses during lectures and tutorial classes. Eighty three undergraduates viewed the lineup for the first research assistant (Appendix B) and 76 undergraduates viewed the lineup for the second research assistant (Appendix C). The mock witnesses were provided with no information, but were required to select the lineup member they thought was the target<sup>9</sup> (the target research assistants were not present). The premise of a mock-witness evaluation is that the target should be selected with a probability no greater than chance if the lineup is unbiased (Malpass et al., 2007). The lineup bias was .18 ( $CI_{.95} = .10, .26$ ) for the first lineup (Appendix B) and .16 ( $CI_{.95} = .08, .24$ ) for the second lineup (Appendix C), indicating that there was no significant structural bias towards the targets or away from the targets.

The nominal size of the lineups was eight members, because this is the number traditionally used in South African police practice (Tredoux, 1996). The effective size—the number of plausible members in the lineup (Malpass, 1981)—was calculated as  $E'$  (Tredoux, 1998, as cited in Malpass et al., 2007). For the first lineup,  $E' = 5.16$  ( $CI_{.95} = 4.31, 6.47$ ), and for the second lineup,  $E' = 6.52$  ( $CI_{.95} = 5.40, 8.22$ ), indicating a sufficient number of plausible foils and a high degree of fairness in both lineups.

### *Eyewitness Tasks*

Participants were required to complete four paper-and-pencil tasks during the final session of the experiment (see Appendix D). Each task was completed twice: once with reference to the stress

---

<sup>9</sup> This is a modification of the traditional mock witness task, where witnesses are provided with a physical description of the target before selecting him or her from the lineup (Doob & Kirshenbaum, 1973, as cited in Malpass et al., 2007).

session, and once with reference to the non-stress session. Written instructions were provided for all three tasks.

The first task involved a free-recall target description. It was emphasized that participants should write a *physical* description of each target research assistant, and it was suggested that they describe such elements as height, build, hair length and colour, and facial features.

The second task involved a set of nine questions pertaining to details about the research assistant, details about the experience of the session itself, and details about the environment in which the session took place. These questions demanded *free recall* from participants. Participants completed one set of questions with reference to the first exposure session and the first research assistant they met, and another set of questions with reference to the second exposure session and the second research assistant they met. Though the questions for the two sessions were highly similar (if not identical), the answers were different. With reference to both sessions, for example, the participant was questioned about the colour of the ink with which he completed the written task, the image on the research assistant's shirt, and the contents of a jar on the table at the data collection station. One of the research assistants always provided the participant with a green pen, wore a Puma t-shirt and placed flowers in the jar. The other research assistant always provided participants with a red pen, wore a UCT sweatshirt and placed pencils in the jar. Each research assistant used the same props for every session and wore the same clothing and hairstyle, so the set of answers associated with her was invariant. Seeing as the research assistants were counterbalanced across conditions, each set of answers was associated with the stress session half of the time and associated with the non-stress session half the time.

The third task involved the set of questions completed in Task 2 simply presented in a multiple choice format<sup>10</sup>. This task differed from Task 2 in that it demanded recognition, not recall. The presentation of cues (in the form of multiple choice options, for example) is known to facilitate memory retrieval (Haist, Shimamura, & Squire, 1992). In this study, it was of interest to know

---

<sup>10</sup> Task 3 was presented to participants only after they had completed and handed in Task 2.

whether memory errors were a result of poor encoding or poor retrieval. In the case of the former, participants may be expected to perform equally poorly on the free recall and multiple choice tasks. In the case of the latter, participants would perform significantly better on the multiple choice task.

The final task involved lineup identifications. Each participant saw two lineups, one for each research assistant. Target-absence versus target-presence is an important element in eyewitness research: Police may or may not have the culprit in their lineup, and it is essential to know how eyewitnesses perform under either set of circumstances. Thus, target-absence versus target-presence was included as a within-subjects variable. Of the two lineups that each participant saw, one was always target-absent, in which the research assistant did not appear, and the other was always target-present, in which the research assistant did appear. The order of target-absent and target-present lineups was randomised across participants and counter-balanced across research assistants (such that each research assistant appeared equally often in a target-present lineup as in a target-absent lineup). Instructions for making lineup selections were provided in writing so as to avoid any verbal or non-verbal cues that may be inadvertently communicated by the lineup administrator. In these instructions, participants were advised that the research assistant may not actually appear in the lineup. They viewed the lineups on a widescreen laptop computer, where the eight lineup members were displayed simultaneously, with a number beneath each one. The order in which the lineup members appeared was randomised across participants, eliminating any confounds associated with lineup order. Participants made their selections on the question paper by placing an X in one of nine boxes, eight of which contained numbers corresponding with the faces on the screen, and one of which contained the option, 'none of them'. Participants viewed the two lineups consecutively, but were only shown the second lineup after making a selection for the first.

### *Scoring Materials*

Standardised descriptions of each target research assistant were generated for the scoring of participants' descriptions. To this end, two tutorial classes of UCT psychology students were

---

asked for their assistance in providing physical descriptions of the targets. Each research assistant was presented to a class of approximately 20 students and each student provided a written description of her physical appearance. Half of the class wrote their descriptions with the target in full view, while the other half wrote it from memory. The procedure was as follows: I entered the room with the target, introduced her to the students and explained that their descriptions would be used in the compilation of a standardised description for an eyewitness experiment. This took approximately two minutes. I then announced that half of the students would be describing the target from memory and asked those members of the class to turn around in their chairs so that they were facing away from the target. All students were then asked to write their descriptions in silence (whether from memory or while looking at the target). During this time, the target research assistant stood still at the front of the class in a relaxed position, smiling occasionally, but trying to maintain a neutral expression. When everyone had finished writing, the memory group was permitted to turn around again and everyone submitted their descriptions. From these descriptions, I extracted the most commonly mentioned characteristics: Items observed by four or more students in either the 'full view' group or the 'memory group' were taken to be correct and included in the standardised description. Whether students were able to look at the target while describing her seemed to have little effect on what they reported: The same characteristics emerged as salient from both sets of descriptions (see Appendix E).

#### Procedure

The protocol for this study was approved by the Psychology Department Ethics Committee at the University of Cape Town. Data collection took place at an all-boys high school in Rondebosch. This study was preceded by a pilot study, which took place at a different high school in the Cape Town area. We tested the feasibility of the protocol on fourteen pilot participants in a classroom setting and made small modifications, such as adjusting the length of exposure to the target.

Some weeks before the study commenced, consent forms were sent home with all students interested in participating (see Appendix F). The form contained details of what would be

---

expected of the participant over the three sessions of the study, the procedure for taking the physiological measurements and his rights as a research participant. Consent was obtained from both participants and their parents. They were informed that the study was about memory (though not specifically eyewitness memory) and public speaking, and that all students would have to give a speech irrespective of whether they chose to take part in the study, as it was part of the class syllabus.

The present experiment was conducted over a period of three weeks, with each participant attending three sessions spaced roughly a week apart. Sessions 1 and 2 constituted the *exposure phase*. Of the two exposure phase sessions, one was designed to be stressful, while the other was designed to be neutral – the ordering of these events was counterbalanced across participants. During these sessions, participants interacted with two target research assistants whom they would have to recognize at session 3 (the *memory phase*). The counterbalancing of condition (whether the participants experienced the stress or non-stress session first) and target exposure (whether the participant met Research Assistant 1 or Research Assistant 2 first) is depicted in Table 1.

Table 1

*Counterbalanced Variables in the Design of Experiment 1*

	Week 1		Week 2		Week 3
	Stress	Non-stress	Stress	Non-stress	
Research Assistant 1	n = 6	n = 6	n = 6	n = 8	
Research Assistant 2	n = 8	n = 6	n = 6	n = 6	Memory Phase
Total	14	12	12	14	(all participants)

Approximately half the participants delivered their speeches in the first week and thus experienced the stress session first, while the other half delivered their speeches in the second week and thus experienced the non-stress session first. All sessions took place at the ‘data collection station’ which was set up in a quiet room down the corridor from the classroom.

In both the stress condition and the non-stress condition, the participant was met by one of the target research assistants when he arrived at the data collection station. The research assistant introduced herself by name, made eye contact and shook the participant's hand. This deliberate attempt to establish interpersonal contact aimed to provide all participants with an equal encoding opportunity.

In both the stress condition and the non-stress condition, the participant was seated at a desk where the research assistant connected him to the AMS. The research assistant adhered to a strict scripted dialogue while interacting with the participant. This included action cues, such as asking the participant to raise his shirt in order to place the electrodes, and necessary information gathering, such as establishing handedness so that the SCL electrodes could be attached to the non-dominant hand. Furthermore, each research assistant asked one contrived question and offered the participant a book to press on, both of which would serve as 'details of the event' during the memory phase. If the participant tried to make conversation, the research assistant gave neutral responses that would not encourage further conversation. If the participant asked questions, the research assistant informed him that I would provide the answers after the session. After connecting the participant to the AMS and carrying out the necessary dialogue, the research assistant retired to an adjacent room where she was out of the participant's sight. This concluded the exposure period for the session. Both research assistants always followed the same procedure, which took approximately three minutes, ensuring that target exposure was uniform across participants and conditions.

While attached to the AMS, the participant was left alone to complete a puzzle task (comprising of age-appropriate word search puzzles and visual illusion tasks) in the non-stress condition, or to complete the STAI in the stress condition. In the non-stress condition, it was emphasised that he would not be evaluated on his performance of the puzzle task to prevent him from experiencing any anxiety. The participant remained connected to the AMS for five minutes in the non-stress condition (long enough to obtain reliable HR and SCL readings), and for as long as it took him to complete the STAI in the stress condition (approximately five minutes).

---

After the appropriate amount of time had elapsed, I disconnected the participant from the AMS. He returned to the classroom to deliver his speech if he was experiencing the stress condition, or to resume normal classroom activities if he was experiencing the non-stress condition.

Participants in the stress condition did not remain connected to the AMS device while delivering their speeches as this had proved unfeasible in the pilot study. It was apparent from the pilot study that wearing the AMS while speaking distracted the participant and the audience, as well as slowing the data collection process considerably (allowing two rather than four students to participate per period). Disconnecting the participant from the AMS before he delivered his speech had no effect on the physiological data that I used in the analysis. As the variable under investigation was stress *at encoding*, the participant's state of physiological activation during his interaction with the research assistant (and in the few minutes immediately thereafter), was of chief interest. Participants in the stress condition were coordinated so that they could deliver their speeches immediately upon their return to the classroom. Thus, participants knew that they would imminently have to give their speeches and were presumably in a state of anticipatory anxiety when they met the research assistant – it was this critical period of physiological activation that was recorded by the AMS. Each participant was asked to send the next participant to the data collection station upon his return to the classroom, allowing the process to run smoothly without much class disruption. For the non-stress condition, the teacher was asked to schedule classroom activities that would be non-aversive and non-anxiety provoking for the participants. The participant was summoned from class a week later and experienced the opposite condition, interacting with the other research assistant; the procedure was as described above.

In the third week, the participant took part in the final session of the experiment, where he was required to complete a series of memory tasks pertaining to the previous two sessions. I conducted the memory session, and no research assistants were present. The participant completed four eyewitness tasks (target description, free recall questions, multiple choice questions and lineup identification). Each task was completed twice: once for the stress session and once for the non-stress session. The participant was then thoroughly debriefed as to the aim of the experiment.

---

## Data Analysis

### *Scoring*

*Target Descriptions.* Target descriptions were scored by the author and an independent rater, using a procedure where we were blind to condition (stress or non-stress). Using the standardized description for each research assistant as the benchmark for truth, we assigned points per correct ‘syntactic unit’ (i.e. any word or phrase that communicates meaningful information). We agreed on a strict scoring protocol ahead of time, and achieved strong correlations on ‘correct items’ for both research assistants (*Pearson’s*  $r_{(RA1)} = .91$ ; *Pearson’s*  $r_{(RA2)} = .91$ ). In addition to ‘correct items’, we coded for ‘subjective items’ and ‘incorrect items’. Relative judgments (such as ‘she was a little shorter than me’) and subjective appraisals (such as ‘she had a friendly face’) were coded as ‘subjective items’, while false facts were coded as ‘incorrect items’. There were significant correlations between our scores for ‘subjective items’ (*Pearson’s*  $r_{(RA1)} = .90$ ; *Pearson’s*  $r_{(RA2)} = .82$ ) and ‘incorrect items’ (*Pearson’s*  $r_{(RA1)} = .89$ ; *Pearson’s*  $r_{(RA2)} = .83$ ).

Given the strong inter-rater agreement on all item types, I calculated the average between our scores for each target description, and used these in subsequent analyses. These averaged scores were also used to calculate the accuracy for each description: The number of correct items was divided by the sum of the number of subjective items and the number of incorrect items, and multiplied by 100.

Each description was also given a score for ‘precision’ and a score for ‘informativeness’. Precision was defined as the degree of detail and specificity present in the description, and was scored on a scale from 1 (extremely vague) to 7 (extremely precise). Informativeness was defined as the usefulness or ‘identification value’ of the description—in other words, the facility with which one could pick the target out of a crowd on the basis of the description. This criterion was also scored on a scale from 1 (useless) to 7 (extremely useful). Scores for precision and informativeness were assigned by the author and two independent raters, again using a blind procedure. The precision and informativeness scores assigned by the three raters were averaged to provide the scores used in the analysis. Pearson’s correlation coefficient was calculated for

---

each pair of raters, and Kendall's coefficient of concordance was used to establish inter-rater reliability among all three raters. Inter-rater agreement for precision is represented in Table 2 below and inter-rater agreement for informativeness is represented in Table 3 below. In both Table 2 and Table 3, the correlation coefficients for Research Assistant 1 are represented in the bottom left portion of the table, and the correlation coefficients for Research Assistant 2 are represented in the top right portion of the table.

Table 2

*Intercorrelations Between Independent Raters of Description Precision*

	Rater 1	Rater 2	Rater 3
Rater 1	—	.74 <sup>b</sup>	.72 <sup>b</sup>
Rater 2	.53 <sup>a</sup>	—	.65 <sup>b</sup>
Rater 3	.68 <sup>a</sup>	.45 <sup>a</sup>	—

<sup>a</sup> Correlation coefficients for Research Assistant 1; Kendall's Coefficient of Concordance = .14

<sup>b</sup> Correlation coefficients for Research Assistant 2; Kendall's Coefficient of Concordance = .22

Table 3

*Intercorrelations Between Independent Raters of Description Informativeness*

	Rater 1	Rater 2	Rater 3
Rater 1	—	.76 <sup>b</sup>	.68 <sup>b</sup>
Rater 2	.82 <sup>a</sup>	—	.84 <sup>b</sup>
Rater 3	.66 <sup>a</sup>	.72 <sup>a</sup>	—

<sup>a</sup> Correlation coefficients for Research Assistant 1; Kendall's Coefficient of Concordance = .22

<sup>b</sup> Correlation coefficients for Research Assistant 2; Kendall's Coefficient of Concordance = .44

*Detail Questions.* One point was assigned per correct answer in each category ('Details of the Research Assistant', 'Details of the Event', and 'Details of the Environment'). Points were summed to yield a total per category and an overall total.

*Lineup Identifications.* Participants' selections were coded as 'correct identification', 'foil identification' or 'false rejection' in the target-present condition, and as 'correct rejection' or 'false positive' in the target-absent condition.

#### *Statistical Analysis*

For all statistical analyses, I used the software package Statistica Version 8 (StatSoft, 2007). The threshold for statistical significance was set at  $\alpha = 0.05$ .

---

## Results

### Measures of Stress

#### *Heart Rate*

Heart rate readings were taken from all participants at the stress session and the non-stress session. As predicted, mean heart rate was higher for the stress session ( $M = 80.23$ ,  $SD = 12.39$ ) than for the non-stress session ( $M = 75.78$ ,  $SD = 9.65$ ). A dependent-samples  $t$  test was conducted to establish whether the mean difference between the conditions (4.45 beats per minute) was significant. The assumption of normality was tested using the *Shapiro-Wilk  $W$  test*, as its power properties are superior to those of other tests of normality (Shapiro, Wilk, & Chen, 1968). A non-significant  $W$  statistic indicates that the data is normally distributed (where  $\alpha$  is set at .05). The heart rate data from the stress condition satisfied the assumption of normality ( $SW-W = 0.986$ ,  $p = .963$ ), as did the heart rate data from the non-stress condition ( $SW-W = 0.936$ ,  $p = 0.109$ ). The results from the dependent-samples  $t$  test confirmed that the stress manipulation significantly increased heart rate in the sample ( $t(25) = 2.424$ ,  $p = .024$ ), yielding a small to medium effect ( $d = 0.39$ ).

#### *Skin Conductance Level*

Skin conductance readings were taken from all participants at the stress session and the non-stress session. As predicted, mean skin conductance level was higher for the stress session ( $M = 10.53$ ,  $SD = 4.00$ ) than for the non-stress session ( $M = 8.35$ ,  $SD = 2.34$ ). A dependent-samples  $t$  test was conducted to establish whether the mean difference between the conditions ( $2.18\mu\text{S}$ ) was significant. The assumption of normality was met for the non-stress condition ( $SW-W = 0.982$ ,  $p = .910$ ), but not for the stress condition ( $SW-W = 0.896$ ,  $p = .013$ ). The distribution of readings in the stress condition was positively skewed, as is typical with skin conductance data (Cacioppo, Tassinary, & Berntson, 2000). The natural log of the data was taken, yielding a more normal distribution ( $SW-W = 0.966$ ,  $p = .520$ ) (see Appendix G). All subsequent analyses were conducted with the log-transformed values for both the stress and non-

---

stress data. The results from the dependent-samples  $t$  test confirmed that the prospect of public speaking significantly increased skin conductance level in the sample ( $t(25) = 3.057, p = .005$ ), yielding a medium to large effect ( $d = 0.69$ ).

#### *Autonomic Reactivity: The Relationship between Heart Rate and SCL*

Subjects' mean heart rate readings for the stress condition were plotted against their mean skin conductance readings for the stress condition, and likewise for the non-stress condition. The stress condition yielded a small positive correlation ( $r = .179$ ), but the non-stress condition revealed no relationship between heart rate and skin conductance ( $r = .014$ ). Given the poor relationship between heart rate and skin conductance within the experimental conditions, it is no surprise that 'change in heart rate' and 'change in skin conductance' appear to be unrelated ( $r = .029, p = .889$ ).

The negligible relationship between change in heart rate and change in skin conductance informs the primary choice of the statistical procedure used in the analysis for this study. Multiple regression takes into account the contribution of autonomic reactivity to memory performance, but does not rely on parallel variation between heart rate and skin conductance (indeed, the analysis is strengthened by the absence of multicollinearity between these two autonomic indicators).

#### *Self-reported Anxiety*

All participants completed the State-Trait Anxiety Inventory at the stress session. It was predicted that mean state-anxiety would be significantly elevated as a result of the stress manipulation. Means and standard deviations are reported in Table 4.

---

Table 4

*State- and Trait-Anxiety: Descriptive Statistics*

	Sample Mean (n = 26)	Sample SD <sup>a</sup> (n = 26)	Norm Mean (N = 202)	Norm SD <sup>a</sup> (N = 202)
State-Anxiety	37.00	10.91	39.45	9.74
Trait-Anxiety	36.58	9.42	40.17	10.53
Difference	0.42	9.11	-0.72	-

<sup>a</sup> Standard deviation

A single-sample *t* test was conducted to establish whether the difference between state- and trait-anxiety was greater for the sample than for the population. If this were the case, it would confirm that participants were experiencing subjective stress. To my knowledge, this approach has not been used before – it is more typical for researchers to administer the STAI repeatedly and then to compare scores within subjects. In the absence of repeated administration, I was obliged to compare the sample mean with the population mean. State anxiety is most meaningful in relation to trait anxiety. To illustrate, if a participant who is generally highly-strung and another who is generally easygoing (as indicated by their ratings of trait-anxiety), score exactly the same for state-anxiety, we will draw different conclusions about their respective levels of stress. It therefore seemed most sensible to investigate the mean difference scores between state- and trait-anxiety in the sample relative to the population (rather than to compare state-anxiety and trait-anxiety individually).

In comparing a sample with a population, a single-sample *t* test can be used when the standard deviation of the population is unknown (Howell, 2004). The population difference score between state- and trait-anxiety ( $\mu = -0.72$ ) was set as a reference constant against which the difference scores for the sample ( $M = 0.42$ ) were tested. The difference scores for the sample were normally distributed ( $SW-W = 0.962, p = .429$ ). The sample mean did not differ significantly from the

population mean ( $t(25) = 0.640, p = .528, d = 0.025$ ) indicating that the sample is representative of an unstressed population. Though it appears that the stress manipulation did not provoke heightened subjective stress, it is worth noting that the sample's mean state- and trait-anxiety scores were lower than the norm (by 2.45 points for state-anxiety and 3.59 points for trait-anxiety). These differences are not statistically significant, but they do yield small and medium effects respectively (State Anxiety:  $t(25) = 1.145, p = .263, d = 0.33$ ; Trait-Anxiety:  $t(25) = 1.944, p = .063, d = 0.54$ ). The implication is that participants may have underreported their levels of anxiety for social desirability reasons, warranting a cautious interpretation of the subjective anxiety results for the adolescent sample.

#### *Overview*

Heart rate, skin conductance level and self-reported anxiety were measured as validity checks on subjects' state of arousal. Heart rate and skin conductance level were significantly elevated in the stress condition. This indicates that the stress manipulation was effective in provoking a state of stress, as represented by physiological activation. Contrary to expectations, participants did not report heightened state-anxiety in response to the stressor. This suggests that the prospect of public speaking did not provoke subjective anxiety in the sample.

### Measures of Memory

#### *Target Descriptions*

All participants completed target descriptions for the stress and non-stress sessions. Descriptions were scored on four dimensions: (i) total number of items correct, (ii) accuracy, (iii) precision and (iv) informativeness. I predicted that participants' target descriptions for the non-stress session would contain more correct items and would be more accurate, precise and informative than those for the stress session. Means and standard deviations are reported in Table 5.

---

Table 5

*Target Description Scores: Descriptive Statistics*

Dimensions of Description Performance	Stress (n = 26)		Non-stress (n = 26)	
	Mean	SD <sup>a</sup>	Mean	SD <sup>a</sup>
(i) Total Correct Items	1.42	1.03	1.56	1.04
(ii) Accuracy	46.45	27.52	45.05	26.12
(iii) Precision	2.63	0.74	2.80	0.89
(iv) Informativeness	2.61	0.84	2.87	0.93

*Note.* Reported means are untransformed.

<sup>a</sup> Standard deviation

As shown in Table 5, description scores for the non-stress condition were elevated above scores for the stress condition on three of the description dimensions (total correct items, precision and informativeness).

*Test: Dependent-samples *t* test.* For each dimension where the direction of the difference was as predicted (total correct items, precision and informativeness), a dependent-samples *t* test was conducted.

*Assumptions.* The assumption of normality was met for total items correct in the non-stress condition ( $SW-W = 0.955$ ,  $p = .300$ ), but not in the stress condition ( $SW-W = 0.828$ ,  $p = .001$ ). The natural log of the data from the stress condition was taken, yielding a more normal distribution ( $SW-W = 0.928$ ,  $p = .068$ ) (see Appendix H). All subsequent analyses were conducted with the log-transformed values for both the stress and non-stress data, but the reported means are untransformed. Precision scores were normally distributed for both the non-stress condition ( $SW-W = 0.961$ ,  $p = .402$ ) and the stress condition ( $SW-W = 0.965$ ,  $p = .505$ ), as were informativeness scores (Non-stress:  $SW-W = 0.970$ ,  $p = .629$ ; Stress:  $SW-W = 0.958$ ,  $p = .348$ ).

*Results.*

Table 6

*Results of the Dependent-samples  $t$  Test for Description Dimensions*

	Stress Mean (n = 26)	Non-stress Mean (n = 26)	$t$	df	$p$
Total correct items	1.42	1.56	-0.50	25	0.31
Precision	2.63	2.80	-1.12	25	0.27
Informativeness	2.61	2.87	-1.74*	25	0.05

*Note.* Reported means are untransformed.

\* $p \leq .05$

As shown in Table 6, the results for total correct items ( $t(25) = -0.501, p = .310$ ) and precision ( $t(25) = -1.121, p = .273$ ) were non-significant; thus, the null hypothesis that no difference exists in total correct items or precision across conditions cannot be rejected. The effect for total correct items was negligible ( $d = -0.09$ ), whereas precision produced a small effect ( $d = -0.19$ ).

Participants provided descriptions that were more informative in the non-stress condition than in the stress condition, as shown by the significant  $t$  test result and the small effect size for informativeness ( $t(25) = -1.743, p = .047, d = -0.29$ ).

*Test: Regression.* As an alternative approach, a multiple regression analysis was conducted to examine the particular role that heart rate and skin conductance played in this pattern of results. It is possible that the experimental conditions did not vary enough in terms of stress exerted on the participants to create clear differences in accuracy, precision or number of correct items. This

would not compromise a multiple regression analysis, as the objective is not to analyse mean differences, but to assess the contribution of predictor variables to changes in dependent variable measures (however small those changes might be). The change (between the stress and non-stress conditions) for each dimension of description performance was regressed on both the change in heart rate and the change in skin conductance level.

*Assumptions.* Regression requires the absence of multicollinearity. In addition, the assumptions of normality, linearity and homoscedasticity of residuals must be met.

