TRUTHFUL VERSUS FABRICATED ACCOUNTS OF VICTIMIZATION:
A PROSPECTIVE INVESTIGATION OF THE CONSISTENCY OF TRAUMATIC
MEMORY REPORTS AND TRAUMA SYMPTOM PROFILES

by

Kristine A. Peace

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

at

Dalhousie University
Halifax, Nova Scotia
November 2006

© Copyright by Kristine A. Peace, 2006
NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author’s permission.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni les extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.
DALHOUSIE UNIVERSITY

To comply with the Canadian Privacy Act the National Library of Canada has requested that the following pages be removed from this copy of the thesis:

Preliminary Pages
Examiners Signature Page (pii)
Dalhousie Library Copyright Agreement (piii)

Appendices
Copyright Releases (if applicable)
Dedication

I would like to dedicate this dissertation to the individuals who shared their traumas with me, and to genuine victims (as a whole) who have been re-victimized by the criminal justice system due to misguided perceptions of credibility. May the search for "truth" be enhanced by this research.
# Table of Contents

List of Tables ................................................................................. x
List of Figures ................................................................................ xi
Abstract ....................................................................................... xiii
List of Abbreviations Used ............................................................ xiv
Acknowledgements ....................................................................... xvi

Chapter 1. Introduction ................................................................... 1

Chapter 2. Perspectives on Remembering .......................................... 8
  Classical Spatial Perspectives of Memory .................................... 8
  Modern Constructive Perspectives of Memory ......................... 12
  Defining Trauma ....................................................................... 17
  Models of Traumatic Memory .................................................. 19
    Traumatic Memory Argument ................................................. 20
    Trauma Equivalency Argument ............................................. 20
    Trauma Superiority Argument ............................................. 21
  Summary .................................................................................. 22

Chapter 3. Credibility and Memory .................................................. 24
  Defining Deception ................................................................... 25
  False Allegations ....................................................................... 26
    Historical False Allegations ............................................... 26
    Historical Approaches to Credibility Assessment .................. 28
    Modern Legal Perspectives on Allegations ......................... 29
  Statement Validity Analysis ...................................................... 32
    Historical Foundations ....................................................... 32
  Components of SVA ............................................................... 34
  Empirical Findings .................................................................... 39
  Factors Influencing CBCA Scores ............................................ 44
  Criticisms of SVA .................................................................... 47
Chapter 6. Introduction to the Present Study ................................................. 121
  Summary and Hypotheses ........................................................................... 121
  Study Design ............................................................................................... 126
  Conclusion .................................................................................................... 129

Chapter 7. Method ......................................................................................... 130
  Participants .................................................................................................. 130
  Measures ..................................................................................................... 131
    Memory Survey Instructions ...................................................................... 131
    Memory Report Forms .............................................................................. 132
    Consistency Questionnaire (CQ) .............................................................. 132
    Coding the Qualitative and Self-Reported Features of Memory Reports .... 134
    Revised Impact of Event Scale (IES-R) .................................................... 136
    Post-Traumatic Stress Disorder Checklist (PCL) ..................................... 137
    Trauma