The assumption concerning multicollinearity was not violated: The low correlation between the two predictors (change in heart rate and change in skin conductance) indicates that tolerance was high (over 95%) and that multicollinearity was absent.

The assumptions of normality, linearity and homoscedasticity of residuals were tested by examining the residuals scatter-plots for each description dimension. Following the removal of two outliers for precision and one outlier for informativeness, these assumptions were upheld for all dimensions (see Appendices I, J, K and L). The subsequent regression analyses for precision and informativeness were conducted with the outliers removed.

---

*Results.*

Table 7

*Regression Results for Target Description Dimensions (where  $\hat{y}$  = the difference in the DV between stress and non-stress)*

	$R^2$	$F$	$p$	Df	$\beta_1$ ( $\Delta$ HR)	$p$	$\beta_2$ ( $\Delta$ SCL)	$p$
$\Delta$ Total correct items	0.15	2.06	.15	2, 23	0.36	.08	-0.17	.40
$\Delta$ Accuracy	0.47	3.34*	.05	2, 23	0.47*	.02	-0.03	.86
$\Delta$ Precision	0.25	3.48*	.05	2, 21	-0.12	.53	-0.48*	.02
$\Delta$ Informativeness	0.11	1.40	.27	2, 22	-0.33	.12	-0.07	.75

*Note.* Delta refers to the change in the description dimension between the stress and the non-stress condition. This was calculated by subtracting the participant's stress score from his non-stress score. Conversely, change in heart rate and skin conductance level was calculated by subtracting the participant's non-stress score from his stress score.  
\* $p \leq .05$

Four different regression models are represented in Table 7, with the same independent variables (change in heart rate and change in skin conductance) fitted to each of the four target description dimensions. From Table 7, it can be seen that only two dimensions of description performance (accuracy and precision) produced significant regression models. The overall regression for change in description accuracy was not significant at the 5% level, but approached significance ( $F(2, 23) = 3.340, p = .053$ ), with variability in the physiological difference scores accounting for 47.4% of the variability in  $\Delta$ Accuracy ( $R^2 = .474$ ). Increases in skin conductance difference predicted negative change in description accuracy across conditions, but did not contribute significantly to the prediction ( $\beta = -0.033, p = .858$ ). Change in heart rate, however, was a significant positive predictor of change in accuracy ( $\beta = 0.474, p = .016$ ). Based on this regression model, greater changes in heart rate across the conditions (stress > non-stress) predicted greater changes in accuracy across the conditions (non-stress > stress).

The overall regression model for change in description precision was significant ( $F(2, 23) = 3.481, p = .049$ ), with variability in the physiological difference scores accounting for almost 25% of the variability in  $\Delta$ Precision ( $R^2 = .249$ ). Both change in heart rate and change in skin conductance level were negative predictors of change in precision. Heart rate difference did not contribute significantly to the prediction ( $\beta = -0.122, p = .525$ ), but change in skin conductance did ( $\beta = -0.479, p = .019$ ). Based on this regression model, greater changes in SCL across the conditions (stress > non-stress) predicted smaller changes in precision across the conditions (non-stress > stress).

#### *Memory for Detail: Free Recall Questions*

All participants answered two sets of free recall detail questions – one set for the stress session and one set for the non-stress session. The questions covered three categories: (i) Details of the Research Assistant, (ii) Details of the Event, and (iii) Details of the Environment. It was predicted that participants would show better recall for details of the non-stress session than for details of the stress session in every category. Furthermore, it was predicted that recall for central details (relating to the RA and the event) would be less impaired in the stress condition than would recall for peripheral details (relating to the environment). Means and standard deviations are reported in Table 8.

Table 8

#### *Free Recall Question Scores: Descriptive Statistics*

Categories of Detail Recall	Stress (n = 26)		Non-stress (n = 26)	
	Mean	SD <sup>a</sup>	Mean	SD <sup>a</sup>
(i) Research assistant	0.635	0.715	0.635	0.521
(ii) Event	0.442	0.554	0.615	0.668
(iii) Environment	0.365	0.481	0.558	0.622
(iv) Total score	1.442	1.003	1.808	1.184

*Note.* Reported means are untransformed.

<sup>a</sup> Standard deviation

As shown in Table 8, participants achieved higher scores in the non-stress condition for details of the event, details of the environment, and total score. The three detail categories (from which total score is comprised) are represented graphically in Figure 5. The percentage change between the stress and non-stress conditions for each detail category is provided.

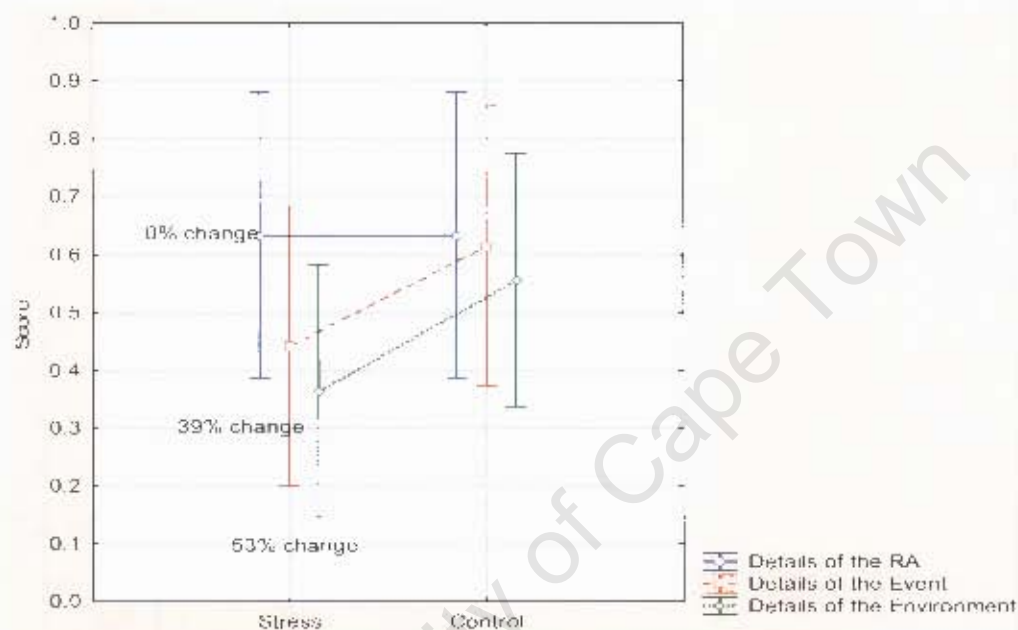


Figure 5. Percentage differences across conditions for free recall questions. Vertical bars denote 95% confidence intervals.

As is apparent in Figure 5, participants performed as well in the stress condition as they did in the non-stress condition for details of the research assistant. The stress condition was, however, associated with a higher standard deviation (as shown in Table 8). The variability in the stress condition was 37% greater than in the non-stress condition, indicating that the stress manipulation differentially affected memory performance for this particular dimension.

Of further interest in Figure 5 is the performance of participants on details of the environment relative to their performance on details of the event and details of the research assistant. For peripheral details, their performance was substantially worse than for central details. While the scores for details of the environment are very low, the variability associated with this dimension is great. As expected, it appears that participants performed particularly poorly and erratically on

peripheral details in general. It was predicted that memory for peripheral details would be more vulnerable to degradation by stress than would memory for central details. Figure 5 does not provide any evidence to suggest this, however: The rate of change for peripheral details is roughly equivalent to that of central details (as indicated by the parallel slopes for details of the environment and details of the event). Thus, the hypothesis that peripheral detail recall would be particularly impaired by stress was not subjected to further testing.

The primary hypothesis that participants would perform better in the non-stress condition was examined for the three detail categories in line with the prediction. Details of the event and details of the environment were tested using a Wilcoxon Matched Pairs Test, while total score was tested with a dependent-samples  $t$  test.

*Test: Wilcoxon Matched-Pairs Signed-Ranks Test.* This nonparametric test was chosen as an alternative to the  $t$  test for dependent samples because it does not have the same stringent requirements concerning normality (even with log-transformations, the scores for both details of the event and details of the environment remained abnormally distributed). Like a dependent-samples  $t$  test, the Wilcoxon test can be used to assess the difference between groups of paired data (scores for the stress condition vs. scores for the non-stress condition) and is more robust than the Sign test.

*Assumptions.* The Wilcoxon test requires that the scale of measurement for the two data sets have the properties of an ordinal scale.

The variables, 'Details of the Research Assistant' and 'Details of the Environment', were measured using an interval scale, which allows for the rank ordering of observations within the stress and non-stress conditions, and also allows for the rank ordering of the difference scores between the stress and non-stress conditions. Thus, the assumption of an ordinal scale is upheld.

*Results.* The result of the test was non-significant both for details of the event ( $T(25) = 21.50$ ,  $Z = 1.022$ ,  $p = .307$ ) and details of the environment ( $T(25) = 19.50$ ,  $Z = 1.200$ ,  $p = .230$ ). The null hypothesis that the distribution of the difference scores between the stress and non-stress

---

condition are symmetric around zero cannot be rejected. This indicates that the stress condition did not differ from the non-stress condition in how it affected participants' memory for these categories of detail. Nonetheless, both details of the event and details of the environment produced medium effects ( $d = -0.41$  and  $d = -0.48$ ), suggesting the presence of a possible trend that this study lacked the power to detect (Power  $(1-\beta) = 0.633$ , where Cohen, 1988, proposed 0.8 as the acceptable level of power for behavioural science research).

*Test: Dependent-samples  $t$  test.* The scores for total detail recall were normally distributed (Stress:  $SW-W = 0.927$ ,  $p = .067$ ; Non-stress:  $SW-W = 0.943$ ,  $p = .158$ ) and were therefore tested using a dependent-samples  $t$  test.

*Results.* The result of the test was non-significant ( $t(25) = -1.516$ ,  $p = .142$ ). The null hypothesis that there was no statistical difference in performance on total detail recall across the stress and non-stress conditions cannot be rejected. However, the total score for details yielded a small to medium effect ( $d = -0.33$ ).

*Test: Regression.* As an alternative approach, a multiple regression analysis was conducted to examine the particular contributions of heart rate and skin conductance to the recall of detail. The change (between the stress and non-stress conditions) for total score was regressed on both the change in heart rate and the change in skin conductance level.

The difference scores for the individual categories of detail recollection (details of the research assistant, event and environment) did not show enough variability to be regressed. Many participants produced difference scores of zero, accounting for the lack of variability. This suggests perhaps that the individual categories of detail recollection (each of which comprised of only three questions) were not sensitive enough to detect differences between the stress and non-stress conditions.

*Assumptions.* The assumption concerning multicollinearity was not violated (tolerance was over 95%). The assumptions of normality, linearity and homoscedasticity of residuals were also upheld (see Appendix M).

---

*Results.*

Table 9

*Regression Results for Total Detail Recollection (where  $\hat{y}$  = the difference in the DV between stress and non-stress)*

	$R^2$	$F$	$p$	df	$\beta_1$ ( $\Delta HR$ )	$p$	$\beta_2$ ( $\Delta SCL$ )	$p$
$\Delta$ Total score	0.25	3.77*	.04	2, 23	0.37*	.05	-0.35	0.07

\* $p \leq .05$

**Positive Predictors**

**Negative Predictors**



*Figure 6.* Graphic representation of the regression of the change in total score for detail recollection.

As shown in Figure 6, increases in heart rate difference predicted positive change in the dependent variable (difference in total score), while increases in SCL difference predicted negative change. Neither the contribution of heart rate difference nor of SCL difference to the prediction was significant at the 5% level ( $\beta_1 = 0.368, p = .054$  and  $\beta_2 = -0.345, p = .069$  respectively), but both contributions approached significance. The effects of heart rate and skin conductance difference are almost equivalent, but they work in opposite directions. Thus, if a participant experienced stress-induced increases in both heart rate and SCL, then we expect no difference in total score across the conditions. If his heart rate increased in the stress session, but his skin conductance remained constant, we predict a greater difference in total score between

the conditions (with non-stress > stress). Alternatively, if his skin conductance increased in the stress session, but his heart rate remained constant, we predict a smaller difference in total score between the conditions (with non-stress > stress). The overall regression for change in total score was significant ( $F(2, 23) = 3.773, p = .038$ ), with variability in the physiological difference scores accounting for 24.7% of the variability in  $\Delta$ total score ( $R^2 = .247$ ).

#### *Memory for Detail: Multiple Choice Questions*

All participants answered two sets of multiple choice detail questions – one set for the stress session and one set for the non-stress session. The questions covered three categories: (i) Details of the Research Assistant, (ii) Details of the Event, and (iii) Details of the Environment. The ‘Details of the Research Assistant’ category will be excluded from the analysis, as it contained only two questions and is therefore of little statistical value.

As with all the other measures of memory performance, it was predicted that participants would perform better for details of the non-stress session than for details of the stress session. Furthermore, it was predicted that memory performance for central details (details of the event) would be less impaired in the stress condition than would memory performance for peripheral details (details of the environment). Finally, it was predicted that participants would score higher in every category for multiple choice than for free recall in both the stress and non-stress conditions. Means and standard deviations are reported in Table 10.

Table 10

#### *Multiple Choice Question Scores for Detail Recollection: Descriptive Statistics*

	Stress (n = 26)		Non-stress (n = 26)	
	Mean	SD	Mean	SD
(i) Event	1.23	0.82	1.31	0.79
(ii) Environment	1.54	1.70	1.19	0.94
(iii) Total score	3.42	1.33	3.31	1.74

*Note.* Reported means are untransformed.

As shown in Table 10, participants achieved higher scores in the non-stress condition only for details of the event. A Wilcoxon test confirmed that this difference was non-significant ( $t(25) = 76.00, Z = 0.024, p = .981$ ) and that the effect produced was negligible ( $d = 0.01$ ). Thus, the null hypothesis that participants would demonstrate no difference in performance between the stress- and non-stress conditions must be accepted. Wilcoxon tests showed that participants' superior performance in the stress condition for details of the environment and total score was non-significant, offering further support for the null hypothesis (Environment:  $t(25) = 66.0, Z = 0.497, p = .619; d = 0.19$ ; Total score:  $t(25) = 79.0, Z = 0.644, p = .520; d = 0.25$ ).

The hypothesis that peripheral detail recall would be particularly degraded by stress must also be rejected, as multiple choice performance for details of the environment was worse in the non-stress condition.

The final prediction – that participants would perform better on the multiple choice than the free recall questions in all categories – was tested using Wilcoxon Matched-Pairs tests and dependent-samples  $t$  tests. Means are reported in Table 11.

Table 11

*Free Recall vs. Multiple Choice Question Scores for Detail Recollection: Descriptive Statistics*

	Stress (n = 26)		Non-stress (n = 26)	
	Free Recall	MCQ	Free Recall	MCQ
(i) Event	0.44	1.23	0.62	1.31
(ii) Environment	0.37	1.54	0.56	1.19
(iii) Total score	1.44	3.42	1.81	3.31

*Note.* Reported means are untransformed.

As shown in Table 11, multiple choice scores were substantially higher than free recall scores in every category, across both the stress and control conditions. The scores for each category are

represented graphically below and discussed in turn.

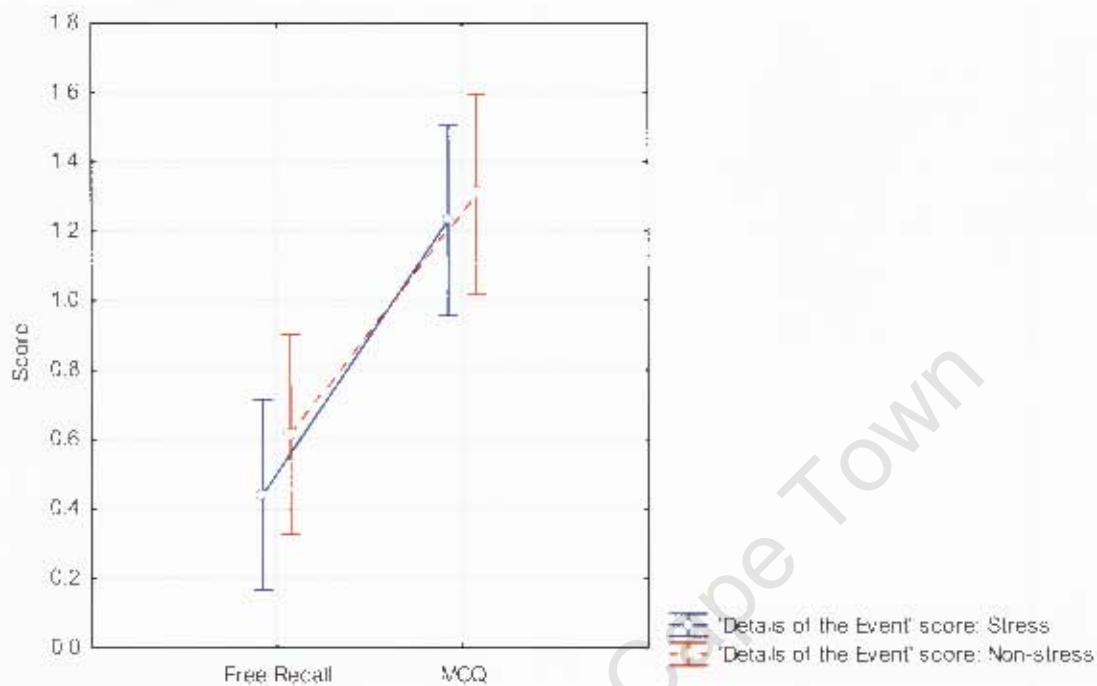


Figure 7. Free recall vs. multiple choice performance for details of the event

Figure 7 confirms that condition (stress vs. control) had no substantial impact on performance for details of the event. There does, however, appear to be a main effect for 'question type' (free recall vs. multiple choice). To investigate this, two Wilcoxon tests were conducted – one within the stress condition and one within the non-stress condition.

*Test: Wilcoxon Matched-Pairs Signed-Ranks Test.* This test was chosen as an alternative to the dependent-samples *t* test, as the scores under consideration were abnormally distributed (even with log-transformations).

*Assumptions.* The assumption of an ordinal scale is upheld (the category, 'Details of the Event', was measured using an interval scale, allowing for the rank ordering of observations).

*Results.* The test produced significant results for the stress condition ( $T(25) = 11.50$ ,  $Z = 3.615$ ,  $p < .001$ ,  $d = 2.01$ ) and the non-stress condition ( $T(25) = 0.00$ ,  $Z = 3.621$ ,  $p < .001$ ,  $d = 2.02$ ). Thus, for details of the event the multiple choice format elicited superior memory performance.

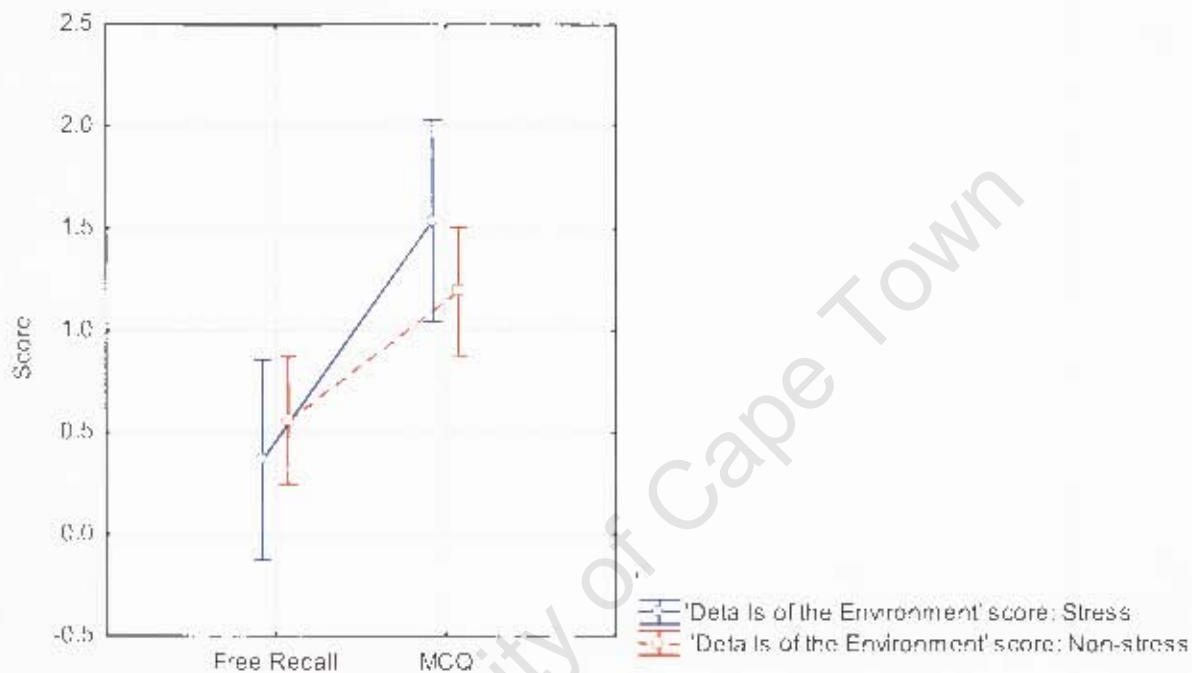


Figure 8. Free recall vs. multiple choice performance for details of the environment

Figure 8 confirms that condition (stress vs. control) had no impact on performance for details of the environment (at least, not in the predicted direction). There does, however, appear to be a main effect for 'question type' (free recall vs. multiple choice). To investigate this, two Wilcoxon tests were conducted – one within the stress condition and one within the non-stress condition.

*Test: Wilcoxon Matched-Pairs Signed-Ranks Test.* This test was chosen as an alternative to the dependent-samples  $t$  test, as the scores under consideration were abnormally distributed (even with log-transformations).

*Assumptions.* The assumption of an ordinal scale is upheld (the category, 'Details of the Environment', was measured using an interval scale, allowing for the rank ordering of observations).

*Results.* The test produced a significant result for the stress condition ( $T(25) = 7.50$ ,  $Z = 3.266$ ,  $p = .001$ ,  $d = 1.67$ ) and for the non-stress condition ( $T(25) = 0.00$ ,  $Z = 3.296$ ,  $p < .001$ ,  $d = 1.69$ ). Thus, for details of the environment the multiple choice question format elicited superior memory performance.

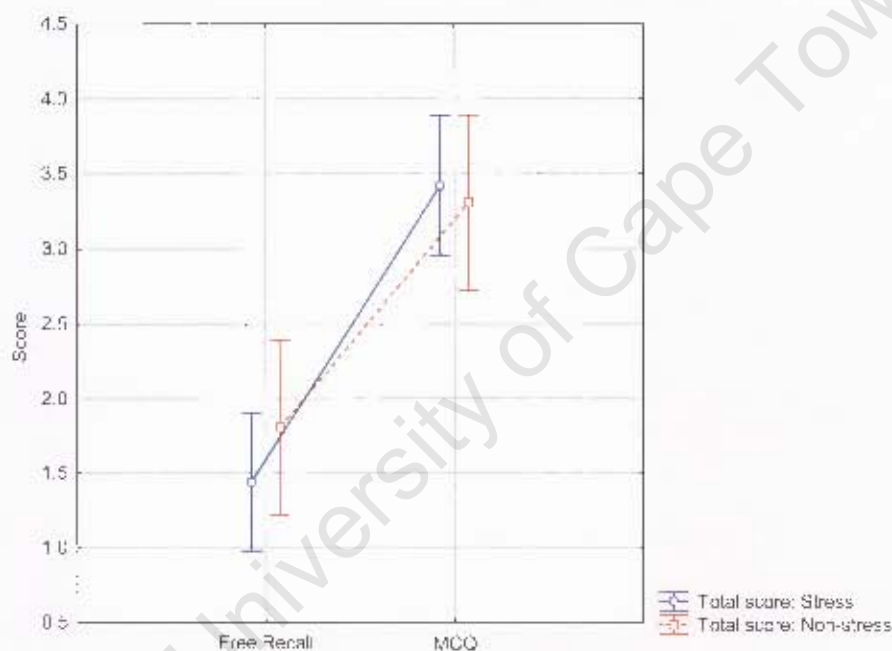


Figure 9. Free recall vs. multiple choice performance for total score

Figure 9 confirms that condition (stress vs. control) had no impact on total detail recollection performance. There does, however, appear to be a main effect for 'question type' (free recall vs. multiple choice).

*Test: Dependent-samples *t* test.* To further investigate the effect of 'question type', two dependent samples *t* tests were conducted – one within the stress condition and one within the non-stress condition.

*Assumptions.* Total Scores for free recall were normally distributed in the stress condition ( $SW-W = 0.927, p = .067$ ) and in the non-stress condition ( $SW-W = 0.943, p = .158$ ). Similarly, total scores for multiple choice were normally distributed in the stress condition ( $SW-W = 0.898, p = .014$ ) and in the non-stress condition ( $SW-W = 0.944, p = 0.167$ ).

*Results.* The test produced significant results and large effect sizes both within the stress condition ( $t(25) = -7.426, p < .001, d = 1.66$ ) and within the non-stress condition ( $t(25) = -6.708, p < .001, d = 0.91$ ). Thus, for total score the multiple choice question format elicited superior memory performance.

In sum, the final prediction – that participants would score higher in every category for multiple choice than for free recall irrespective of condition – was unequivocally supported.

#### *Lineup Identifications*

All participants made two lineup identifications during the memory phase of the experiment – one identification for the research assistant they met during the stress session (henceforth called the *stress RA*), and one identification for the research assistant they met during the non-stress session (henceforth called the *non-stress RA*).

Target-absence versus target-presence was counterbalanced across condition (stress vs. non-stress). Half of the sample saw the target-absent lineup for the stress RA and the target-present lineup for the non-stress RA. The other half saw the target-present lineup for the stress RA and the target-absent lineup for the non-stress RA.

It was predicted that participants would be more accurate in their lineup identifications for the non-stress condition than for the stress condition. It was also predicted that participants would be more accurate in their identifications for target-present lineups than for target-absent lineups. It was thus expected that participants would perform particularly poorly when the condition was ‘stress’ and the lineup type was ‘target absent’.

---

Lineup selections are presented in Table 12 and Table 13. Please note that each participant is represented once in Table 12 and once in Table 13 as each participant saw one target-present lineup and one target-absent lineup.

Table 12

*Lineup Identifications in the Target-Present Condition: Descriptive Statistics*

	True Hits	False Rejections	Foil IDs	Number of Participants
Stress	46.1% (6)	15.4% (2)	38.5% (5)	13
Non-stress	76.9% (10)	15.4% (2)	7.7% (1)	13
Number of Participants	16	4	6	26

*Note.* Frequencies in parentheses.

Table 13

*Lineup Identifications in the Target-Absent Condition: Descriptive Statistics*

	False Positives	Foil IDs	Correct Rejections	Number of Participants
Stress	8.7% <sup>a</sup> (1.13)	60.6% (7.87)	30.7% (4)	13
Non-stress	8.7% <sup>a</sup> (1.13)	60.6% (7.87)	30.7% (4)	13
Number of Participants	2.26	15.74	8	26

*Note.* Frequencies in parentheses.

<sup>a</sup> As there was no designated target replacement in the target-absent condition, the number of error identifications (9 for stress and 9 for non-stress) was divided by the number of lineup members (8). This value (1.13) is reflected as a percentage of the total (13) in the 'False Positives' column. The remaining error identifications (7.87) were treated as foil identifications and are reflected as a percentage of the total (13) in the 'Foil IDs' column.

As shown in Table 12, participants conformed to the prediction for ‘true hits’ and ‘foil identifications’ in the target-present condition. A Comparison of Percentages test confirmed that the higher frequency with which participants selected the target in the non-stress condition was almost significant at the 5% level ( $p = .060$ ). The lower frequency with which they selected foils in the non-stress condition was significant ( $p = .037$ ). As shown in Table 13, participants did not conform to the prediction at all in the target-absent condition, showing identical performance in the stress and non-stress conditions. Thus, the prediction that participants would perform particularly poorly for the stress condition when the lineup-type was ‘target-absent’ was unsupported.

In order to analyse the differences in accuracy across the stress and non-stress conditions, I tabulated percentage accuracy. ‘True hits’ were considered accurate in the target-present condition, and ‘correct rejections’ were considered accurate in the target-absent condition. Comparison of Percentages tests were used to determine the significance of the accuracy statistics reported in Table 14.

---

Table 14

*Lineup Accuracy*

	Target-present	Target-absent	Overall
Stress	46.2% (n = 13)	30.8% (n = 13)	38.5% (n = 26)
Non-stress	76.9% (n = 13)	30.8% (n = 13)	53.8% (n = 26)
Overall	61.5% (n = 26)	30.8% (n = 26)	46.2% (n = 26)

From Table 14 it can be seen that participants were more accurate in the non-stress condition when viewing target-present lineups. This finding is in the predicted direction and, while not quite significant at the 5% level ( $p = .053$ ), it seems to support the hypothesis that stress compromises accuracy. For target-present lineups, the odds ratio indicated that an accurate selection in the stress condition was almost 75% less likely than an accurate selection in the non-stress condition, but was not statistically significant ( $\theta = 0.257, p = .113$ ). Participants performed with equal accuracy in the stress and non-stress conditions when the lineup type was target-absent ( $\theta = 1.00, p = .664$ ) (see Appendix N for the contingency tables from which the odds ratios were calculated).

Participants were generally more accurate for target-present lineups than for target-absent lineups. A Comparison of Percentages test revealed that the difference was not significant for the stress condition ( $p = .214$ ), but it was significant for the non-stress condition ( $p = .013$ ). Calculation of the odds-ratio confirmed that the effect for the stress condition was negligible ( $\theta = 1.929, p = .344$ ) but that the effect for the non-stress condition was substantial ( $\theta = 7.500, p = .024$ ) (see Appendix O for the contingency tables from which the odds ratios were calculated).

*Differential Accuracy.* In order to analyse the overall accuracy of participants, while considering their performance for the stress and the non-stress condition simultaneously, a ‘differential accuracy’ approach was used. Lineup selections for the stress and non-stress conditions were coded, with 1 representing an accurate selection (‘true hit’ or ‘correct rejection’) and 0 representing an inaccurate selection (‘false rejection’, ‘foil identification’, or ‘false positive’). Each participant’s coded score for the non-stress condition was subtracted from his coded score for the stress condition, yielding a ‘differential accuracy’ score. This new variable, differential accuracy, has three levels and is represented in Table 15.

Table 15

*Calculation of Differential Accuracy Scores (where an X represents an accurate selection)*

Coded Score: Stress	subtract	Coded Score: Non-stress	Differential Accuracy Score
<input type="checkbox"/> 0	-	<input checked="" type="checkbox"/> 1	= -1
<input checked="" type="checkbox"/> 1	-	<input checked="" type="checkbox"/> 1	= 0
<input type="checkbox"/> 0	-	<input type="checkbox"/> 0	= 0
<input checked="" type="checkbox"/> 1	-	<input type="checkbox"/> 0	= 1

*Test: Single-sample t test.* The mean differential accuracy score for the sample indicates whether participants were, on the whole, more accurate for the non-stress condition, the stress condition, or neither. The null hypothesis for this test is that there is no difference between the two conditions, and is represented by a mean score of zero. The research hypothesis for this test is that participants were superior in the non-stress condition, and is represented by a negative mean score significantly different from zero.

*Assumptions.* The assumption of normality is not satisfied. Since the variable under consideration is non-continuous (having only three levels), the data does not fall in a Gaussian distribution. Nonetheless, the data is distributed relatively symmetrically around zero, and this is deemed sufficient to proceed with the test. See Appendix P.

*Results.* The differential accuracy mean ( $M = -0.154$ ) does not differ significantly from zero ( $t(25) = -1.07, p = .147$ ). This result is in the predicted direction and, though not statistically significant, it does yield a medium effect ( $d = -0.427$ ).

*Gravity of Error.* Lineup identifications can also be examined from a more applied angle, looking at the implications of participants' lineup choices. A simple analysis of accuracy does not take into account the fact that some identification errors are relatively benign, while others could lead to a grave miscarriage of justice. In order to examine the gravity of error across conditions, I assigned numerical scores to each lineup selection possibility. The positive or negative sign before the score denotes whether the selection could lead to a delivery of justice or a miscarriage of justice respectively. The absolute value relates to the magnitude of the consequences, and is, to a large degree, subjective. A true hit was deemed the best possible outcome, as it could lead to the conviction of a guilty perpetrator. It was thus assigned the highest value (+4). A foil identification was regarded as a slightly negative outcome: Though no one is incriminated by such a selection, it does undermine the credibility of the witness. It was thus assigned a value of -1. A fairly substantial negative value of -3 was assigned for a false rejection. The implication of rejecting a perpetrator-present lineup is the likely exoneration of a guilty culprit. A correct rejection was assigned the highest positive value (+4) because it means that the witness rejected a lineup containing an innocent suspect, thus preventing a miscarriage of justice. Finally, the lowest score (-6) was assigned for a false positive identification: This lineup choice could lead to the prosecution and possible conviction of an innocent suspect and was thus regarded as the most calamitous outcome. Gravity of error scores are presented in Table 16.

---

Table 16

*Gravity of Error for Lineup Selections*

	Stress		Non-stress	
	Frequency	Gravity of Error	Frequency	Gravity of Error
True Hit (+4)	6	24	10	40
Foil ID (-1)	12.87	-12.87	8.87	-8.87
False Rejection(-3)	2	-6	2	-6
Correct Rejection (+4)	4	16	4	16
False Positive (-6)	1.13	-6.78	1.13	-6.78
Mean		0.58		1.35

The mean values in Table 16 indicate that participants committed slightly more grave identification errors in the stress condition, though on the whole, their errors were relatively benign. A single-sample  $t$  test (testing the difference scores between the stress and non-stress sessions against zero) confirmed that this difference was non-significant ( $t(25) = 0.877$ ,  $p = .389$ ). It seems that the presence of stress at exposure does not affect the type of identification errors a witness might commit, but this conclusion must be drawn with caution as a small effect was found in the predicted direction despite the statistical non-significance of the results ( $d = -0.254$ ).

## CHAPTER FOUR: EXPERIMENT 2

### *Method*

#### Rationale, Aim, and Hypotheses

*Rationale.* The first experiment employed a novel manipulation to induce stress in participants: a public speaking task derived from the 'Trier Social Stress Test'. While advantageous because it can be used ethically with juveniles, this manipulation failed to produce a defensive, activation-mode dominant state of stress complete with heightened cognitive anxiety. To establish the utility of using this manipulation in future studies of stress and eyewitness memory, I adjusted the protocol slightly (augmenting the element of social evaluative threat) and applied it to a university sample.

*Aim.* I aimed to establish whether an ethically defensible stress manipulation could produce the same adverse effects on eyewitness memory as the more extreme stress manipulations used in previous studies with young adults.

*Hypotheses.* With regard to the effect of the stress manipulation, I expected that participants would show evidence of activation mode dominance. This was expected to manifest as elevated heart rate, skin conductance and state-anxiety.

With regard to eyewitness memory performance, I expected that participants would perform worse on all eyewitness memory measures following the presence of stress at encoding. Specifically, I predicted that:

---

- 1) Participants' target descriptions of the research assistant met at the stressful exposure session would be less accurate, less precise and less informative than their descriptions of a research assistant met at the non-stressful exposure session.
  
- 2a) Participants would show poorer recall for details of the stressful exposure session than for details of the non-stressful exposure session.
  - b) Recall for central details of the stressful exposure session would be less impaired than recall for the peripheral details of the stressful exposure session.
  
- 3a) Participants would be less accurate in their lineup identification of the research assistant from the stressful exposure session than in their lineup identification of the research assistant from the non-stressful exposure session.
  - b) Participants would be less accurate in their identifications when the lineup type was target-absent than target-present.
  - c) Participants would perform particularly poorly when making identifications in the target absent condition of the research assistant for the stressful exposure session.

### Design

The design in the present experiment is identical to that used in Experiment 1: A repeated measures design with two independent variables (event type and lineup type) and three dependent variables (description performance, detail recollection, and identification accuracy) has been used.

---

## Participants

Twenty six white male undergraduates participated in this study (mean age = 20.4 years), but only 22 were included in the final analysis<sup>11</sup>. The demographic restrictions mentioned in relation to the high school sample applied to the university sample.

Approximately half the participants were approached and recruited on campus, while the other half were recruited via email. Participants were offered R100 reimbursement for attending three sessions over the course of three weeks. In addition, they were told that one participant would win a cash prize of R1500.

## Measures, Instruments and Materials

### *Physiological Measures*

As in Experiment 1, the Ambulatory Monitoring System (AMS) was used for the measurement of bio-signals (Groot, de Geus & de Vries, 1998).

### *Self-reported Anxiety*

As in Experiment 1, the Spielberger State-Trait Anxiety Inventory (STAI) (Spielberger et al., 1983) was used to measure participants' subjective levels of anxiety. This measure has been applied extensively with university students (Spielberger et al., 1983).

### *Psychosocial Stressor*

As in Experiment 1, a public speaking task based on the 'Trier Social Stress Test' (Kirschbaum et al., 1993) was used to induce psychosocial stress in the participants. A slightly modified

---

<sup>11</sup> Four participants had to be excluded from the analysis for the following reasons: the AMS device failed to record any physiological data for one participant; two participants had met the research assistants in a social capacity before enrolling in the experiment; and one subject came to the exposure sessions after exercising, which interfered with his physiological activation.

---

version of the TSST protocol was employed. As in the original TSST, participants were required to speak about their strengths and weaknesses (as though applying for a job), but no cognitive task (e.g. mental arithmetic) was included.

#### *Research Assistants*

The research assistants employed to take participants' physiological measurements and to serve as the target individuals in Experiment 1 fulfilled the same functions in the present experiment.

#### *Evaluators*

Twenty seven psychology undergraduates served as confederate evaluators of participants' performance during the mock job interview. Each confederate attended one of four stress sessions, ensuring that there was an audience of between five and eight evaluators present for each stress session. The confederates participated in return for course credit.

#### *Stimulus Materials*

The lineups used in Experiment 1 were used again in the present experiment.

#### *Eyewitness Tasks*

The paper-and-pencil tasks used in Experiment 1 were used again in the present experiment, but the multiple choice component was excluded. Thus, the university sample completed three eyewitness tasks: a target description, a set of free recall questions and a lineup identification. As in Experiment 1, each task was completed twice, once with reference to the stress session and once with reference to the non-stress session.

---

### *Scoring Materials*

The standardised descriptions of the targets used in Experiment 1 were used again in the present experiment.

### Procedure

The protocol for this study was approved by the Psychology Department Ethics Committee at the University of Cape Town. Data collection took place in the ACSSENT laboratory at the University of Cape Town.

The design for the present experiment was identical to that of Experiment 1. Participants experienced a stress session, a non-stress session and a memory session over three weeks, with the sessions spaced roughly a week apart. The ordering of the stress and non-stress sessions was counterbalanced across participants. Target exposure was also counterbalanced across participants. See Table 17.

Table 17

#### *Counterbalanced Variables in the Design of Experiment 2*

	Week 1		Week 2		Week 3
	Stress	Non-stress	Stress	Non-stress	
Research Assistant 1	n = 6	n = 6	n = 4	n = 6	
Research Assistant 2	n = 6	n = 4	n = 6	n = 6	Memory Phase
Total	12	10	10	12	(all participants)

Prior to commencing the first session of the experiment, the participant was asked to read and sign a consent form. This form contained details of what would be expected of him over the three sessions of the study, the procedure for taking the physiological measurements, his rights as a

research participant and his entitlement to remuneration. The participant was informed that the study was about memory (though not specifically eyewitness memory) and involved a job interview component.

If the participant was attending the non-stress session, I explained that he would be allowed to leave as soon as his physiological measurements had been taken. If he was attending the stress session, however, I warned him that he would have to undergo a job interview with a panel of between five and ten evaluators after his physiological measurements had been taken. The participant was told that each evaluator had a score card and would be rating his performance on a number of dimensions, such as confidence, originality and sincerity. On each dimension, he could be given a rating of 'poor' (1 point), 'moderate' (2 points), 'good' (3 points) or 'excellent' (4 points). The participant was given a score card to peruse so he would be familiar with the criteria on which he was being evaluated. The participant was then told that his eligibility for the R1500 cash prize was contingent on his job interview performance: He would need to achieve an average score of three or above to be considered for the prize. This deception was necessary to intensify his perception of the audience as a social evaluative threat, which is integral to the elicitation of a stress response. The participant was informed that he would have five minutes to convince the audience of why he deserved the job, and would then be questioned by the audience for a further five minutes. This information was always delivered just before the participant entered the lab to experience the session, and was expected to provoke a sense of anticipatory stress which would then be present when he met the research assistant.

In both conditions, the participant was ushered into the lab and was met by one of the target research assistants, who introduced herself by name, made eye contact and shook the participant's hand. This deliberate attempt to establish interpersonal contact aimed to provide all participants with an equal encoding opportunity.

The participant was seated at a desk where the research assistant connected him to the AMS. As in Experiment 1, the research assistant adhered to a strict scripted dialogue while interacting with the participant. This dialogue was identical to that used in Experiment 1 with the exception of the contrived question, which was modified to be more appropriate to a university context. After

---

connecting the participant to the AMS and carrying out the necessary dialogue, the research assistant retired to a part of the lab where she was out of the participant's sight. This concluded the exposure period for the session. Both research assistants always followed the same procedure, which took approximately three minutes, ensuring that target exposure was uniform across participants and conditions.

While attached to the AMS, the participant was left alone to complete a puzzle task (comprising of riddles, anagrams and Sudoku) in the non-stress condition, or to complete the STAI in the stress condition. In the non-stress condition, it was emphasized that he would not be evaluated on his performance of the puzzle task to prevent him from experiencing any anxiety. The participant remained connected to the AMS for five minutes in the non-stress condition (long enough to obtain reliable HR and SCL readings), and for as long as it took him to complete the STAI in the stress condition (approximately five minutes).

After the appropriate amount of time had elapsed, I disconnected the participant from the AMS and allowed him to leave if it was the non-stress session, or sent him to be interviewed if it was the stress session. As in Experiment 1, it was not feasible to leave the participant connected to the AMS during his interview, as this would halve the number of subjects who could participate per session (as it was, each new participant was connected to the AMS as soon as his predecessor was sent to be interviewed). Following his interview (five minutes of speaking and five minutes of questioning), the participant had completed the stress session and was free to go. The participant returned the following week, experienced the opposite condition, and met the other research assistant; the procedure was as described above.

In the third week the participant experienced the memory phase, and no research assistants were present. I administered the three eyewitness tasks to the participant (target description, free recall questions and lineup identification). Each task was completed twice: once for the stress session and once for the non-stress session. The participant was then paid and thoroughly debriefed as to the nature of the experiment and the stress manipulation. He was assured that he would be in the running for the cash prize, irrespective of his job interview performance.

---

## Data Analysis

### *Scoring*

*Target Descriptions.* Using the same procedure and scoring protocol used in Experiment 1, target descriptions were scored by the author and an independent rater. Inter-rater reliability was high for ‘correct items’ (*Pearson’s*  $r_{(RA1)} = .99$ ; *Pearson’s*  $r_{(RA2)} = .85$ ), ‘subjective items’ (*Pearson’s*  $r_{(RA1)} = .87$ ; *Pearson’s*  $r_{(RA2)} = .72$ ) and ‘incorrect items’ (*Pearson’s*  $r_{(RA1)} = .95$ ; *Pearson’s*  $r_{(RA2)} = .81$ ).

Given the strong inter-rater agreement on all item types, I calculated the average between our scores for each target description, and used these in subsequent analyses. These averaged scores were also used to calculate the accuracy for each description.

As in Experiment 1, each participant’s description was also given a score for precision and a score for informativeness. The precision and informativeness scores assigned by three independent raters were averaged to provide the scores used in the analysis. Inter-rater agreement for precision is represented in Table 18 and inter-rater agreement for informativeness is represented in Table 19. In both Table 18 and Table 19, the correlation coefficients for Research Assistant 1 are represented in the bottom left portion of the table, and the correlation coefficients for Research Assistant 2 are represented in the top right portion of the table.

---

Table 18

*Intercorrelations Between Independent Raters of Description Precision*

	Rater 1	Rater 2	Rater 3
Rater 1	—	.94 <sup>b</sup>	.86 <sup>b</sup>
Rater 2	.67 <sup>a</sup>	—	.87 <sup>b</sup>
Rater 3	.80 <sup>a</sup>	.68 <sup>a</sup>	—

<sup>a</sup> Correlation coefficients for Research Assistant 1; Kendall's Coefficient of Concordance = .19

<sup>b</sup> Correlation coefficients for Research Assistant 2; Kendall's Coefficient of Concordance = .16

Table 19

*Intercorrelations Between Independent Raters of Description Informativeness*

	Rater 1	Rater 2	Rater 3
Rater 1	—	.88 <sup>b</sup>	.77 <sup>b</sup>
Rater 2	.79 <sup>a</sup>	—	.80 <sup>b</sup>
Rater 3	.73 <sup>a</sup>	.81 <sup>a</sup>	—

<sup>a</sup> Correlation coefficients for Research Assistant 1; Kendall's Coefficient of Concordance = .42

<sup>b</sup> Correlation coefficients for Research Assistant 2; Kendall's Coefficient of Concordance = .24

*Detail Questions.* One point was assigned per correct answer in each category ('Details of the Research Assistant', 'Details of the Event', and 'Details of the Environment'). Points were summed to yield a total per category and an overall total.

*Lineup Identifications.* Participants' selections were coded as a 'correct identification', 'foil identification' or 'false rejection' in the target-present condition, and as a 'correct rejection' or 'false positive' in the target-absent condition.

#### *Statistical Analysis*

For all statistical analyses, I used the software package Statistica Version 8 (StatSoft, 2007). The threshold for statistical significance was set at  $\alpha = 0.05$ .

University of Cape Town

---

## Results

### Measures of Stress

#### *Heart Rate*

Heart rate readings were taken from all participants at the stress session and the non-stress session. As predicted, mean heart rate was higher for the stress session ( $M = 95.55$ ,  $SD = 17.88$ ) than for the non-stress session ( $M = 82.45$ ,  $SD = 13.84$ ). A dependent-samples  $t$  test was conducted to establish whether the mean difference between the conditions (13.1 beats per minute) was significant. The assumption of normality was met in the non-stress condition ( $SW-W = 0.972$ ,  $p = .749$ ), but not in the stress condition ( $SW-W = 0.887$ ,  $p = .017$ ). As is typical with heart rate data, the distribution of heart rate readings in the stress condition was leptokurtic and positively skewed (Jennings, Stringfellow, & Graham, 1974). The natural log of the data was taken, yielding a more normal distribution ( $SW-W = 0.914$ ,  $p = .056$ ) (see Appendix Q). All subsequent analyses were conducted with the log-transformed values for both the stress and non-stress data, but the reported means are untransformed. The results from the dependent-samples  $t$  test confirmed that the stress manipulation significantly increased heart rate in the sample ( $t(25) = 4.03$ ,  $p = .001$ ), yielding a large effect ( $d = 0.83$ ).

#### *Skin Conductance Level*

Skin conductance readings were taken from all participants at the stress session and the non-stress session. It was predicted that mean skin conductance would be higher for the stress condition than for the non-stress condition. Contrary to the prediction, mean skin conductance level was slightly lower for the stress session ( $M = 8.97$ ,  $SD = 4.52$ ) than for the non-stress session ( $M = 9.58$ ,  $SD = 3.91$ ). A dependent-samples  $t$  test confirmed that this difference was non-significant ( $t(21) = -0.799$ ,  $p = .433$ ,  $d = -0.18$ ), supporting the null hypothesis. The skin conductance readings for the stress condition, though lower on average than for the non-stress

---

condition, showed greater variability ( $SD = 4.52$ ). This statistic indicates that the stress manipulation varied substantially in how much it affected participants.

*Autonomic Reactivity: The Relationship between Heart Rate and SCL*

Subjects' mean heart rate readings for the stress condition were plotted against their mean skin conductance readings for the stress condition, and likewise for the non-stress condition. The stress condition revealed no relationship between heart rate and skin conductance ( $r = -.56$ ), while the non-stress condition yielded a small positive correlation ( $r = .284$ ). Given the poor relationship between heart rate and skin conductance within the experimental conditions, it is no surprise that 'change in heart rate' and 'change in skin conductance' appear to be unrelated ( $r = -.087, p = .701$ ). As in Experiment 1, multiple regression was chosen as the primary statistical test used in the analysis for this study, as it does not rely on parallel variation between the two autonomic indicators.

*Self-reported Anxiety*

All participants completed the State-Trait Anxiety Inventory at the stress session. It was predicted that mean state-anxiety would be significantly elevated as a result of the stress manipulation. Means and standard deviations are reported in Table 20.

---

Table 20

*State- and Trait-Anxiety: Descriptive Statistics*

	Sample Mean (n = 22)	Sample SD <sup>a</sup> (n = 22)	Norm Mean (N = 324)	Norm SD <sup>a</sup> (N = 324)
State-Anxiety	39.23	8.95	36.47	10.02
Trait-Anxiety	36.59	7.16	38.30	9.18
Difference	2.64	6.95	-1.83	-

<sup>a</sup> Standard deviation

A single-sample *t* test was conducted to establish whether the difference between state- and trait-anxiety was greater for the sample than for the population. If this were the case, it would confirm that participants were experiencing subjective stress. The population difference score between state- and trait-anxiety ( $\mu = -1.83$ ) was set as a reference constant against which the difference scores for the sample ( $M = 2.64$ ) were tested. The difference scores for the sample were normally distributed ( $SW-W = 0.938$ ,  $p = .179$ ). The result of the *t* test was significant ( $t(21) = 3.01$ ,  $p = .007$ ), producing a large effect ( $d = 1.31$ ). The null hypothesis—that the sample is representative of an unstressed population—can be decisively rejected. The experimental manipulation clearly succeeded in provoking subjective stress in the sample.

*Overview*

Heart rate, skin conductance level and self-reported anxiety were measured as validity checks on subjects' state of arousal. Heart rate was significantly elevated in the stress condition, but skin conductance level was equivalent across conditions. An explanation for the latter result is offered

in the Discussion. Participants reported heightened state-anxiety at the stress session, indicating that they experienced subjective anxiety in anticipation of the job interview.

### Measures of Memory

#### *Target Descriptions*

All participants completed target descriptions for the stress and non-stress sessions. Descriptions were scored on four dimensions: (i) total number of items correct, (ii) accuracy, (iii) precision and (iv) informativeness. It was predicted that participants' target descriptions for the non-stress session would contain more correct items and would be more accurate, precise and informative than those for the stress session. Means and standard deviations are reported in Table 21.

Table 21

#### *Target Description Scores: Descriptive Statistics*

Dimensions of Description Performance	Stress (n = 22)		Non-stress (n = 22)	
	Mean	SD <sup>a</sup>	Mean	SD <sup>a</sup>
(i) Total correct items	1.86	1.41	1.88	1.07
(ii) Accuracy	56.82	32.56	54.76	26.64
(iii) Precision	2.72	1.21	2.91	0.86
(iv) Informativeness	2.75	1.18	2.89	0.78

*Note.* The means reported here are untransformed.

<sup>a</sup> Standard deviation

As shown in Table 21, description scores for the non-stress condition were elevated above scores for the stress condition on three of the description dimensions (total correct items, precision and informativeness). For accuracy, however, participants performed slightly better in the stress

condition. Of interest is the difference in variability between the conditions: On all four dimensions, the standard deviation in the stress condition exceeds that in the non-stress condition (by more than 60% for total correct items and accuracy, and more than 70% for precision and informativeness). The implication of this is that participants' performances were more uniform in the non-stress condition and more erratic in the stress condition.

*Test: Single-sample  $t$  test.* For each dimension where the direction of the difference was as predicted (total correct items, precision and informativeness), a single sample-sample  $t$  test was conducted. This test was chosen because it eliminates any problems associated with unequal variance. On each dimension, the difference score between the stress and non-stress condition was compared against 0.

*Assumptions.* The assumption of normality was met for precision ( $SW-W = 0.985, p = .975$ ) and informativeness ( $SW-W = 0.987, p = .986$ ), but not for total items correct ( $SW-W = 0.889, p = .018$ ). For this dimension, one outlier was removed, yielding a more normal distribution ( $SW-W = 0.933, p = .161$ ) (see Appendix R). The outlying score was excluded from the subsequent analysis.

*Results.* Single-sample  $t$  tests confirmed that there was no significant difference in description performance on any of the dimensions. In support of the null hypothesis, participants provided descriptions that were as correct ( $t(20) = 0.998, p = .330$ ), precise ( $t(21) = 0.613, p = .547$ ) and informative ( $t(21) = 0.481, p = .635$ ) for the stress session as for the non-stress session. Small effects in the predicted direction were produced for total correct items ( $d = -0.23$ ), precision ( $d = -0.18$ ) and informativeness ( $d = -0.14$ ).

*Test: Regression.* As an alternative approach, a multiple regression analysis was conducted to examine the particular role that heart rate and skin conductance played in this pattern of results. The change (between the stress and non-stress conditions) for each dimension of description performance was regressed on both the change in heart rate and the change in skin conductance level.

---

*Assumptions.* The assumption concerning multicollinearity was not violated (tolerance was over 95%) and the assumptions of normality, linearity and homoscedasticity of residuals were upheld for all dimensions (see Appendices S, T, U and V).

*Results.*

Table 22

*Regression Results for Target Description Dimensions (where  $\hat{y}$  = the difference in the DV between stress and non-stress)*

	$R^2$	$F$	$p$	df	$\beta_1$ ( $\Delta HR$ )	$p$	$\beta_2$ ( $\Delta SCL$ )	$p$
$\Delta$ Total correct items	0.04	0.36	0.70	2,19	-0.12	0.64	0.15	0.52
$\Delta$ Accuracy	0.12	1.30	0.30	2,19	-0.11	0.61	0.32	0.16
$\Delta$ Precision	0.12	1.25	0.31	2,19	-0.20	0.36	-0.29	0.19
$\Delta$ Informativeness	0.07	0.71	0.51	2,19	-0.13	0.57	-0.24	0.29

\* $p \leq .05$

From Table 22, it can be seen that no significant regression models emerged for description performance. On all dimensions, variability in the physiological measures accounted for less than 13% of the variability in the dependent variables. In sum, neither change in heart rate nor change in skin conductance level contributed substantially to change in description performance on any dimension.

*Memory for Detail: Free Recall*

All participants answered two sets of detail questions – one set for the stress session and one set for the non-stress session. The questions covered three categories: (i) Details of the Research Assistant, (ii) Details of the Event, and (iii) Details of the Environment. It was predicted that participants would show better total recall for details of the non-stress session than for details of the stress session. Furthermore, it was predicted that recall for central details (relating to the research assistant and the event) would be less impaired in the stress condition than would recall for peripheral details (relating to the environment). Means and standard deviations are reported below.

Table 23

*Free Recall Scores for Detail Recollection: Descriptive Statistics*

Categories of Detail Recall	Stress (n = 22)		Non-stress (n = 22)	
	Mean	SD <sup>a</sup>	Mean	SD <sup>a</sup>
(i) Research assistant	0.59	0.50	0.61	0.53
(ii) Event	0.43	0.64	0.27	0.55
(iii) Environment	0.05	0.21	0.02	0.11
(iv) Total score	1.07	0.92	0.91	0.70

*Note.* The means reported here are untransformed.

<sup>a</sup> Standard deviation

As shown in Table 23, mean scores for the non-stress condition were higher than mean scores for the stress condition only for details of the research assistant. A Wilcoxon Matched Pairs Test confirmed that the difference of 0.023 was non-significant ( $T(22) = 9.00$ ,  $Z = 0.314$ ,  $p = .377$ ;  $d = 0.13$ ).

For details of the event and the environment, as well as total score, participants performed better in the stress condition. Wilcoxon tests confirmed the non-significance of these differences for details of the event ( $T(22) = 21.00$ ,  $Z = 1.067$ ,  $p = .286$ ) and total score ( $T(22) = 52.50$ ,  $Z = 0.426$ ,  $p = .670$ ). A Sign Test confirmed the non-significance of the difference for details of the environment ( $Z = -0.707$ ,  $p = .480$ ). The effect for total score was small ( $d = 0.18$ ), but details of the event and details of the environment yielded small to medium effects ( $d = 0.47$  and  $d = 0.31$ ). Since recall performance for details of the environment is *worse* in the non-stress condition, the hypothesis that peripheral detail recall would be particularly degraded by stress is clearly unsupported.

*Test: Regression.* A multiple regression analysis was conducted to examine the particular contributions of heart rate and skin conductance to recall of detail. The change (between the stress and non-stress conditions) for total score was regressed on both the change in heart rate and the change in skin conductance level.

The difference scores for the individual categories of detail recollection (details of the research assistant, details of the event and details of the environment) did not show enough variability to be regressed. Many participants produced difference scores of zero, accounting for the lack of variability.

*Assumptions.* The assumption concerning multicollinearity was not violated (tolerance was over 95%) and the assumptions of normality, linearity and homoscedasticity of residuals were upheld (see Appendix W).

---

*Results.*

Table 24

*Regression Equation for Total Detail Recollection (where  $\hat{y}$  = the difference in the DV between stress and non-stress)*

	$R^2$	$F$	$p$	df	$\beta_1$ ( $\Delta$ HR)	$p$	$\beta_2$ ( $\Delta$ SCL)	$p$
$\Delta$ Total score	0.13	1.35	0.28	2, 19	0.27	0.22	-0.20	0.36

\* $p \leq .05$

As shown in Table 24, the regression model for change in total score was non-significant: Neither change in heart rate nor change in skin conductance level accounted for substantial variability in the change of detail recollection performance.

*Lineup Identifications*

As in Experiment 1, it was predicted that participants would be more accurate in their lineup identifications for the non-stress condition than for the stress condition. It was also predicted that participants would be more accurate in their identifications for target-present lineups than for target-absent lineups. It was thus expected that participants would perform particularly poorly when the condition was 'stress' and the lineup type was 'target absent'.

The counterbalancing across lineup type (target-present vs. target-absent) and condition (stress vs. non-stress) was slightly unsymmetrical, owing to the necessary exclusion of certain participants from the analysis. Of the included sample, 55% saw the target-absent lineup for the stress RA and the target-present lineup for the non-stress RA, while 45% saw the target-present lineup for the stress RA and the target-absent lineup for the non-stress RA.

Lineup selections are presented in Table 25 and Table 26. Please note that each participant is represented once in Table 25 and once in Table 26 as each participant saw one target-present lineup and one target-absent lineup.

Table 25

*Lineup Identifications in the Target-Present Condition: Descriptive Statistics*

	True Hits	False Rejections	Foil IDs	No. of participants
Stress	50% (5)	20% (2)	30% (3)	10
Non-stress	33.3% (4)	41.7% (5)	25% (3)	12
No. of participants	9	7	6	22

*Note.* Frequencies in parentheses.

Table 26

*Lineup Identifications in the Target-Absent Condition: Descriptive Statistics*

	False Positives	Foil IDs	Correct Rejections	No. of participants
Stress	8.3% <sup>a</sup> (1)	58.3% (7)	33.3% (4)	12
Non-stress	10% <sup>a</sup> (1)	70% (7)	20% (2)	10
No. of participants	2	14	6	22

*Note.* Frequencies in parentheses.

<sup>a</sup> As there was no designated target replacement in the target-absent condition, the number of error identifications (8 for stress and 8 for non-stress) was divided by the number of lineup members (8). This value (1) is reflected as a percentage of the total in the 'False Positives' column (e.g. 1/12). The remaining error identifications were treated as foil identifications and are reflected as a percentage of the total in the 'Foil IDs' column.

As shown in Table 25, participants conformed to the prediction only for ‘foil identifications’ in the target-present condition. A Comparison of Percentages test confirmed that their less frequent selection of foils in the non-stress condition was non-significant ( $p = .398$ ). As shown in Table 26, participants did not conform to the prediction at all in the target-absent condition: They showed superior performance for the stress RA, committing fewer false positive errors and foil identifications, and making more correct rejections. Thus, the prediction that participants would perform particularly poorly for the stress condition when the lineup-type was ‘target-absent’ was clearly unsupported.

In order to analyse the differences in accuracy across the stress and non-stress conditions, percentage accuracy has been tabulated in Table 27. ‘True hits’ were considered accurate in the target-present condition, and ‘correct rejections’ were considered accurate in the target-absent condition.

Table 27

*Lineup Accuracy*

	Target-present	Target-absent	Overall
Stress	50% (n = 10)	33.3% (n = 12)	40.9% (n = 22)
Non-stress	33.3% (n = 12)	20% (n = 10)	27.3% (n = 22)
Overall	40.9% (n = 22)	27.3% (n = 22)	34.1% (n = 22)

From Table 27 it can be seen that participants made more accurate lineup identifications in the stress condition, irrespective of whether the lineup was target-present or target-absent. This

finding was statistically non-significant for both target-present lineups ( $p = .214$ ) and target-absent lineups ( $p = .243$ ). Calculation of the odds-ratios confirmed the non-significance of these differences (Target-presence:  $\theta = 2.00$ ,  $p = .890$ ; Target-absence:  $\theta = 2.00$ ,  $p = .881$ ) (see Appendix X for the contingency tables from which the odds ratios were calculated).

Participants were generally more accurate for target-present lineups than for target-absent lineups in both the stress condition ( $p = .214$ ) and the non-stress condition ( $p = .251$ ), but the effects for these results were inconsequential (Stress:  $\theta = 2.00$ ,  $p = 0.361$ ; Non-stress:  $\theta = 2.00$ ,  $p = .417$ ) (see Appendix Y for the contingency tables from which the odds ratios were calculated).

*Differential Accuracy.* As in Experiment 1, a ‘differential accuracy’ approach was used to analyse the overall accuracy of participants, while considering their performance for the stress and the non-stress condition simultaneously.

*Test: Single-sample t test.* The mean differential accuracy score for the sample indicates whether participants were, on the whole, more accurate for the non-stress condition, the stress condition, or neither. The null hypothesis for this test is that there is no difference between the two conditions, and is represented by a mean score of zero. The research hypothesis for this test is that participants were superior in the non-stress condition, and is represented by a negative mean score significantly different from zero.

*Assumptions.* The assumption of normality is not satisfied. Since the variable under consideration is non-continuous (having only three levels), the data does not fall in a Gaussian distribution. Nonetheless, the data is distributed relatively symmetrically around zero, and this is deemed sufficient to proceed with the test. See Appendix Z.

*Results.* The differential accuracy mean ( $M = 0.136$ ) does not differ significantly from zero ( $t(21) = 1.00$ ,  $p = .329$ ,  $d = 0.44$ ). From this, it seems that lineup identification performance was unaffected by the stress manipulation.

---

*Gravity of Error.* The same ‘gravity of error’ values employed in Experiment 1 were applied to participants’ lineup choices in the present experiment. Gravity of error scores are presented in Table 28.

Table 28

*Gravity of Error for Lineup Selections*

	Stress		Non-stress	
	Frequency	Gravity of Error	Frequency	Gravity of Error
True Hit (+4)	5	20	4	16
Foil ID (-1)	10	-10	10	-10
False Rejection(-3)	2	-6	5	-15
Correct Rejection (+4)	4	16	2	8
False Positive (-6)	1	-6	1	-6
Mean		0.64	22	-0.32

The mean values in Table 28 indicate that participants committed more grave identification errors in the non-stress condition, but a single-sample *t* test (testing the difference scores between the stress and non-stress sessions against zero) confirmed the non-significance of this difference ( $t(21) = -1.316, p = .202, d = 0.57$ ). Thus, the null hypothesis—that the gravity of participants’ errors is unaffected by stress at encoding—cannot be rejected.

## DISCUSSION

In the present study, I aimed to ascertain the effect of stress at the time of exposure to an unknown individual on eyewitness memory. This aim was applied to an adolescent sample in the first experiment because adolescents have been underrepresented in eyewitness research (Pozzulo & Lindsay, 1998). The aim was also applied to a young adult sample (Experiment 2), because I wanted to establish the effectiveness of the chosen stressor in replicating the results of previous studies. Based on the robust finding that stress impairs eyewitness memory in adults (Deffenbacher et al., 2004), I predicted that participants would show impoverished eyewitness memory performance following a psychosocial stress manipulation. Adolescents are as competent as adults on most eyewitness tasks (Pozzulo & Lindsay, 1998), so it seemed reasonable to expect that they would perform like adults in response to stress at encoding. The predictions for the study were based on empirical findings, but were also underpinned by Deffenbacher's integrative theory of stress and memory (1994). According to this theory, a combination of moderate physiological activation and cognitive anxiety produce dramatic decrements in memory.

I will first discuss the success of the stress manipulation as indicated by measures of physiological arousal and cognitive anxiety in participants. I will then address each dependent variable in turn, discussing whether the findings conformed to my predictions.

### Statistics: Summary and Discussion

#### *Elicitation of a Stress Response*

Within the adolescent sample, the public speaking stressor produced a modest elevation in both heart rate and skin conductance. Such autonomic responses are characteristic of activation mode

---

dominance (Deffenbacher, 1994), signifying that stress has been elicited. Participants' responses to the State-Trait Anxiety Inventory did not confirm this, however. The mean state-anxiety score for the sample did not differ from that of an unstressed population, implying an absence of cognitive anxiety. It is possible that participants' answers to the STAI were not a true reflection of their subjective stress levels. Many participants obtained scores for both state- and trait-anxiety that were below the norms for their age-group, suggesting that a social desirability effect might have been at play.

In the second experiment, cognitive anxiety was successfully induced by increasing the consequences associated with the public speaking task and heightening the social evaluative threat. The mean state-anxiety score for the young adult sample was significantly higher than that of an unstressed population. The modified stressor successfully produced cognitive anxiety, but it also produced a puzzling set of physiological data. While the mean heart rate was dramatically increased by the stress manipulation, the mean skin conductance level showed no elevation.

A lack of parallel variation between physiological measures is not uncommon (Lazarus, Speisman, & Mordkoff, 1963). Indeed, Lacey and Lacey (1958) found various physiological measures to be remarkably independent of one another, pointing out that extensive "studies of relationships among autonomic measures show that no single physiological measure correlates well with others" (p. 67). Skin conductance appears to be the most independent of the autonomic responses, despite being one of the most popular indices of autonomic activation (Lacey & Lacey, 1958). The negligible relationship between heart rate and skin conductance levels in the stress condition can best be explained by the principle of *autonomic response-stereotypy*. According to this principle, individuals respond to different stressors with a consistent, idiosyncratic pattern of autonomic activation: They tend to respond with maximal activation in one autonomic response channel (heart rate *or* skin conductance, for example) (Lacey, Bateman, & Van Lehn, 1953; Lacey & Lacey, 1958). These individualised patterns of response are typically hierarchical, with subjects exhibiting over-activation in some physiological measures, average activation in others, and under-activation in still others (Lacey & Lacey, 1958). Of the

---

young adults, three participants responded to the stressor with elevations in skin conductance only, while fifteen responded with elevations in heart rate only. Thus, the majority of the sample (68%) responded to the stressor with significant elevation in only one of the two autonomic response channels. This finding does not imply the absence of stress, but rather represents an idiosyncratic pattern of autonomic activation in accord with the principle of autonomic response-stereotypy.

In sum, the measures of stress did not behave entirely as expected in either experiment. Nonetheless, the substantial effects for both physiological measures in the adolescent study, and for heart rate and subjective anxiety in the young adult study, indicate that participants *were* affected by the stress manipulation. The lack of mean skin conductance increase in Experiment 2 may simply represent pervasive under-activation in one autonomic response channel. Even so, an incomplete picture of stress emerged from each study, warranting a cautious interpretation of performance on the eyewitness memory measures.

### *Target Descriptions*

I predicted that participants would provide descriptions that contained fewer correct items, were less accurate, less precise and less informative for the research assistant from the stress session than for the research assistant from the non-stress session. Each of these hypotheses was tested individually. In the adolescent study (Experiment 1), the null hypothesis could only be decisively rejected for informativeness: Participants' performance on this dimension was significantly inferior in the stress condition, producing a small effect. It would be premature to draw any conclusions from this finding alone, but, coupled with the small effect found for precision in the predicted direction, it provides grounds for further investigation into the effect of stress on adolescents' person descriptions. In the young adult study (Experiment 2), the magnitude of the difference between the stress and non-stress conditions was not significant for any of the dimensions; thus, the null hypothesis could not be rejected.

---

A regression analysis was included to highlight the role of autonomic activation in predicting eyewitness performance on person description measures. The change for each dimension of description performance (between the stress and non-stress sessions) was regressed on the changes in heart rate and skin conductance level (between the stress and non-stress sessions). Given that physiological activation served as an index to participants' state of stress, I expected increases in the physiological measures to predict decreases in description performance. This expectation was not met in the young adult study (Experiment 2), but it was partially met in the adolescent study (Experiment 1).

In Experiment 1, changes in heart rate and skin conductance together contributed substantially to changes in the precision and accuracy of participants' target descriptions. These contributions were statistically significant, with physiological activation explaining 25% of variability in precision and 47% of variability in accuracy. These findings point to the existence of a relationship between stress and target description performance in adolescents. I anticipated that this relationship would be negative, based on findings from the literature (e.g. Deffenbacher et al., 2004) and the theoretical formulation discussed earlier (Deffenbacher, 1994). As expected, increases in heart rate predicted decreased description accuracy. Contrary to what I expected, however, increases in skin conductance level predicted *enhanced* precision. The incongruity of these findings can be explained in terms of the lack of agreement between the two physiological measures. Consistent with the literature on stress-induced physiological activation, it seems that heart rate and skin conductance did not operate in concert in response to the stressor, accounting for inconsistencies in the regression analysis.

### *Interrogative Recall*

I predicted that participants would show poorer recall for details of the stress session than for details of the non-stress session. This pattern was expected to apply to both central and peripheral details, but the contrast was expected to be greater for peripheral than for central details. The latter hypothesis received no support from either experiment.

---

Free recall for details generally conformed to the predicted pattern in the adolescent sample (Experiment 1). In three of the four detail categories, recall was better for the non-stress session. Details of the event, details of the environment and total detail recall produced small to medium effects in the predicted direction. Despite the lack of statistical significance, the presence of such effects suggests that a significant trend might be detected with a larger sample. This finding was not replicated in the second experiment. For the young adults, detail recall for the stress session was, in fact, superior, yielding small to medium effects for both details of the event and details of the environment.

In Experiment 1, the multiple choice question format elicited better memory performance than the free recall format, as anticipated. This is evident from the large effect sizes associated with each type of detail recall, and indicates that poor memory retrieval was the primary source of error in detail recollection. Counter to expectations, however, was the finding that multiple choice performance did not mirror free recall performance, but rather conformed to the null hypothesis that there would be no difference between the stress and non-stress condition. This finding can perhaps be explained by the tendency of participants to guess when confronted with multiple choice options. In response to the free recall questions, participants tended to fill in only the answers they knew (which were almost always correct) and to leave the rest blank; in response to the multiple choice questions, they tended to select an answer for every question, guessing when they were uncertain. Despite the use of negative marking, such guessing behavior reduced the diagnostic value of the multiple choice question measure and perhaps caused differences between the stress and non-stress conditions to be obscured.

For the adolescent sample (Experiment 1), the regression analysis revealed a strong association between autonomic activation and memory for detail. Together, heart rate and skin conductance contributed significantly to changes in total detail recall, explaining 25% of the variability in this memory measure. This finding provides evidence of a relationship between stress and memory for detail in adolescent eyewitnesses. In accord with the overarching hypothesis for this study (that stress degrades eyewitness memory), I expected this relationship to be negative. This was indeed the case for change in heart rate (between the stress and non-stress sessions). Increases in

---

heart rate predicted decreases in detail recollection. Contrary to expectations, however, increases in skin conductance level predicted *improved* detail recollection. As with the regression results for the target descriptions, this anomalous finding can be explained by the lack of covariance between the two physiological measures of stress.

In contrast to the adolescent study (Experiment 1), the regression analysis from the young adult study (Experiment 2) produced no evidence of a relationship between physiological activation and detail recollection.

#### *Lineup Identifications*

Firstly, I predicted that participants would be more accurate in their lineup identifications for the non-stress condition than for the stress condition. Secondly, I predicted that more accurate identifications would be made from target-present lineups than from target-absent lineups. Thirdly, I predicted that participants would perform particularly poorly when the condition was 'stress' and the lineup type was 'target absent'.

The first prediction gained no support from either experiment. In both the adolescent study and the young adult study, the odds ratios for lineup selections were statistically non-significant. Participants were as likely to make an accurate identification in the stress condition as in the non-stress condition, irrespective of whether the lineup type was target-absent or target-present. The second prediction gained partial support: The direction of superiority was as predicted in both experiments, with participants generally demonstrating more accuracy when viewing target-present lineups than target-absent lineups. However, this finding was only statistically significant for the non-stress condition in the adolescent study. These two predictions formed the basis of the third. In line with the overarching hypothesis of the study, I expected participants to perform worse in the stress condition than in the non-stress condition (Prediction 1), and in line with the

---

eyewitness literature (e.g. Lindsay et al., 1997), I expected them to perform worse in the target-absent condition than the target-present condition (Prediction 2). I therefore anticipated that participants would be least accurate when the condition was 'stress' and the lineup type was 'target-absent' (Prediction 3). Given that Prediction 1 received no support and Prediction 2 received limited support, Prediction 3 was rendered untenable. Indeed, adolescents performed with equal accuracy in the stress and non-stress conditions when the lineup type was 'target-absent' and the young adults were least accurate in the *non-stress* condition when the lineup type was 'target-absent'.

In Experiment 1, the pattern of accuracy obtained by the adolescents replicated the pattern found in previous studies of eyewitness performance following stress at encoding. When the target was present in the lineup, the hit rate was higher for the non-stress research assistant than for the stress-related research assistant. When the target was absent from the lineup, however, the correct response rate was equivalent across conditions. In Peters' child study (1987), participants made significantly more correct responses for the neutral target than for the stress-related target only when the lineup was target-present. When participants were shown a blank lineup, no difference in accuracy between the targets emerged. Similarly, Peters' (1988) young adult participants were more accurate in identifying the researcher than the nurse who administered the inoculation, but only in the target-present condition; a null difference was found for target-absent lineups. These results are in keeping with Deffenbacher et al.'s (2004) meta-analytic finding: A substantial effect was produced for target-present lineups, but a negligible effect was found for target-absent lineups, leading the authors to conclude that "the correct rejection rate for TA lineups was unaffected by stress level" (p. 695). Though the lineup identification results for the present study were not significant, the performance patterns resemble those found by previous researchers, suggesting promise for further research with larger adolescent samples.

Relative to the adolescents, the young adult participants showed deflated accuracy performance (62% [adolescents] vs. 41% [young adults] overall across the stress and non-stress conditions). One possible explanation for this is that the young adults' memory representations of the targets suffered from retroactive interference. Upon walking out of the laboratory, they presumably

---

encountered dozens of unknown females within the same age range as the target research assistants, seeing as data collection took place at the busiest time of day on campus. Owing to the interference of co-occurrent traces, it is possible that participants' memory representations of the targets' faces were never properly consolidated. This explanation makes sense in light of the higher accuracy of the adolescent sample: In Experiment 1, data collection took place at a single-sex high school campus, where participants were unlikely to encounter female faces other than those highly familiar to them (such as teachers), preventing the likelihood of retroactive interference. Alternatively, it is possible that the adolescent participants identified the targets more accurately because they attended to them more closely during the interaction. This may have been a consequence of the novelty of encountering young women in a predominantly male environment, or perhaps because the schoolboys took the experiment more seriously than did the university participants.

### *Summary*

Taken as a whole, the findings from the adolescent study (Experiment 1) are somewhat reconcilable with the hypothesis of this thesis. To the extent that autonomic activation represents a state of stress, the regression models for target descriptions and detail recollection indicate a relationship between stress and eyewitness memory. The effects for description precision and informativeness reinforced this finding, as did the results from comparisons of detail recall. In general, inferior performance was associated with the stress condition. This was not extended to lineup identifications, however: There was no significant difference between participants' performance for the stress condition versus the non-stress condition. The null hypothesis for lineup identifications must be accepted warily, however, as the likelihood of committing a Type II error was high (Power  $(1-\beta) = 0.49$ , where the desired power value is conventionally set at 0.8, as proposed by Cohen, 1988). Further research on this measure of eyewitness performance in adolescents is advisable. In sum, it seems likely that stress compromises the quality of person descriptions and memory for detail in adolescents, but future studies with larger samples are needed to confirm this.

---

It can be stated unequivocally that the findings from the young adult study (Experiment 2) do not support the hypothesis of this thesis. None of the dependent variable measures produced evidence to suggest that stress impairs eyewitness memory. To revisit the original rationale for conducting the young adult study, I aimed to establish whether a psychosocial stressor could effectively replicate the results of previous studies, specifically those which have found a negative relationship between stress and eyewitness memory. The answer to this is unambiguous: The psychosocial stressor failed to evince the memory degradation produced previously through violent imagery (Clifford & Hollin, 1981), electric shocks (Brigham et al., 1983), inoculations (Peters, 1988), and coercive interrogations (Morgan et al., 2004). As to why this was the case, two alternatives warrant consideration: Firstly, it is possible that the stress manipulation did not produce a defensive state of activation-mode dominance; secondly, it is possible that a Type II error was committed, given the small sample size and subsequent lack of power. The first of these speculations is improbable: The heart rate and subjective anxiety increases bear testament to the presence of a defensive state, rather than a mere orienting response. Thus, the null results can most plausibly be attributed to the study's minimal power – this will be addressed as one of the chief limitations of this research.

#### Limitations and Directions for Future Research

The first possible shortcoming of this study was the limited number of physiological indices of arousal. Heart rate and skin conductance proved useful in signaling the presence of stress, but the lack of parallel variation between these two response channels made the results somewhat difficult to interpret. Including additional measures of sympathetic activation would presumably diminish such ambiguity in the physiological data. In particular, measures of cortisol can be exceptionally informative in the study of stress (especially in relation to memory function). Measures of salivary cortisol were originally included in the present study, but had to be abandoned for various reasons. Within the adolescent study, cortisol sampling was prevented by the constraints of working in a school context. Cortisol levels fluctuate dramatically during the day, necessitating sampling at the same time across conditions. Conducting the stress and non-

---

stress sessions at the same time of day for each participant was impossible, given the class schedule. The solution would have been to take a baseline cortisol sample for the stress session and a second baseline sample for the non-stress session, but the processing of four cortisol samples per participant was beyond the budget of the study. Within the young adult sample, the assay technique used for saliva collection proved unreliable. This problem is not unique to the present study: Previous students have lost large amounts of cortisol data for this reason (e.g. almost 20% of Bonito-Atwood's (2007) sample yielded insufficient quantities of saliva). As the reliability of assay techniques improves, future researchers will be able to include measures of cortisol to verify the presence of stress and to improve the validity of their findings.

The second limitation of this study concerns the administration of the State-Trait Anxiety Inventory. In both experiments, the STAI was administered at the stress session only. While I was able to make inferences about participants' subjective stress based on state-anxiety elevation relative to the norm, it would have been more useful to be able to compare state-anxiety across conditions. Future studies using similar research designs should administer the STAI at both the stress and non-stress sessions. This will permit the inclusion of state-anxiety in the regression analysis, revealing the contribution of subjective stress to variations in eyewitness memory performance.

The third potential weakness of this study was the limited sensitivity of two of the dependent variable measures. For target descriptions, the mean number of descriptors provided was 4.4 in the adolescent sample and 5.6 in the young adult sample. Compared with the number of descriptors provided by actual witnesses to crimes (documented as a median of eight in Van Koppen & Lochun, 1997), participants' free recall descriptions were somewhat deficient in detail. Future studies could probe participants' memories of the perpetrator's appearance more thoroughly using a feature checklist, for instance. This method would prompt witnesses to characterise features that they might otherwise have overlooked (Meissner et al., 2007), and it would eliminate subjective appraisals, personality inferences and vague statements, making scoring easier and more reliable. Alternatively, participants could be provided with appearance 'anchors', such as colour charts from which they could select the shade of the target's hair or

---

eyes (Meissner et al., 2007). This would assist witnesses (particularly juveniles) who lack the vocabulary to make fine-grained descriptive distinctions. For interrogative recall, each category of detail contained only three questions, restricting the potential for detecting memory differences among the three categories. Future studies should include more questions for each type of detail. This could be facilitated by adding more props to the data collection procedure (to serve as peripheral details), and asking participants more forensically relevant questions about the event, such as estimated duration (to serve as central details).

The fourth limitation of the current study is the smallness of the sample. The number of participants for each experiment was determined by logistical constraints. In the adolescent study, such factors as teacher cooperation, timetable clashes and limited class time reduced the number of participants. In the young adult study, the availability of the laboratory, equipment and confederates restricted the number of 'stress' sessions that could be held over a two week period, and thus the number of subjects that could participate. Repeated measures designs are typically advantageous in that they can achieve the same degree of power as independent-sample designs with fewer subjects (Howell, 2004). Despite this, the power levels associated with the findings from both experiments (Power  $(1 - \beta) < 0.7$ ) did not meet the value regarded as desirable for social science research (0.8) (Cohen, 1988). The implication here is that the likelihood of correctly rejecting the null hypothesis was reduced. There is thus a reasonable chance that a relationship between stress and eyewitness memory (which truly exists in the population) went undetected in these experiments, and further research is warranted. In order to produce results from which solid conclusions can be drawn, future researchers should recruit larger samples than those used in the present study.

## Conclusion

The unique contribution of this research lies in its application of a novel and ethical stress manipulation to the study of eyewitness memory. The psychosocial stressor was largely effective

---

in producing evidence of arousal with both the adolescent and young adult samples. There is, however, vast potential for further investigation. In the research presented here, the 'stress' construct was reduced to autonomic activation and general cognitive anxiety. There are almost certainly other elements of arousal experienced by eyewitnesses that were untapped in the present study. Measures of autonomic activation are useful in that they are objective indices of stress, but stress is ultimately a subjective experience. This was taken into account somewhat by the use of a self-report measure of anxiety, but what of emotional arousal, for instance? Surely eyewitnesses experience the emotions of shock and fear, among others. Naturally, this is affected by the type of crime, the degree of danger and the extent of the witness's involvement, but perhaps future researchers should turn their attention first towards understanding the nature of 'stress' experienced by eyewitnesses before pursuing its effect on memory.

Perhaps qualitative data could be used to advance future experimental investigations. Witnesses to actual crimes could be asked to characterise their experiences. What was the predominant emotion involved? Were they overcome with panic, or was there an unexpected sense of calm that descended (as often happens to people in crises), allowing them to behave in the way most appropriate to the situation? We also want to know about the cognitions specific to the event: Were they thinking of escape or engrossed in the scene unfolding before them? Did it occur to them that they might have to describe or recognise the perpetrator later? The first task facing future researchers in this field is to reconceptualise the construct of 'stress' as it applies to eyewitnesses. The methodology used in the present study may then be constructive. The psychosocial manipulation could be used—with innovative modifications, if need be—to produce a generalized state of stress equivalent (in nature if not in magnitude) to that reported by actual eyewitnesses. In addition to measures of physiological activation, validity checks on subjective aspects of stress should be added: For example, post-event interviews could be used to probe participants' levels of cognitive and emotional arousal.

All stress researchers walk the tightrope between ethical defensibility and forensic validity. The research presented here offers a method of investigation that satisfies the first criterion, but the question of forensic validity presents an eternal conundrum. By getting to the heart of what eyewitnesses actually experience and developing the present methodology to provoke an

---

analogous response, we may be one step closer to solving it. Whilst not making a major contribution to the state of knowledge concerning stress and eyewitness memory, this study has possibly laid the groundwork for future research in this regard.

University of Cape Town

---

## REFERENCES

- Adolescent (n.). In *WorldNet-Online*. Retrieved February 19, 2009, from <http://wordnet-online.com/adolescent.shtml>
- Beal, C. R., Schmitt, K. L., & Dekle, D. J. (1995). Eyewitness identification of children: Effects of absolute judgments, nonverbal response options, and event encoding. *Law and Human Behavior, 19*, 197 – 216.
- Bonito Attwood, C. (2008). *The Impact of Acute Psychological Stress on Spatial Cognition*. Unpublished M.Soc.Sci. thesis, University of Cape Town.
- Brigham, J. C., Maass, A., Martinez, D., & Whittenberger, G. (1983). The effect of arousal on facial recognition. *Basic and Applied Social Psychology, 4*, 279 – 293.
- Burke, A., Heuer, F., & Reisberg, D. (1992). Remembering emotional events. *Memory and Cognition, 20*, 277 – 290.
- Cacioppo, J. T., Tassinary, L. G., & Berntson, G. G. (2000). *Handbook of Psychophysiology (2<sup>nd</sup> ed.)*. Cambridge: Cambridge University Press.
- Carey, S., Diamond, R., & Woods, B. (1980). Development of face recognition— a maturational component? *Developmental Psychology, 16*, 257 – 269.
- Cassel, W. S. & Bjorklund, D. F. (1995). Developmental patterns of eyewitness memory and suggestibility: An ecologically based short-term longitudinal study. *Law and Human Behavior, 19*, 507 – 532.
- Ceci, S. J. & Bruck, M. (1993). Suggestibility of the child witness: A historical review and synthesis. *Psychological Bulletin, 113*, 403 – 439.
-

- Chance, J. E. & Goldstein, A. G. (1984). Face-recognition memory: Implications for children's eyewitness testimony. *Journal of Social Issues, 40*, 69 – 85.
- Christiansen, S-A. (1992). Emotional stress and eyewitness memory: A critical review. *Psychological Bulletin, 112*, 284 – 309.
- Christianson, S-A. & Loftus, E.F. (1991). Remembering emotional events: The fate of detailed information. *Cognition and Emotion, 5*, 81 – 108.
- Christianson, S-A., Loftus, E. F., Hoffman, H., & Loftus, G. R. (1991). Eye fixations and memory for emotional events. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*, 693 – 701.
- Clifford, B. R. & Hollin, C.R. (1981). Effects of the type of incident and the number of perpetrators on eyewitness memory. *Journal of Applied Psychology, 64*, 364 – 370.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2<sup>nd</sup> ed.)*. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Cohen, R. L. & Harnick, M. A. (1980). The susceptibility of child witnesses to suggestion: An empirical study. *Law and Human Behavior, 4*, 201 – 210.
- Cutler, B. L., Penrod, S. D., & Martens, T. K. (1987). The reliability of eyewitness identification: The role of system and estimator variables. *Law and Human Behavior, 11*, 233 – 258.
- Deffenbacher, K. A. (1994). Effects of arousal on everyday memory. *Human Performance, 7*, 141 – 161.
-

- Deffenbacher, K. A., Bornstein, B. H., Penrod, S. D., & McGorty, E. K. (2004). A meta-analytic review of the effects of high stress on eyewitness memory. *Law and Human Behavior, 28*, 687 – 706.
- Dekle, D. J., Beal, C. R., Elliott, R., & Huneycutt, D. (1996). Children as witnesses: A comparison of lineup versus showup identification methods. *Applied Cognitive Psychology, 10*, 1 – 12.
- Dysart, J. E. & Lindsay, R. C. L. (2001). A preidentification questioning effect: Serendipitously increasing correct rejections. *Law and Human Behavior, 25*, 155 – 165.
- Easterbrook, J. A. (1959). The effect of emotion on cue utilization and the organization of behavior. *Psychological Review, 66*, 184 – 201.
- Ellis, H., Shepherd, J. & Bruce, A. (1973). The effects of age and sex upon adolescents' recognition of faces. *Journal of Genetic Psychology, 123*, 173 – 174.
- Flin, R. H. (1980). Age effects in children's memory for unfamiliar faces. *Developmental Psychology, 16*, 373 – 374.
- Goodman, G. S. & Reed, R. S. (1986). Age differences in eyewitness testimony. *Law and Human Behavior, 10*, 317 – 332.
- Grisso, T. (1997). The competence of adolescents as trial defendants. *Psychology, Public Policy and Law, 3*, 3 – 32.
- Groot, P. F. C., de Geus, E. J. C., & de Vries, J. (1998). *User manual for the Ambulatory Monitoring System*. Amsterdam: Vrije Universiteit.
- Gross, J. & Hayne, H. (1996). Eyewitness identification by 5- to 6-year-old children. *Law and Human Behavior, 20*, 359 – 373.
-

- Haist, F., Shimamura, A. P., Squire, L. D. (1992). On the relationship between recall and recognition memory. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 18, 691 – 702.
- Hardy, L. & Parfitt, G. (1991). A catastrophe model of anxiety and performance. *British Journal of Psychology*, 82, 163 – 178.
- Howell, D. C. (2004). *Fundamental statistics for the behavioural sciences (5<sup>th</sup> ed.)*. Belmont, CA: Brooks/Cole.
- Heuer, F. & Reisberg, D. (1990). Vivid memories of emotional events: The accuracy of remembered minutiae. *Memory and Cognition*, 18, 496 – 506.
- Huff, C. R. (2004). Wrongful convictions: The American experience. *Canadian Journal of Criminology and Criminal Justice*, 46, 107 – 120.
- Innocence Project (2008). *Understand the Causes: Eyewitness Misidentification*. Retrieved February 16, 2009, from the Innocence Project website:  
<http://www.innocenceproject.org/understand/Eyewitness-Misidentification.php>
- Jennings, J. R., Stringfellow, J. C. & Graham, M. (1974). A comparison of the statistical distributions of beat-by-beat heart rate and heart period. *Psychophysiology*, 11, 207 – 210.
- Kim, J. J. & Yoon, K. S. (1998). Stress: metaplastic effects in the hippocampus. *Trends in Neuroscience*, 21, 505–509.
- Kirschbaum, C., Kudielka, B. M., Gaab, J., Schommer, N. C., & Hellhammer, D. (1999). Impact of gender, menstrual cycle phase and oral contraceptives on the activity of the hypothalamic-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. and Hellhammer, D.H. (1996). Stress- and treatment-induced elevations of cortisol levels associated with impaired declarative memory in healthy adults. *Life Sciences*, 58, 1475 – 1483.
- Kuehn, L. L. (1974). Looking down a gun barrel: person perception and violent crime. *Perceptual and Motor Skills*, 39, 1159 – 1164.
- Lacey, J. I., Bateman, D. E., & Van Lehn, R. (1953). Autonomic response specificity: An experimental study. *Psychosomatic Medicine*, 15, 8 – 21.
- Lacey, J. I. & Lacey, B. C. (1958). Verification and extension of the principle of autonomic response-stereotypy. *American Journal of Psychology*, 71, 50 – 73.
- Lazarus, R. S., Speisman, J. C., & Mordkoff, A. M. (1963). The relationship between autonomic indicators of psychological stress: Heart rate and skin conductance. *Psychosomatic Medicine*, 25, 19 – 30.
- Libkuman, T. M., Nichols-Whitehead, P., Griffith, J., & Thomas, R. (1999). Source of arousal and memory for detail. *Memory & Cognition*, 27, 166 – 190.
- Lindberg, M. A., Jones, S., Collard, L. M., & Thomas, S. W. (2001). Similarities and differences in eyewitness testimonies of children who directly versus vicariously experience stress. *Journal of Genetic Psychology*, 162, 314 – 333.
-

- Lindsay, R. C. L., Pozzulo, J. D., Craig, W., Lee, K., & Corber, S. (1997). Simultaneous lineups, sequential lineups, and showups: Eyewitness identification decisions of adults and children. *Law and Human Behavior, 4*, 391 – 404.
- Loftus, E. F. (1979). *Eyewitness Testimony*. Cambridge, MA: Harvard University Press.
- Loftus, E. E., & Burns, T. E. (1982). Mental shock can produce retrograde amnesia. *Memory and Cognition, 10*, 318 – 323.
- Lovullo, W. R. & Thomas, T. L. (2000). Stress hormones in psychobiological research: Emotional, behavioral and cognitive implications. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of Psychophysiology (2<sup>nd</sup> ed.)* (pp. 342 – 367). Cambridge: Cambridge University Press.
- Lupien, S. J. & McEwen, B. S. (1997). The acute effects of corticosteroids on cognition: Integration of animal and human model studies. *Brain Research Reviews, 24*, 1 – 27.
- MacLin, O. H., MacLin, M. K., & Malpass, R. S. (2001). Race, arousal, attention, exposure, and delay: An examination of factors moderating face recognition. *Psychology, Public Policy, and Law, 7*, 134 – 152.
- Malpass, R. S. (1981). Effective size and defendant bias in eyewitness identification lineups. *Law and Human Behaviour, 5*, 299 – 309.
- Malpass, R. S., Tredoux, C., & McQuiston-Surrett, D. (2007). Lineup construction and lineup fairness. In R. Lindsay, D. F. Ross, J. D. Read, and M. P. Toglia (Eds.), *Handbook of eyewitness psychology, Vol. 2: Memory for people* (pp. 155 – 178). Mahwah, NJ: Lawrence Erlbaum Associates.
-

- Marin, B. V., Holmes, D. L., Guth, M., & Kovac, P. (1979). The potential of children as eyewitnesses: A comparison of children and adults on eyewitness tasks. *Law and Human Behavior, 3*, 295 – 306.
- McEwen, B.S. & Magarinos, A.M. (1997). Stress effects on morphology and function of the hippocampus. *Annals of the New York Academy of Sciences, 821*, 271 – 284.
- McMurtrie, J. (2005). The role of the social sciences in preventing wrongful convictions. *American Criminal Law Review, 42*, 1271 – 1288.
- McQuiston-Surrett, D., Malpass, R. S., & Tredoux, C. G. (2006). Sequential vs. simultaneous lineups: A review of methods, data, and theory. *Psychology, Public Policy, and Law, 12*, 137 – 169.
- Meissner, C.A. & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law, 7*, 3 – 35.
- Meissner, C. A., Sporer, S. L., & Schooler, J. W. (2007). Person descriptions as eyewitness evidence. In R. Lindsay, D. F. Ross, J. D. Read, and M. P. Toglia (Eds.), *Handbook of eyewitness psychology, Vol. 2: Memory for people* (pp. 3 – 34). Mahwah, NJ: Lawrence Erlbaum Associates.
- Merritt, K. A., Ornstein, P. A., & Spicker, B. (1994). Children's memory for a salient medical procedure: Implications for testimony. *Pediatrics, 94*, 17 – 23.
- Miley, W. M. (1999). *The psychology of well-being*. London: Praeger.
- Morgan, C. A., Hazlett, G., Doran, A., Garrett, S., Hoyt, G., Thomas, P., Baranoski, M., & Southwick, S. M. (2004). Accuracy of eyewitness memory for persons encountered
-

during exposure to highly intense stress. *International Journal of Law and Psychiatry*, 27, 265 – 279.

Oitzl, M. S. & de Kloet, R. (1992). Selective corticosteroid antagonists modulate specific aspects of spatial orientation learning. *Behavioral Neuroscience*, 106, 62 – 71.

Parker, J. F. & Ryan, V. (1993). An attempt to reduce guessing behavior in children's and adults' eyewitness identifications. *Law and Human Behavior*, 17, 11 – 26.

Peters, D. P. (1987). The impact of naturally occurring stress on children's memory. In S. J. Ceci, M. P. Toglia and D. F. Ross (Eds.), *Children's eyewitness memory* (pp. 122 – 141). New York: Springer-Verlag.

Peters, D.P. (1988). Eyewitness memory and arousal in a natural setting. In M.M. Gruneberg, P.E. Morris and R.N. Sykes (Eds.), *Practical aspects of memory: Current research and issues, Volume 1* (pp. 89 – 94). Chichester: John Wiley & Sons.

Pozzulo, J. D. & Lindsay, R. C. L. (1998). Identification accuracy of children versus adults: A meta-analysis. *Law and Human Behavior*, 22, 549 – 570.

Pozzulo, J. D. & Lindsay, R. C. L. (1999). Elimination lineups: An improved identification procedure for child eyewitnesses. *Journal of Applied Psychology*, 84, 167 – 176.

Pozzulo, J. D. & Warren, K. L. (2003). Descriptions and identifications of strangers by youth and adult eyewitnesses. *Journal of Applied Psychology*, 88, 315 – 323.

*R. v. Mputing* 1960 (1) (SA) 785 (T).

*R. v. Shekelele and another* 1953 (1) (SA) 636 (T).

*R. v. T.* 1958 (2) (SA) 676 (A).

---

Robinson, J. A. (1997). *The law of children and young persons in South Africa*. Durban: Butterworths.

*S. v. Mehlape* 1963 (2) (SA) 29 (A).

*S. v. Molelekwa and others* 2007 (JOL) 19040 (T).

Safer, M. A., Christianson, S-A., Autry, M. W., & Osterlund, K. (1998). Tunnel memory for traumatic events. *Applied Cognitive Psychology, 12*, 99 – 117.

Scott, E. S., Repucci, N. D., & Woolard, J. L. (1995). Evaluating adolescent decision making in legal contexts. *Law and Human Behavior, 19*, 221 – 244.

Shapiro, S. S., Wilk, M. B., & Chen, H. J. (1968). A comparative study of various tests for normality. *Journal of the American Statistical Association, 63*, 1343 – 1372.

Shrimpton, S., Oates, K., & Hayes, S. (1998). Children's memory of events: Effects of stress, age, time delay and location of interview. *Applied Cognitive Science, 12*, 133 – 143.

Spielberger, C. D., Gorsuch, R.L., Lushene, R.E., Vagg, P.R., & Jacobs, G.A. (1983). *State-trait anxiety inventory for adults: Manual, test, scoring key*. Redwood City, CA: Mind Garden, Inc.

Sporer, S. L. (1992, March). *An archival analysis of person descriptions*. Paper presented at the Biennial Meeting of the American Psychology-Law Society in San Diego, California.

Sporer, S. L. (1996). Psychological aspects of person descriptions. In S. L. Sporer, R. S. Malpass, & G. Koehnken (Eds.), *Psychological issues in eyewitness identification* (pp. 53 – 86). Mahwah, NJ: Lawrence Erlbaum Associates.

---

- Squire, L. R., Ojemann, J. G., Miezin, F. M., Petersen, S. E., Videen, T. O., & Raichle, M. E. (1992). Activation of the hippocampus in normal humans: A functional anatomical study of memory. *Proceedings of the National Academy of Sciences of the United States of America*, *89*, 1837 – 1841.
- StatSoft, Inc. (2007). STATISTICA Version 8. Tulsa, OK: StatSoft.
- Stebly, N. M. (1992). A meta-analytic review of the weapon focus effect. *Law and Human Behavior*, *16*, 413 – 424.
- Stebly, N., Dysart, J., Fulero, S., & Lindsay, R. C. L. (2001). Eyewitness accuracy rates in sequential and simultaneous lineup presentations: A meta-analytic comparison. *Law and Human Behavior*, *25*, 459 – 473.
- Tredoux, C. G. (1996). *Evaluating the fairness of identification parades with measures of facial similarity*. Unpublished Ph.D. thesis, University of Cape Town.
- Tredoux, C. & Chiroro, P. (2005). Eyewitness testimony. In C. Tredoux, D. Foster, A. Allan, A. Cohen & D. Wassenaar (Eds.), *Psychology and Law* (pp. 193 – 225). Lansdowne, South Africa: Juta.
- Tucker D.M. & Williamson, P.A. (1984). Asymmetric neural control systems in human self-regulation. *Psychological Review*, *91*, 185 – 215.
- U.S. Department of Justice, Office of Justice Programs, Office of Juvenile Justice and Delinquency Prevention (2006). *Juvenile offenders and victims: 2006 National Report*. Retrieved February 13<sup>th</sup>, 2009, from <http://www.ojjdp.ncjrs.gov/ojstatbb/nr2006/downloads/NR2006.pdf>
- Van Koppen, P. J. & Lochun, S. L. (1997). Portraying perpetrators: The validity of offender descriptions by witnesses. *Law and Human Behavior*, *21*, 661 – 685.
-

- Wagstaff, G. F., MacVeigh, J., Boston, R., Scott, L., Brunas-Wagstaff, J., & Cole, J. (2003). Can laboratory findings on eyewitness testimony be generalized to the real world? An archival analysis of the influence of violence, weapon presence, and age on eyewitness accuracy. *Journal of Psychology, 137*, 17 – 28.
- Wells, G. L. & Leippe, M. R. (1981). How do triers of fact infer the accuracy of eyewitness identifications? Using memory for peripheral detail can be misleading. *Journal of Applied Psychology, 66*, 682 – 687.
- Wells, G. L. & Olson, E. A. (2003). Eyewitness Testimony. *Annual Review of Psychology, 54*, 277 – 295.
- Wells, G. L., Seelau, E. P., Rydell, S. M., & Luus, C. A. E. (1994). Recommendations for properly conducted lineup identification tasks. In D. F. Ross, J. D. Read, and M. P. Toglia (Eds.), *Adult eyewitness testimony: Current trends and developments* (pp. 223 – 244). New York: Cambridge University Press.
- Wiggs, C. L., Weisberg, J., & Martin, A. (1999). Neural correlates of semantic and episodic memory retrieval. *Neuropsychologia, 37*, 103 – 118.
- Wolkowitz, O.M., Reus, V.I. and Weingartner, H. (1990). Cognitive effects of corticosteroids. *American Journal of Psychiatry, 147*, 1297 – 1303.
- Yuille, J. C. & Cutshall, J. L. (1989). A case study of eyewitness memory of a crime. *Journal of Applied Psychology, 71*, 291 – 301.
- Yuille, J. C., Davies, G., Gibling, F., Marxsen, D., & Porter, S. (1994). Eyewitness memory of police trainees for realistic role plays. *Journal of Applied Psychology, 6*, 931 – 936.
-

Appendix A

The Spielberger State-Trait Anxiety Inventory: State Scale

mind garden

**SELF-EVALUATION QUESTIONNAIRE**  
STAI Form Y-1

Please provide the following information:

Name \_\_\_\_\_ Date \_\_\_\_\_ S \_\_\_\_\_

Age \_\_\_\_\_ Gender (Circle) M F T \_\_\_\_\_

**DIRECTIONS:**

A number of statements which people have used to describe themselves are given below. Read each statement and then circle the appropriate number to the right of the statement to indicate how you feel *right* now, that is, *at this moment*. There are no right or wrong answers. Do not spend too much time on any one statement but give the answer which seems to describe your present feelings best.

- |   |   |   |   |   |
|---|---|---|---|---|
| 1. I feel calm.....                                       | 1 | 2 | 3 | 4 |
| 2. I feel secure .....                                    | 1 | 2 | 3 | 4 |
| 3. I am tense .....                                       | 1 | 2 | 3 | 4 |
| 4. I feel strained.....                                   | 1 | 2 | 3 | 4 |
| 5. I feel at ease.....                                    | 1 | 2 | 3 | 4 |
| 6. I feel upset.....                                      | 1 | 2 | 3 | 4 |
| 7. I am presently worrying over possible misfortunes..... | 1 | 2 | 3 | 4 |
| 8. I feel satisfied .....                                 | 1 | 2 | 3 | 4 |
| 9. I feel frightened.....                                 | 1 | 2 | 3 | 4 |
| 10. I feel comfortable.....                               | 1 | 2 | 3 | 4 |
| 11. I feel self-confident .....                           | 1 | 2 | 3 | 4 |
| 12. I feel nervous .....                                  | 1 | 2 | 3 | 4 |
| 13. I am jittery.....                                     | 1 | 2 | 3 | 4 |
| 14. I feel indecisive.....                                | 1 | 2 | 3 | 4 |
| 15. I am relaxed.....                                     | 1 | 2 | 3 | 4 |
| 16. I feel content .....                                  | 1 | 2 | 3 | 4 |
| 17. I am worried .....                                    | 1 | 2 | 3 | 4 |
| 18. I feel confused .....                                 | 1 | 2 | 3 | 4 |
| 19. I feel steady .....                                   | 1 | 2 | 3 | 4 |
| 20. I feel pleasant .....                                 | 1 | 2 | 3 | 4 |

VERY MUCH SO  
MODERATELY SO  
SOMEWHAT  
NOT AT ALL

## The Spielberger State-Trait Anxiety Inventory: Trait Scale

mind garden

SELF-EVALUATION QUESTIONNAIRE  
STAI Form Y-2

Name \_\_\_\_\_ Date \_\_\_\_\_

## DIRECTIONS

A number of statements which people have used to describe themselves are given below. Read each statement and then circle the appropriate number to the right of the statement to indicate how you *generally* feel.

- |  |   |              |           |       |               |
|--|---|--------------|-----------|-------|---------------|
|  |   | ALMOST NEVER | SOMETIMES | OFTEN | ALMOST ALWAYS |
| 21. I feel pleasant .....  | 1 | 2            | 3         | 4     |               |
| 22. I feel nervous and restless.....   | 1 | 2            | 3         | 4     |               |
| 23. I feel satisfied with myself.....  | 1 | 2            | 3         | 4     |               |
| 24. I wish I could be as happy as others seem to be .....  | 1 | 2            | 3         | 4     |               |
| 25. I feel like a failure.....   | 1 | 2            | 3         | 4     |               |
| 26. I feel rested.....   | 1 | 2            | 3         | 4     |               |
| 27. I am "calm, cool, and collected" .....   | 1 | 2            | 3         | 4     |               |
| 28. I feel that difficulties are piling up so that I cannot overcome them.....                   | 1 | 2            | 3         | 4     |               |
| 29. I worry too much over something that really doesn't matter.....                              | 1 | 2            | 3         | 4     |               |
| 30. I am happy .....   | 1 | 2            | 3         | 4     |               |
| 31. I have disturbing thoughts .....   | 1 | 2            | 3         | 4     |               |
| 32. I lack self-confidence .....   | 1 | 2            | 3         | 4     |               |
| 33. I feel secure .....  | 1 | 2            | 3         | 4     |               |
| 34. I make decisions easily.....   | 1 | 2            | 3         | 4     |               |
| 35. I feel inadequate.....   | 1 | 2            | 3         | 4     |               |
| 36. I am content.....  | 1 | 2            | 3         | 4     |               |
| 37. Some unimportant thought runs through my mind and bothers me .....                           | 1 | 2            | 3         | 4     |               |
| 38. I take disappointments so keenly that I can't put them out of my mind.....                   | 1 | 2            | 3         | 4     |               |
| 39. I am a steady person .....   | 1 | 2            | 3         | 4     |               |
| 40. I get in a state of tension or turmoil as I think over my recent concerns and interests..... | 1 | 2            | 3         | 4     |               |

Appendix B

Lineup for Research Assistant 1 (Target is in position 4)



4



8



3



7



2



6



1



5

Appendix C

Lineup for Research Assistant 2 (Target is in position 7).

 1	 2	 3	 4	 5	 6	 7	 8
--	--	---	--	---	---	--	---

Appendix D

Lineup Tasks used in Experiments 1 and Experiment 2

<b>NAME:</b> _____	<b>DATE:</b> _____
<b>TASK 1</b>	
<p>Over the past two sessions of this experiment, you interacted with 2 different research assistants who connected you to the AMS.</p> <p>Below is a space for you to write a <b>physical</b> description of each. Please write as much as you remember, as though you were doing a culprit description for the police. You can mention things like height and build, hair length and colour, facial features etc.</p>	
<b>RESEARCH ASSISTANT 1:</b>	
_____ _____ _____ _____ _____ _____ _____ _____	
<b>RESEARCH ASSISTANT 2:</b>	
_____ _____ _____ _____ _____ _____ _____ _____	

**TASK 2**

We are interested in seeing how much you remember about the other sessions of the experiment.

There are 2 sets of questions below. Answer the first set of questions thinking about the first session of the experiment, and the second set of questions thinking about the second session.

**SESSION 1:**

- 1) What was the research assistant's name? \_\_\_\_\_
- 2) You were asked to take a clock out of the bag and put it on the desk. The clock displayed the wrong time -- what time did it say? \_\_\_\_\_
- 3) What logo was on the research assistant's T-shirt? \_\_\_\_\_
- 4) You were offered a book to press on while completing the written task. What was on the cover of the book? \_\_\_\_\_
- 5) What was the colour of the ink you used to complete the task you were given while hooked up to the AMS? \_\_\_\_\_
- 6) You also removed a jar from the bag and placed it on the desk. What was in the jar? \_\_\_\_\_
- 7) What question did the research assistant ask you about yourself while connecting you to the AMS? \_\_\_\_\_
- 8) Was the research assistant wearing her hair loose or tied up? \_\_\_\_\_
- 9) There was a striped scarf hanging on the back of the chair you sat on. What colour were the stripes? \_\_\_\_\_

*Note.* For both Sessions 1 and 2, the category, 'Details of the Research Assistant' comprised of questions 1, 3, and 8. The category, 'Details of the Event' comprised of questions 4, 5, and 7. The category, 'Details of the Environment' comprised of questions 2, 6, and 9.

**SESSION 2:**

- 1) What was the research assistant's name? \_\_\_\_\_
- 2) You were asked to take an out-of-date calendar out of the bag and put it on the desk. What month and year were displayed on the calendar?  
\_\_\_\_\_
- 3) What was written on the research assistant's sweatshirt?  
\_\_\_\_\_
- 4) You were offered a book to press on. What was the cover of the book?  
\_\_\_\_\_
- 5) What was the colour of the ink you used to complete the written task while hooked up to the AMS?  
\_\_\_\_\_
- 6) You also removed a jar from the bag and placed it on the desk. What was in the jar?  
\_\_\_\_\_
- 7) What question did the research assistant ask you about yourself while connecting you to the AMS?  
\_\_\_\_\_
- 8) Was the research assistant wearing her hair loose or tied up?  
\_\_\_\_\_
- 9) There was an umbrella hanging on the back of the chair you sat on. What colour was the umbrella?  
\_\_\_\_\_

**TASK 3**

The next set of questions is **exactly the same** as those you just answered, only in multiple choice format.

If you realise that you got one of the answers wrong in the last task, don't worry – just circle the correct answer this time around.

**SESSION 1:**

- 1) The research assistant's name was:
- A) Joanne
  - B) Kelly
  - C) Nicola
  - D) Janine
  - E) None of the above
- 2) The clock you placed on the desk said:
- A) 3:00
  - B) 5:00
  - C) 10:00
  - D) 12:00
  - E) None of the above
- 3) On the research assistant's T-shirt, there was a:
- A) Nike logo
  - B) Lacoste logo
  - C) Adidas logo
  - D) Puma logo
  - E) None of the above
- 4) The book you pressed on was called *Animation Now!*  
On the cover was:
- A) Buzz Lightyear
  - B) Shrek
  - C) Mr. Incredible
  - D) Nemo
  - E) None of the above

5) The ink you used to complete the task was:

- A) Green
- B) Purple
- C) Black
- D) Blue
- E) None of the above

6) In the jar on the desk, there were:

- A) Paintbrushes
- B) Flowers
- C) Feathers
- D) Beans
- E) None of the above

7) The research assistant asked:

- A) What's your favourite TV show?
- B) What's your favourite subject?
- C) What's your favourite sport?
- D) What's your favourite food?
- E) None of the above

8) The scarf hanging on the chair was:

- A) Blue and white striped
- B) Green and grey striped
- C) Red and white striped
- D) Pink and grey striped
- E) None of the above

**SESSION 2:**

- 1) The research assistant's name was:
- A) Caryn
  - B) Jessica
  - C) Candice
  - D) Sarah
  - E) None of the above
- 2) The calendar on the table said:
- A) November 2004
  - B) June 2007
  - C) January 2001
  - D) May 2002
  - E) None of the above
- 3) The research assistant's sweatshirt said:
- A) Fuller Hall
  - B) W. P. Netball
  - C) University of Cape Town
  - D) Sandringham Matric
  - E) None of the above
- 4) The book you pressed on was a Jamie Oliver cookbook.  
The title was:
- A) 'Jamie's Little Book of Big Treats'
  - B) 'Jamie's Kitchen'
  - C) 'Jamie's Dinners'
  - D) 'Jamie's Italy'
  - E) None of the above
- 5) The ink you used to complete the task was:
- A) Black
  - B) Red
  - C) Purple
  - D) Blue
  - E) None of the above

6) In the jar on the desk, there were:

- A) Marbles
- B) Pencils
- C) Coins
- D) Khoki pens
- E) None of the above

7) The research assistant asked:

- A) How old are you?
- B) How many brothers/sisters do you have?
- C) How far away do you live?
- D) How long have you been at Bishops?
- E) None of the above

8) The umbrella hanging on the chair was:

- A) Red
- B) White
- C) Black
- D) Pink
- E) None of the above

**TASK 4**

You will now be shown 2 lineups on the computer. The first lineup is for the first research assistant you met, and the second lineup is for the second research assistant.

In one of the boxes below, please mark the number of the research assistant's photograph with an X.

**PLEASE NOTE: the research assistant may not actually be in the lineup at all. If you don't think she is there, put an X in the box which says 'None of them'.**

**LINEUP 1**

1	2	3	4	5	6	7	8	NONE OF THEM
---	---	---	---	---	---	---	---	--------------------

**LINEUP 2**

1	2	3	4	5	6	7	8	NONE OF THEM
---	---	---	---	---	---	---	---	--------------------

## Appendix E

Frequency Table of Physical Characteristics Mentioned by Students for Research Assistant 1

Characteristic	Full-view Group	Memory Group
Round face	3	1
Oval face	1	2
Heart-shaped face	1	0
<i>Pointy nose</i>	6	4
Small nose	3	1
<i>Clear complexion</i>	5	3
<i>Fair skin</i>	5	3
Rosy cheeks	3	2
<i>Blue eyes</i>	5	3
Small eyes	3	1
Large eyes	1	0
Wide-set eyes	1	0
<i>Medium height<sup>a</sup></i>	7	5
<i>Medium weight/build<sup>b</sup></i>	6	5
<i>Blonde hair</i>	10	7
Highlighted hair	3	2
<i>Shoulder-length hair</i>	9	7
Straight hair	2	2
<i>Side-swept fringe</i>	4	4
Side part	2	1
<i>Thin eyebrows</i>	4	3
<i>Shaped/defined eyebrows</i>	4	1
Dark eyebrows	1	1
Delicate/small hands	1	2
Straight teeth	0	3
<i>Thin lips</i>	8	3

*Note.* Characteristics that were included in the standardised description are italicized.

<sup>a</sup> The research assistant's actual height was 164cm. Participants were awarded a full point for estimates within 5cm, and half a point for vague descriptions (such as 'average' or 'medium').

<sup>b</sup> The research assistant's actual weight was 65 kg. Participants were awarded a full point for estimates within 3 kg, and half a point for vague descriptions (such as 'average' or 'normal').

Frequency Table of Physical Characteristics Mentioned by Students for Research Assistant 2

Characteristic	Full-view Group	Memory Group
Oval face	0	1
Round face	3	0
<i>Long hair</i>	7	4
Hair below shoulders	3	0
Reddish brown hair	1	0
Light brown hair	1	0
<i>Dark brown hair</i>	7	3
Straight hair	1	1
Curly	0	2
<i>Brown eyes</i>	9	5
Small eyes	1	3
Shadows under eyes	0	1
<i>Thin eyebrows</i>	7	5
<i>Pierced ears</i>	5	1
<i>Dimples</i>	7	6
<i>Thin lips</i>	4	5
<i>Pale skin</i>	6	5
Clear complexion	1	1
Long neck	1	0
Sharp/pointy nose	2	1
Straight nose	0	2
Hooked nose	0	1
<i>Medium height<sup>a</sup></i>	5	5
Medium build	2	2
<i>Slim/slender/small build<sup>b</sup></i>	8	4

*Note.* Characteristics that were included in the standardised description are italicized.

<sup>a</sup> The research assistant's actual height was 167cm. Participants were awarded a full point for estimates within 5cm, and half a point for vague descriptions (such as 'average' or 'medium').

<sup>b</sup> The research assistant's actual weight was 59 kg. Participants were awarded a full point for estimates within 3 kg, and half a point for vague descriptions (such as 'slender' or 'slim').

## Appendix F

Consent Form for Experiment 1

---

**UNIVERSITY OF CAPE TOWN**

---

**Department of Psychology**

University of Cape Town Rondebosch 7701 South Africa

MARIANNE BALL (B.A. Hons.)

☎ 0725217552 (c); 6852418 (h)

💻 [Mariannc.ball@gmail.com](mailto:Mariannc.ball@gmail.com)

05 April 20098

**The Effects of Public Speaking on Memory**

Dear Parents/Legal Guardians,

Your son is invited to participate in a research study about how public speaking affects memory in adolescent boys.

We are researchers from the UCT Psychology Department, and will be conducting research at Bishops during October 2008 (under the supervision of Professor Colin Tredoux). Funding for this study comes from the National Research Foundation.

With the assistance of the Beverly Kemball, a special Life Orientation program has been designed for the purposes of the study. All students (whether they choose to take part in the study or not) will participate in this program, as it has been worked into the Life Orientation

---

curriculum for the fourth quarter. If your son chooses to take part in the study, he will be agreeing to have some physiological measurements taken, to fill out a questionnaire on how he feels about public speaking, and to complete a short computerised memory task.

The physiological measurements that will be taken are heart rate and skin conductance. To record your son's heart rate and skin conductance, he will be connected to a piece of equipment called an Ambulatory Monitoring System (AMS). This involves attaching 3 electrodes to his torso and wrapping a special velcro 'strap' around his finger to measure electrodermal responsiveness (sweating). He will remain attached to the AMS for 5 minutes.

Two sets of measurements will be taken, a week apart. This will allow us to see whether your son's heart rate and skin conductance are different on the day of the speech compared with an ordinary school day. The physiological measurements are completely harmless, and we do not anticipate that your son will experience any discomfort during or after the procedure.

This study will help us learn more about how public speaking affects memory – this has important implications within a school context as well as a broader social context.

Your son is not obliged to participate in this study. If he does agree to be a participant, he may withdraw at any point.

If you and your son decide that he would like to participate in the study, please sign below and return this form. If you have any questions or concerns, please feel free to contact Marianne Ball at 072 521 7552 or email [marianne.ball@gmail.com](mailto:marianne.ball@gmail.com).

---

**I have been informed about this study and agree to my son's participation. I authorize the collection and use of my son's physiological and memory performance data for research purposes.**

---

Learner's Parent/Legal Guardian (signature)  
(PLEASE ALSO PRINT NAME)

---

Date

**I have been informed about this study and agree to be a research participant.**

---

Learner (signature)  
(PLEASE ALSO PRINT NAME)

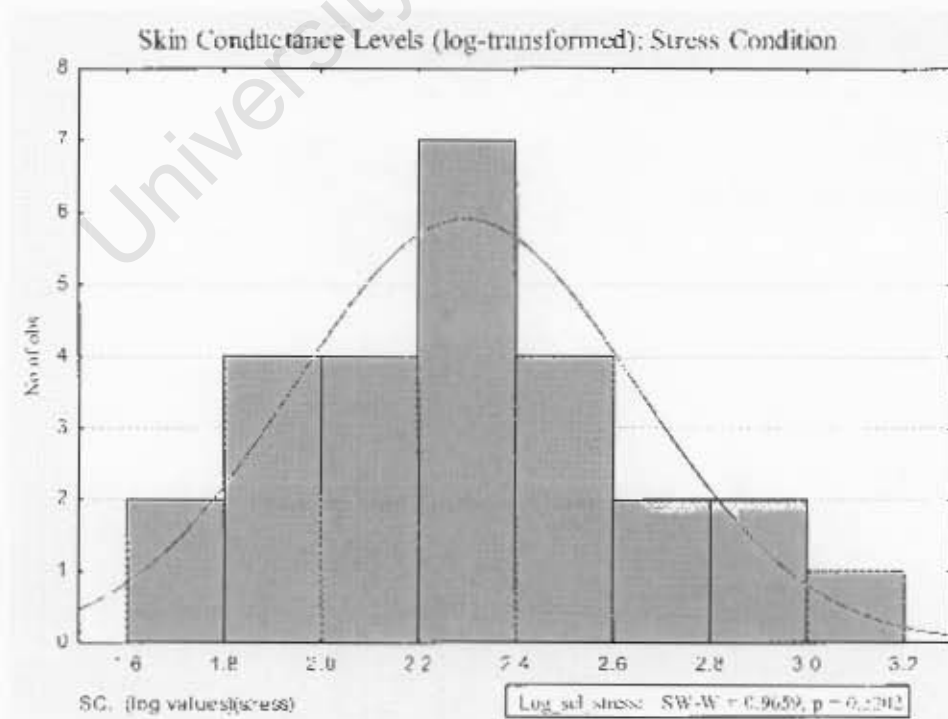
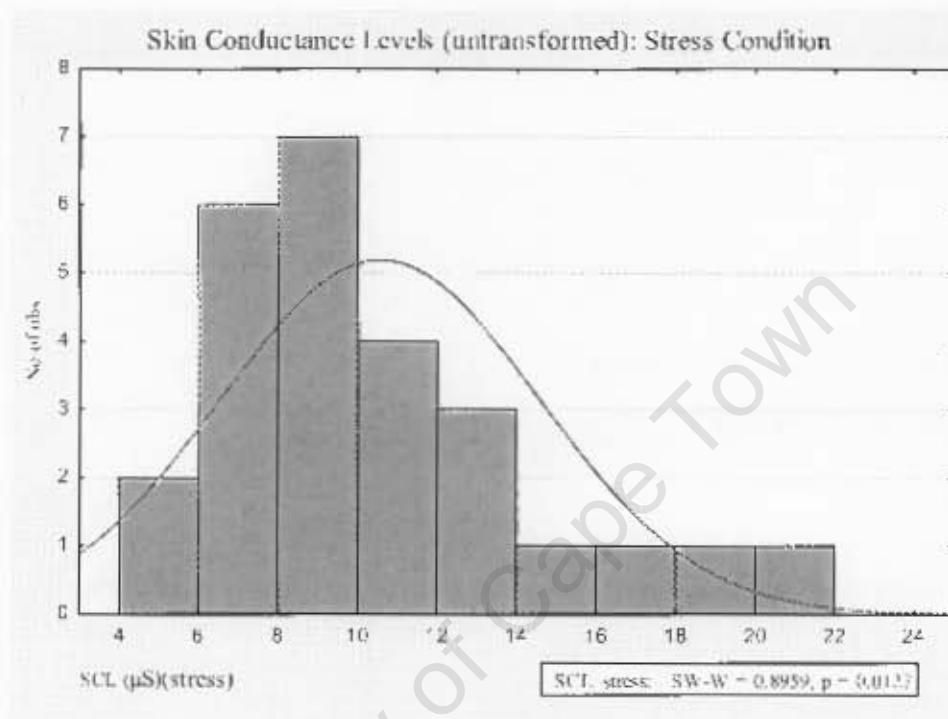
---

Date

---

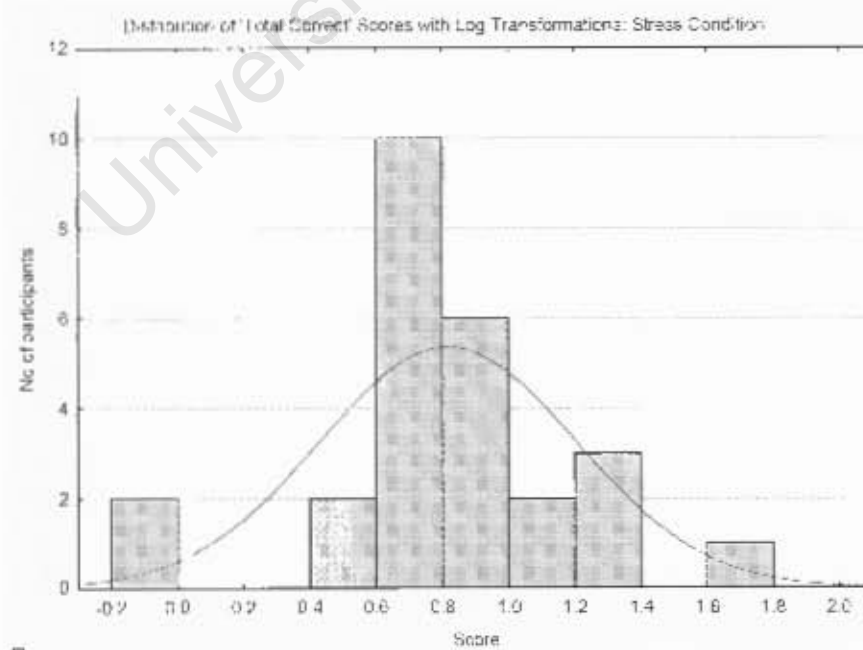
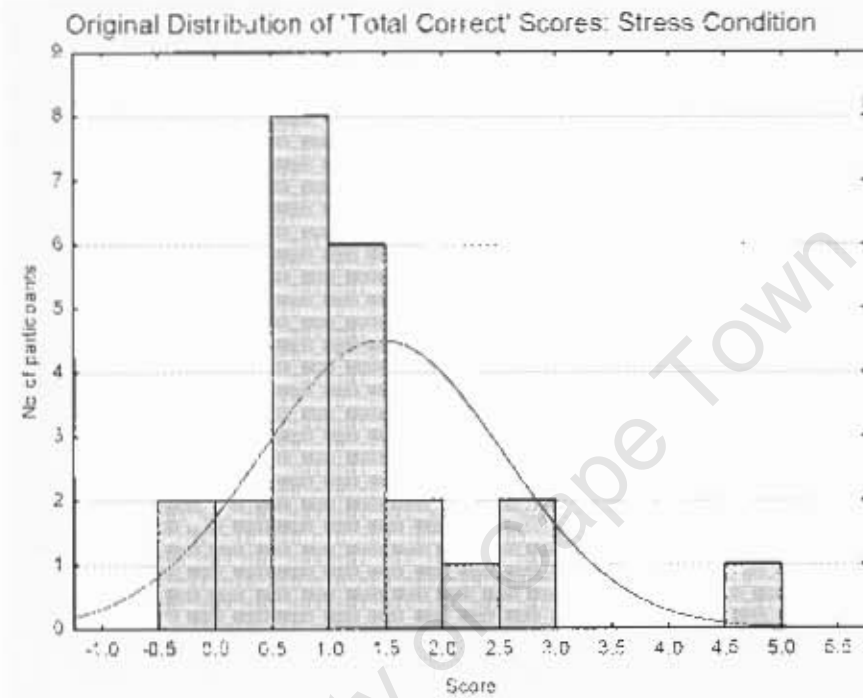
## Appendix G

## Log-transformation of Skin Conductance Data (Experiment 1)



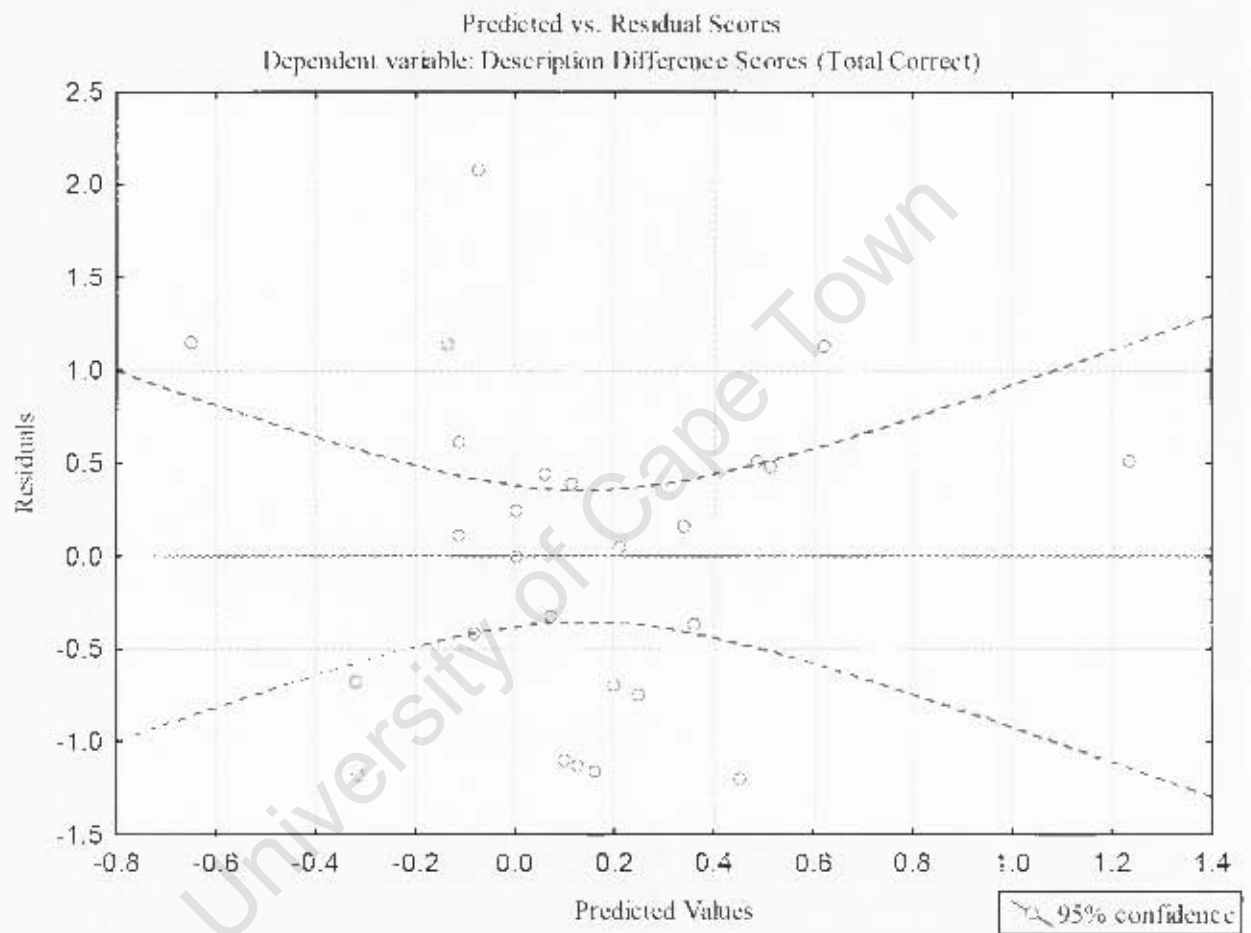
## Appendix H

## Log-transformation of Total Correct Scores (Experiment 1)



## Appendix 1

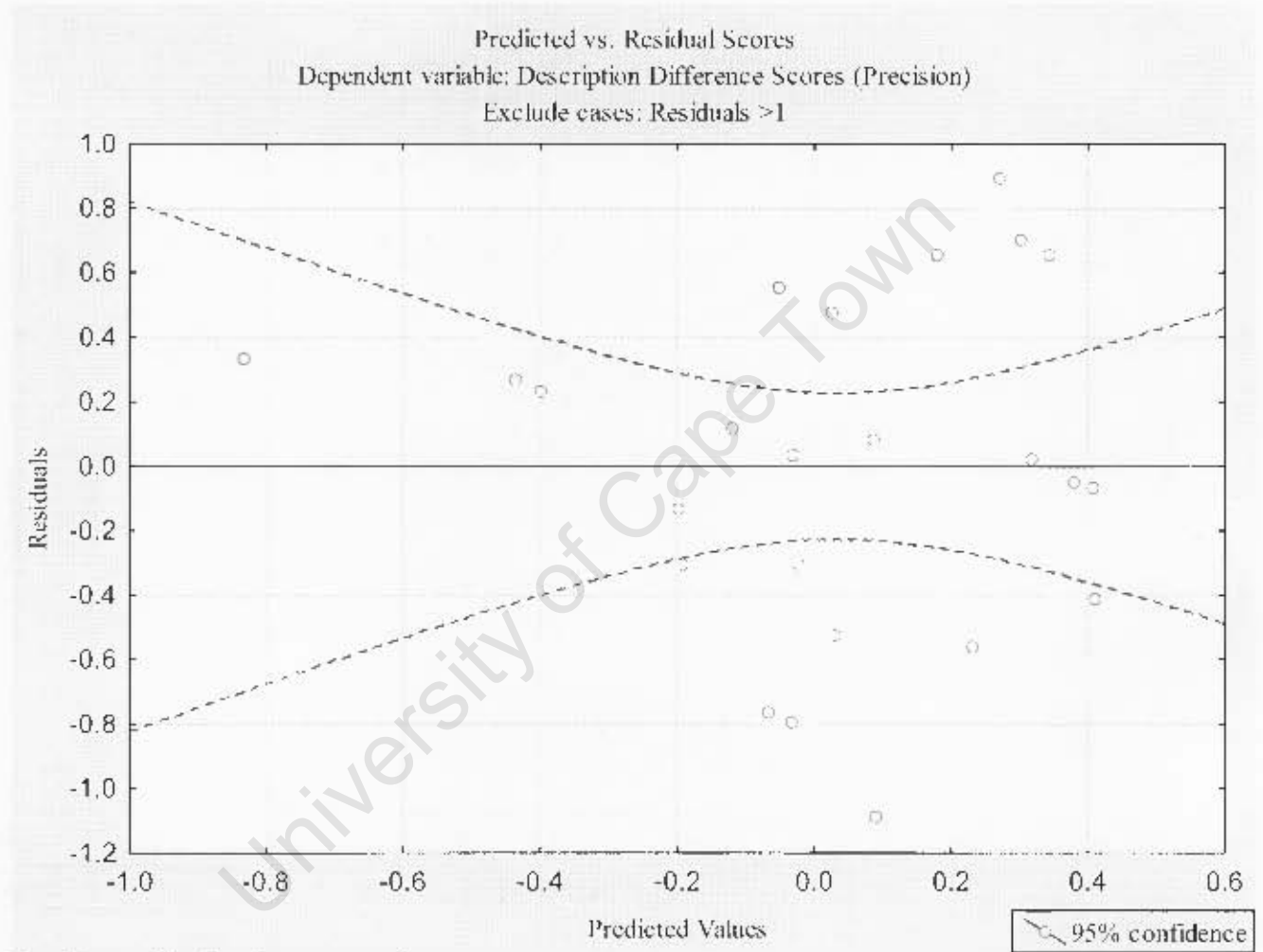
Plot of Residuals for Total Correct Items (Experiment 1)





## Appendix K

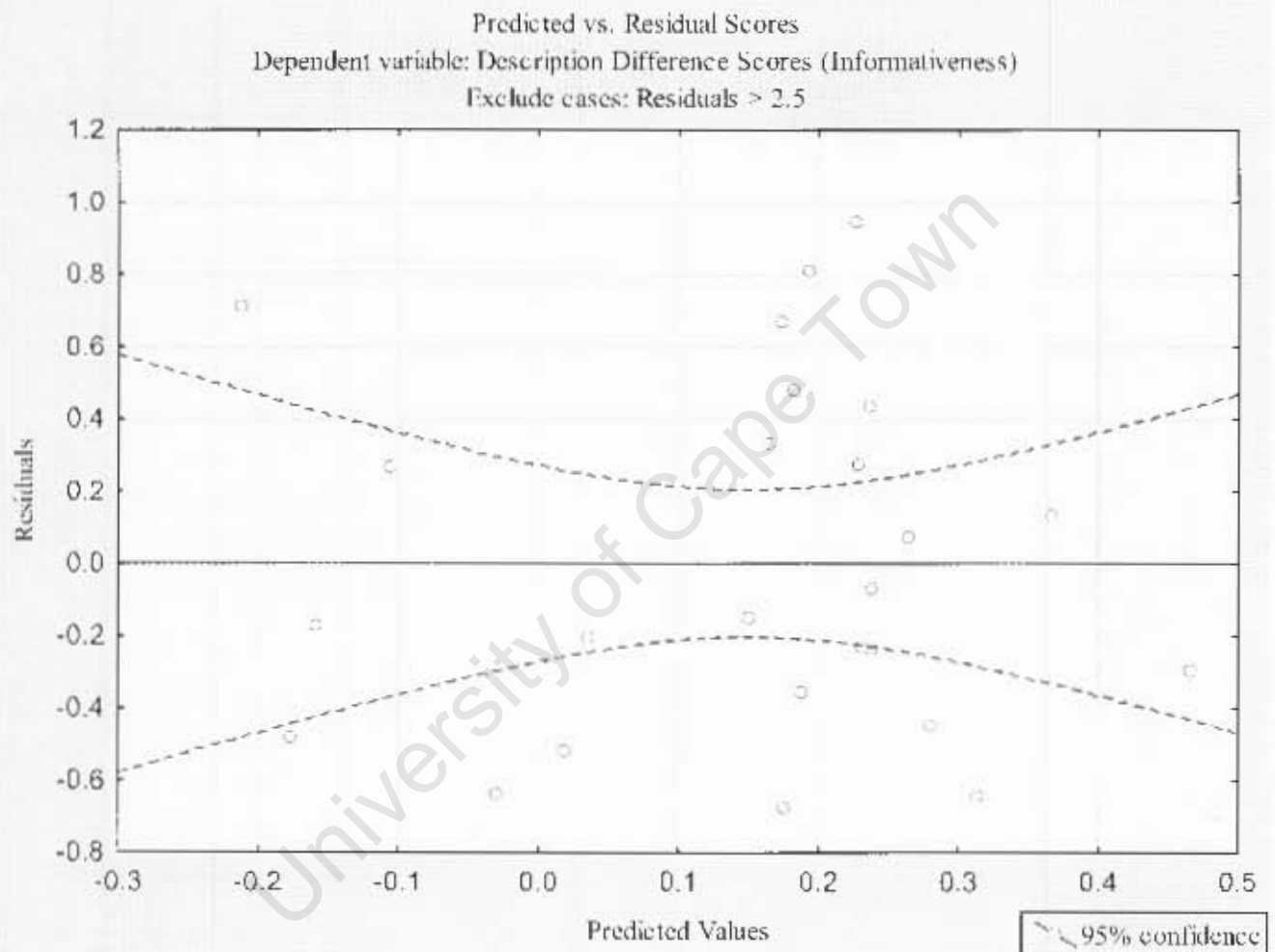
## Plot of Residuals for Precision (Experiment 1)



Note. Two outliers have been removed.

## Appendix I.

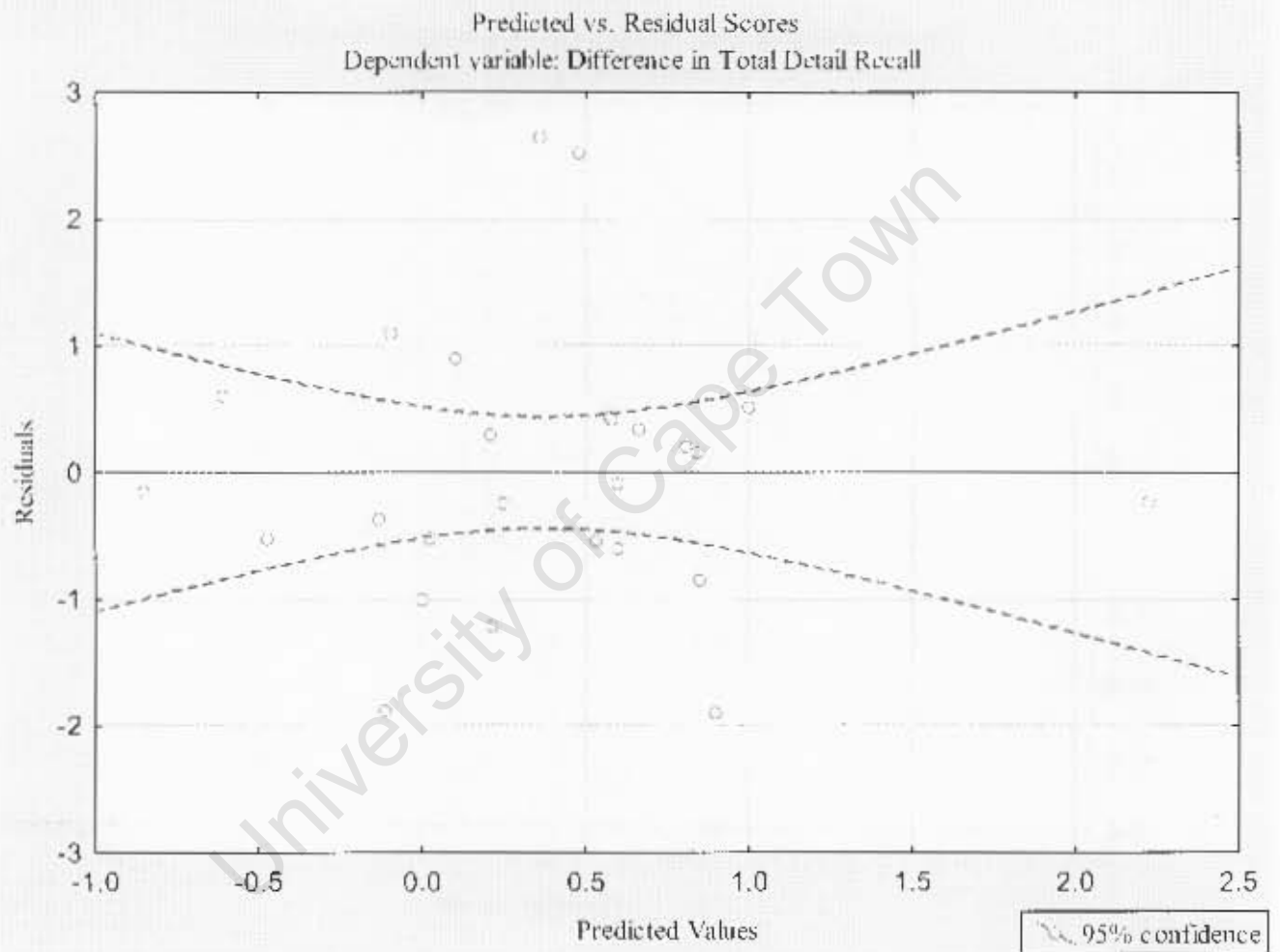
## Plot of Residuals for Informativeness (Experiment 1)



Note. One outlier has been removed.

## Appendix M

Plot of Residuals for Total Detail Recall (Experiment 1)



## Appendix N

Contingency Table for Lineup Accuracy: Target-Present Condition (Experiment 1)

	Accurate Selection	Inaccurate Selection
Stress	6	7
Non-stress	10	3

*Note:* Odds ratio = 0.257,  $p = .113$  (one-tailed)<sup>12</sup>

Contingency Table for Lineup Accuracy: Target-Absent Condition (Experiment 1)

	Accurate Selection	Inaccurate Selection
Stress	4	9
Non-stress	4	9

*Note:* Odds ratio = 1.00,  $p = .664$  (one-tailed)

<sup>12</sup> *Note:* Probability values were calculated using Fisher's Exact Test.

## Appendix O

Contingency Table for Lineup Accuracy: Stress Condition (Experiment 1)

	Accurate Selection	Inaccurate Selection
Target-Present	6	7
Target-Absent	4	9

*Note:* Odds ratio = 1.929,  $p = .344$  (one-tailed)

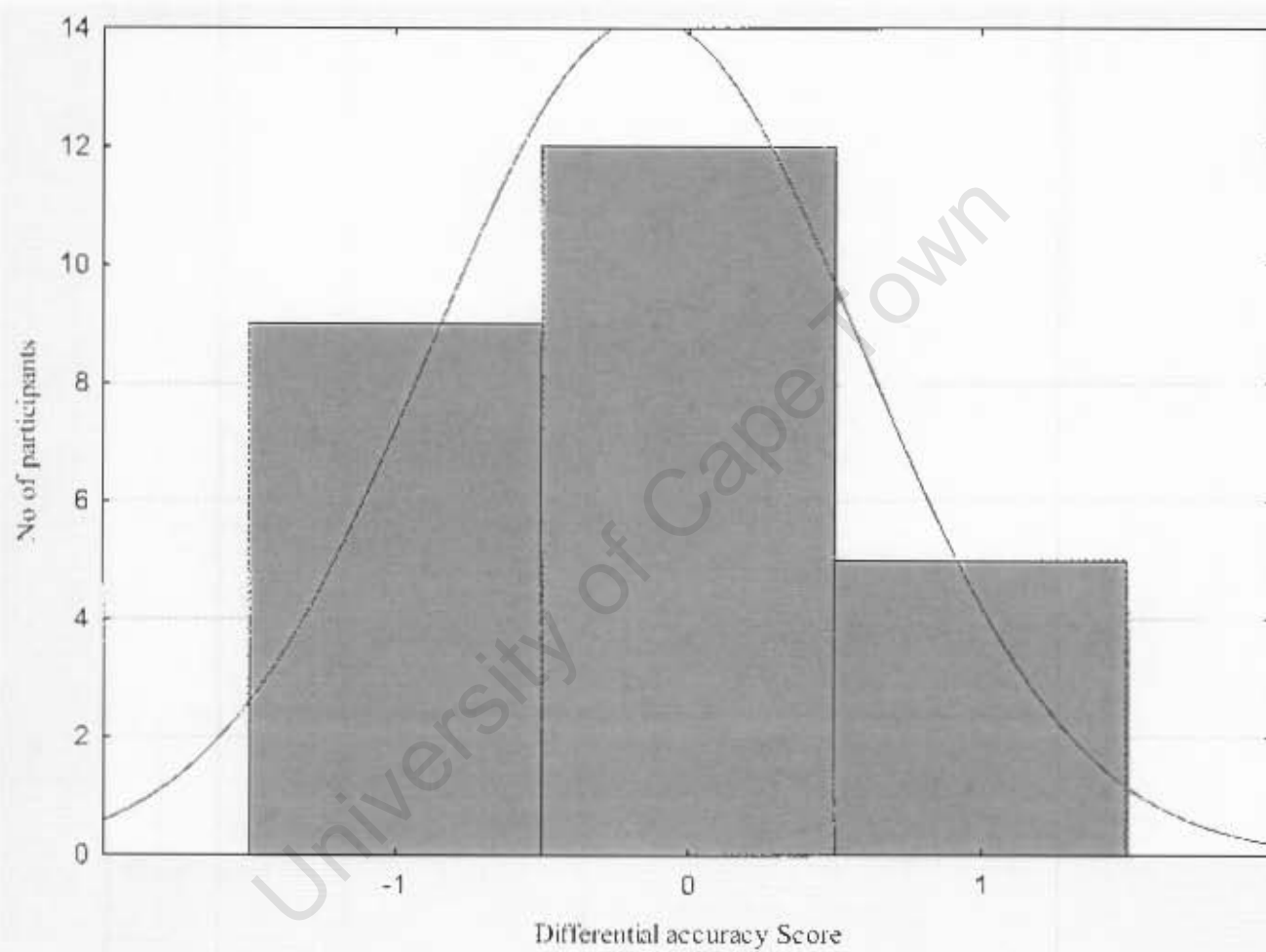
Contingency Table for Lineup Accuracy: Non-Stress Condition (Experiment 1)

	Accurate Selection	Inaccurate Selection
Target-Present	10	3
Target-Absent	4	9

*Note:* Odds ratio = 7.500,  $p = .024$  (one-tailed)

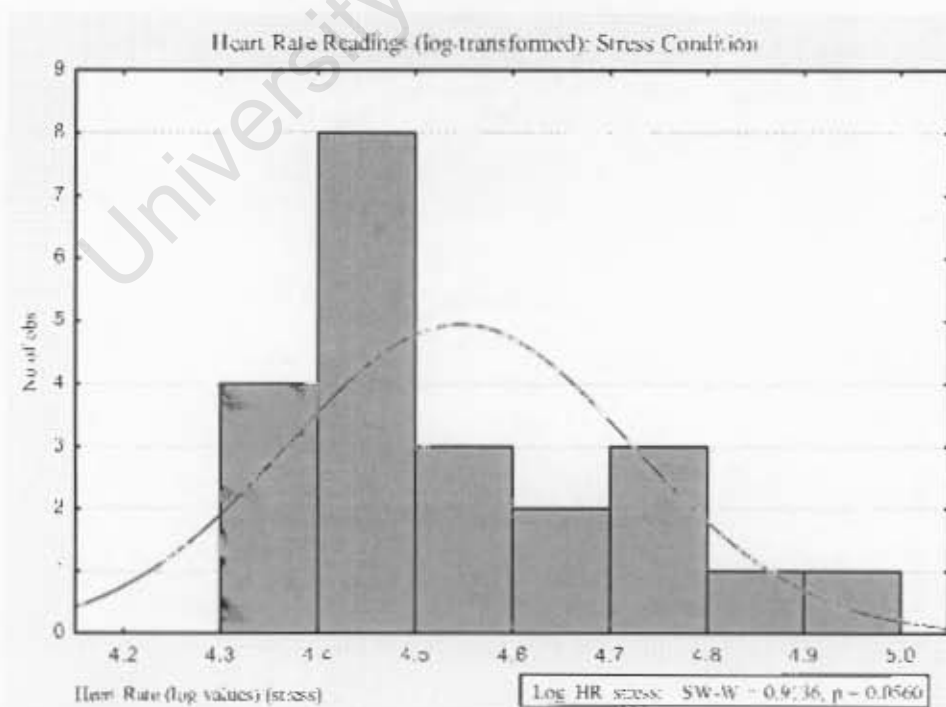
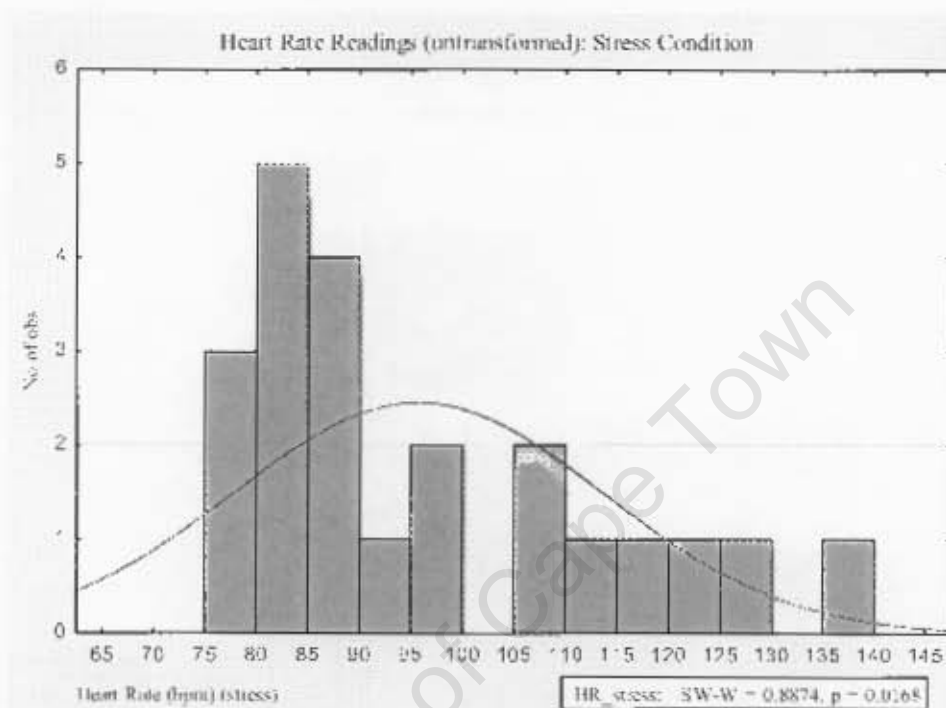
## Appendix P

Distribution of Differential Accuracy Scores (Experiment 1)



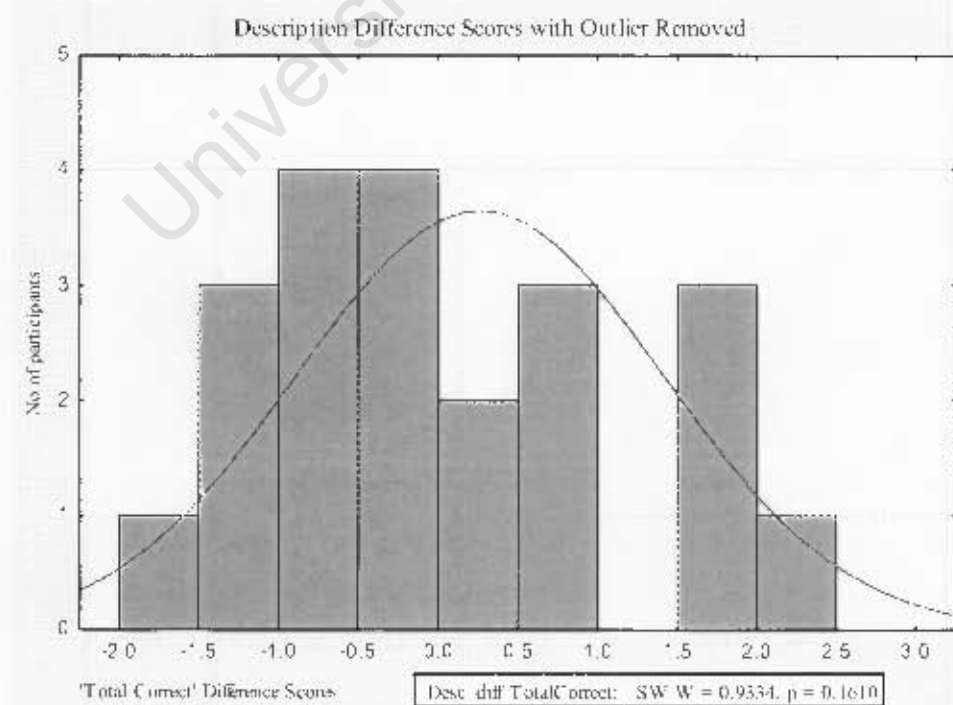
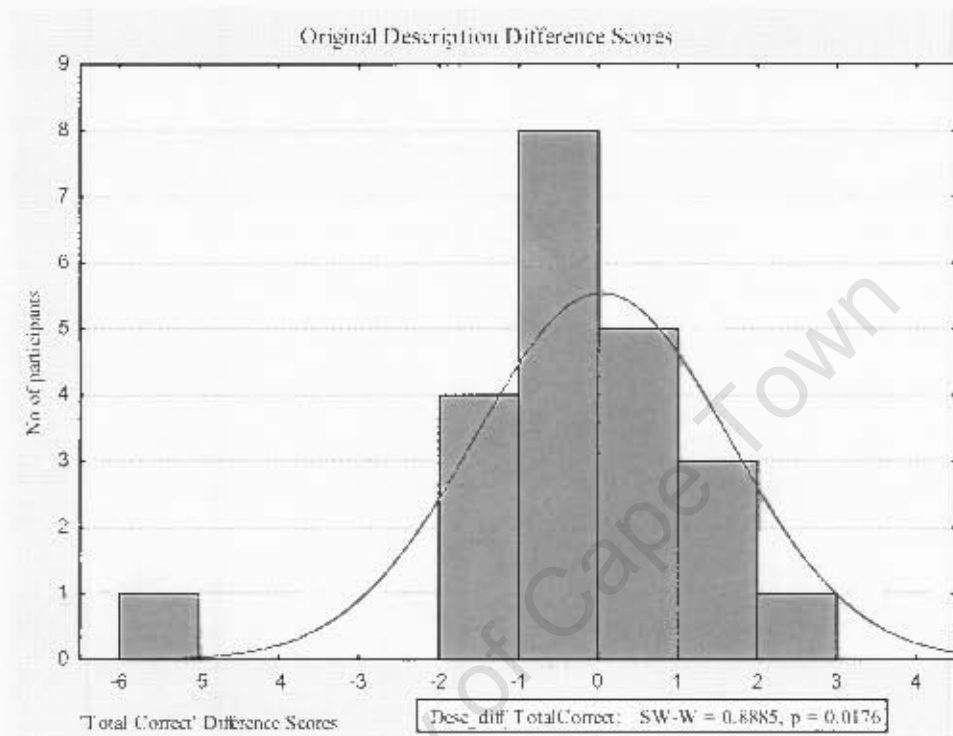
## Appendix Q

## Log Transformations of Heart Rate Data (Experiment 2)



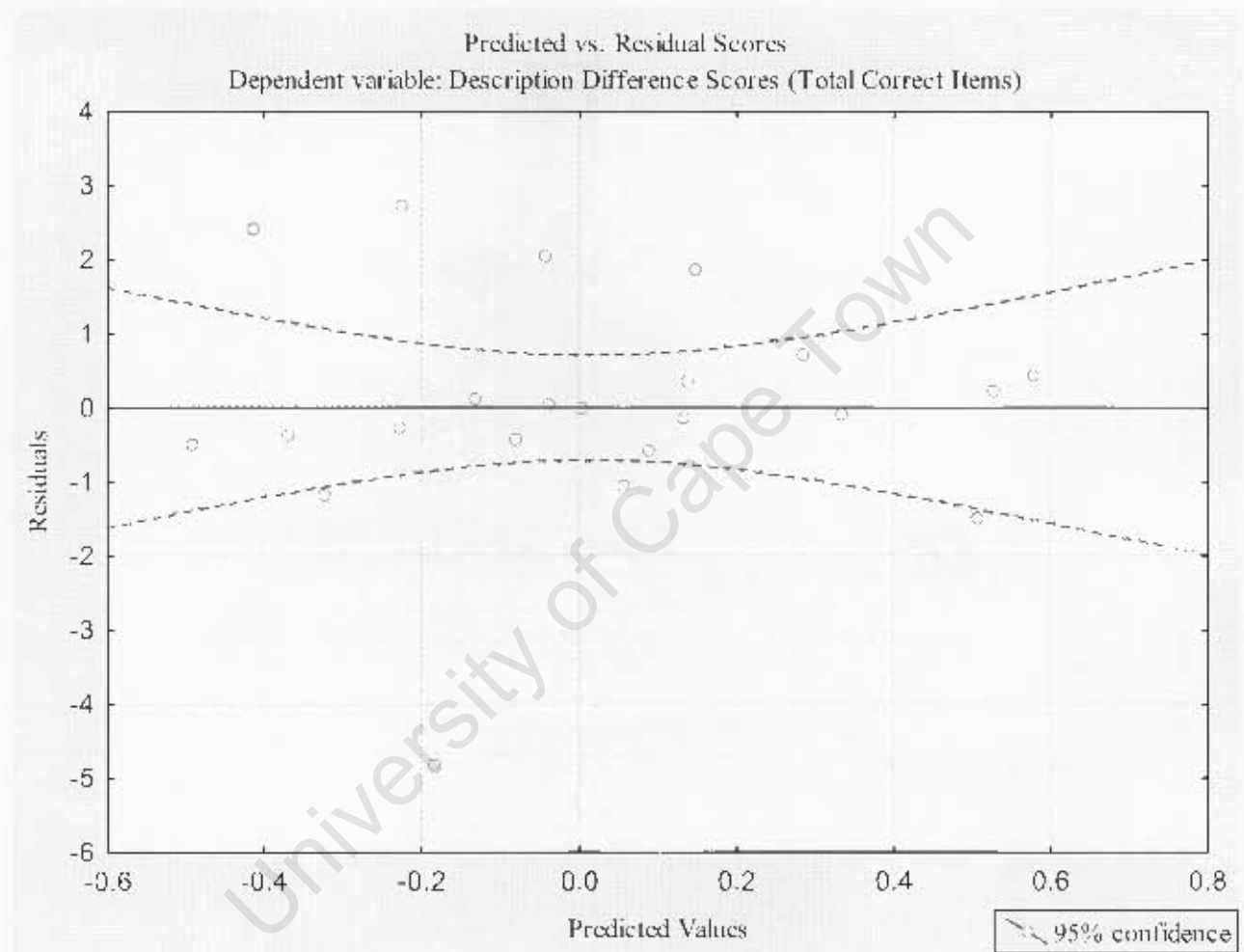
## Appendix R

## Distribution of Difference Scores for Total Correct Items (Experiment 2)



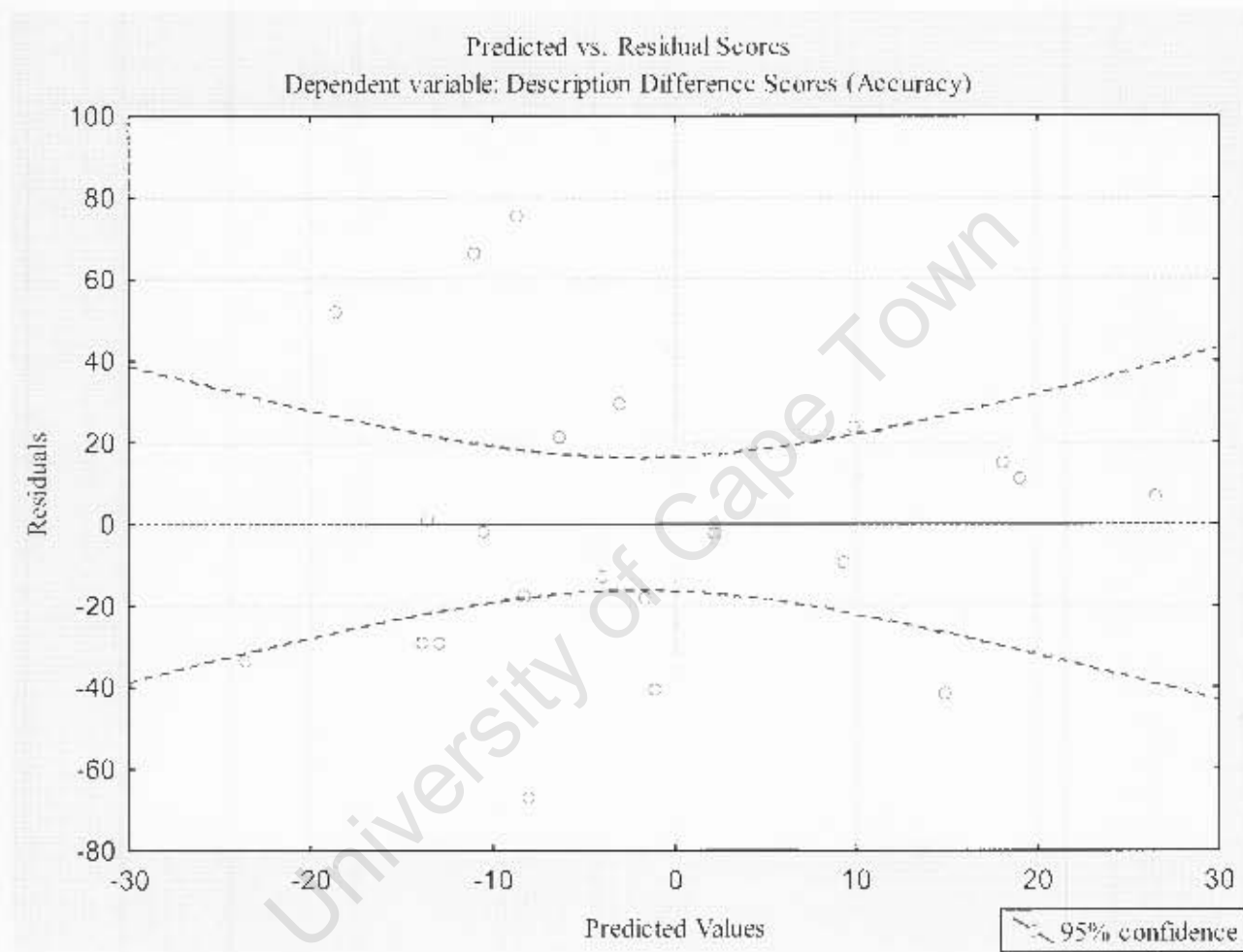
## Appendix S

Plot of Residuals for Total Correct Items (Experiment 2)



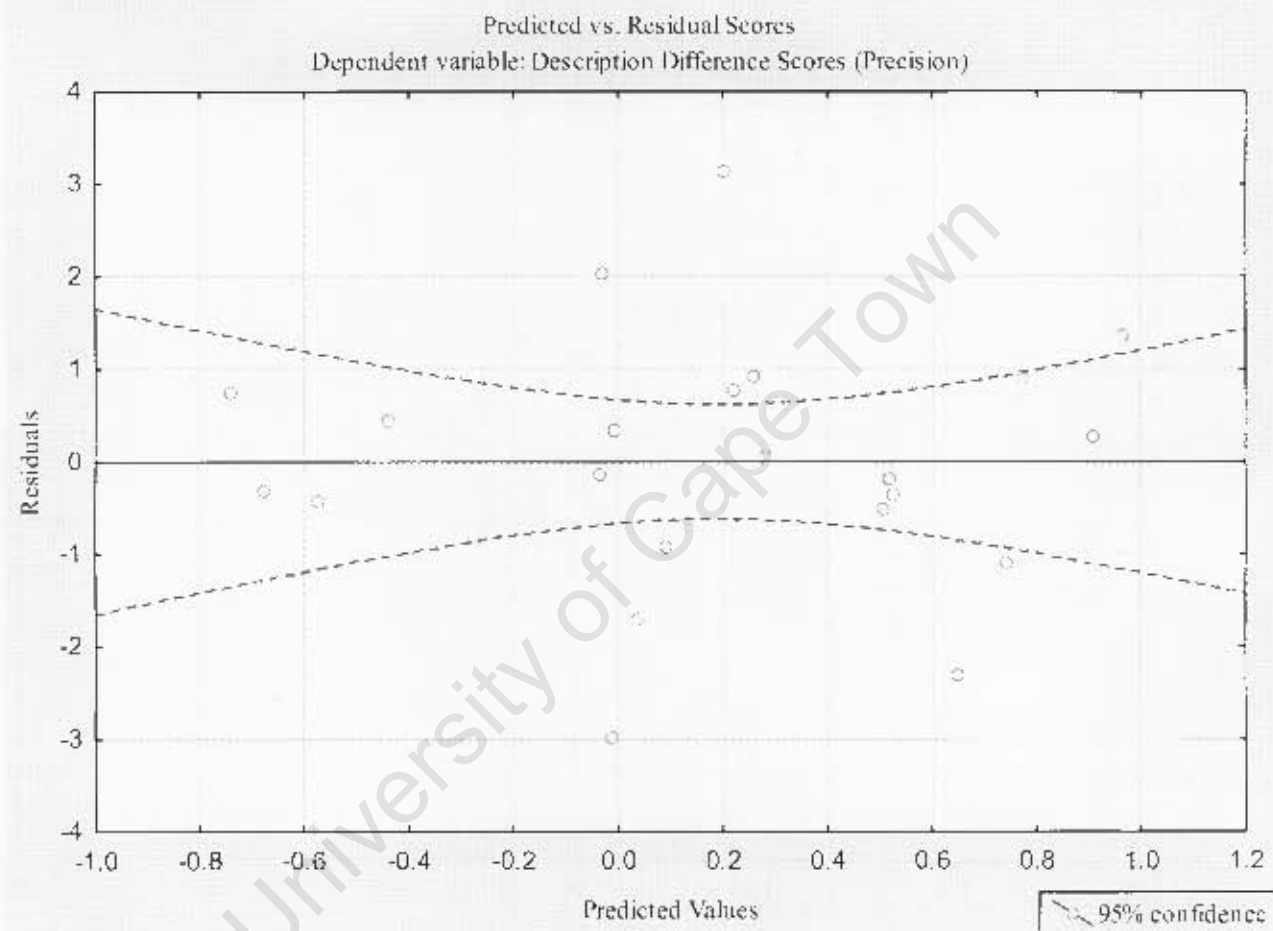
## Appendix T

## Plot of Residuals for Accuracy (Experiment 2)



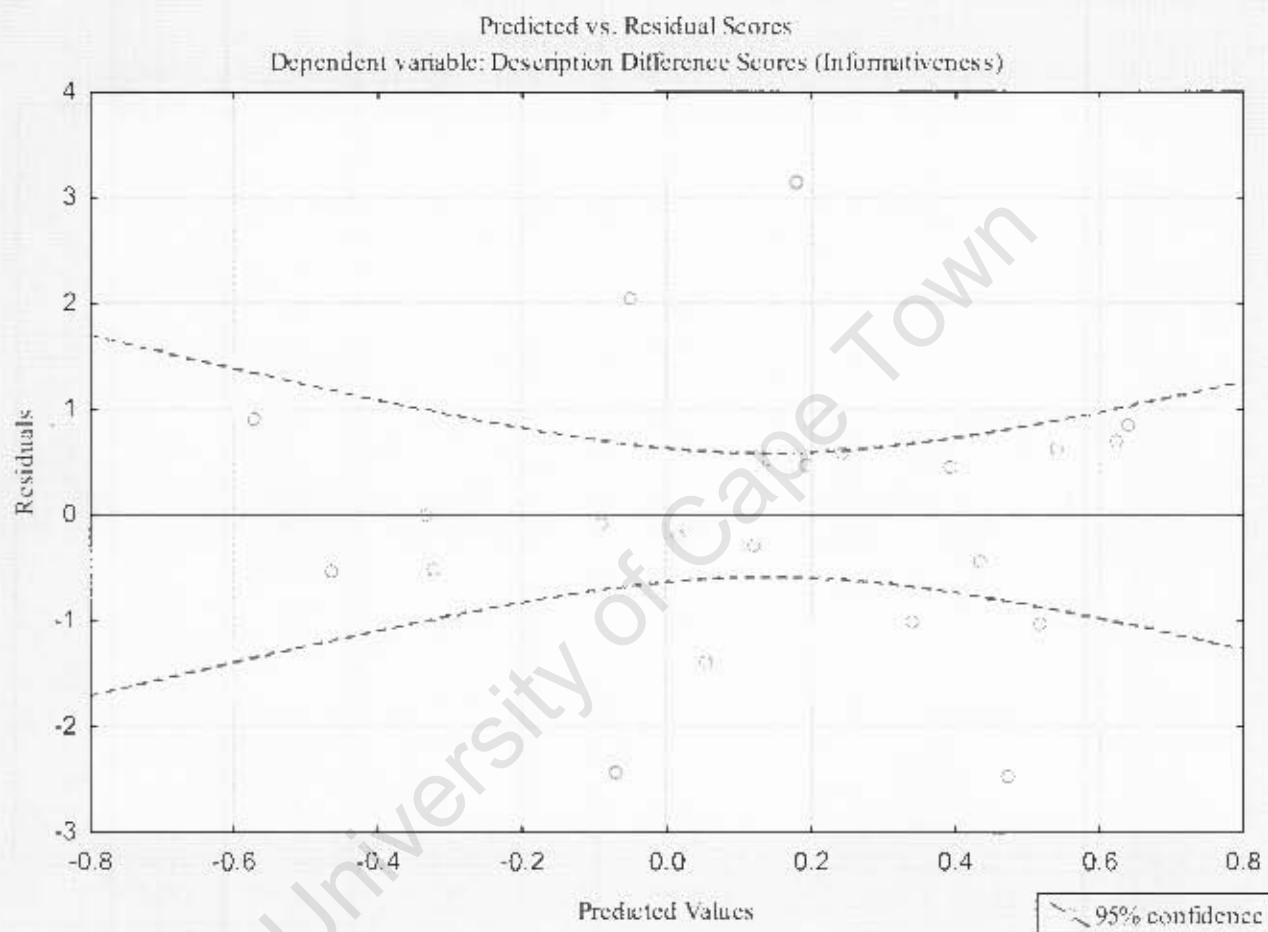
## Appendix U

## Plot of Residuals for Precision (Experiment 2)



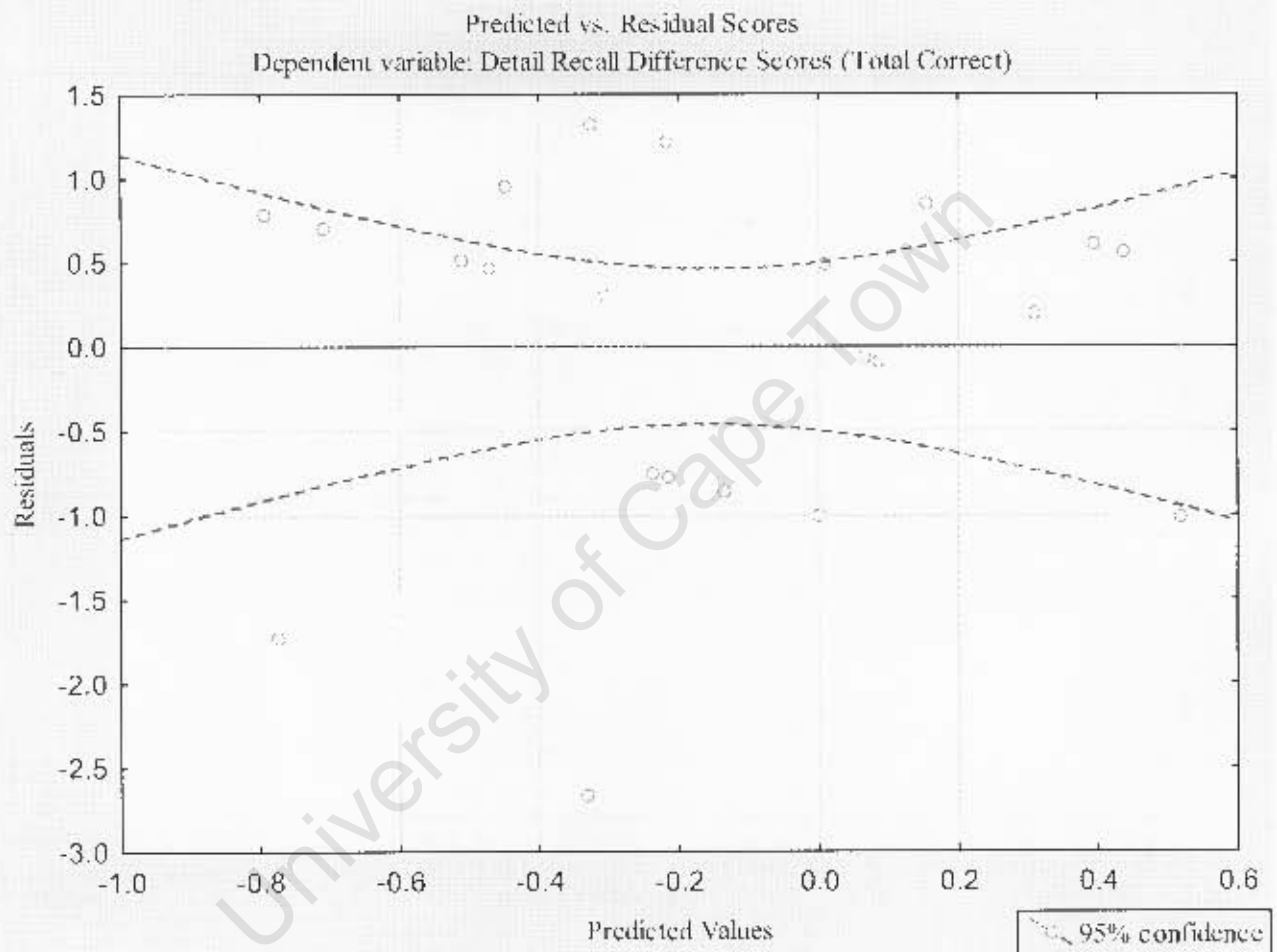
## Appendix V

## Plot of Residuals for Informativeness (Experiment 2)



## Appendix W

Plot of Residuals for Detail Recall Total Score (Experiment 2)



## Appendix X

Contingency Table for Lineup Accuracy: Target-Present Condition (Experiment 2)

	Accurate Selection	Inaccurate Selection
Stress	5	5
Non-stress	4	8

Note: Odds ratio = 2.00,  $p = .890$  (one-tailed)

Contingency Table for Lineup Accuracy: Target-Absent Condition (Experiment 2)

	Accurate Selection	Inaccurate Selection
Stress	4	8
Non-stress	2	8

Note: Odds ratio = 2.00,  $p = .881$  (one-tailed)

## Appendix Y

Contingency Table for Lineup Accuracy: Stress Condition (Experiment 2)

	Accurate Selection	Inaccurate Selection
Target-Present	5	5
Target-Absent	4	8

*Note:* Odds ratio = 2.00,  $p = .361$  (one-tailed)

Contingency Table for Lineup Accuracy: Non-Stress Condition (Experiment 2)

	Accurate Selection	Inaccurate Selection
Target-Present	4	8
Target-Absent	2	8

*Note:* Odds ratio = 2.00,  $p = .417$  (one-tailed)

## Appendix Z

Distribution of Differential Accuracy Scores (Experiment 2)

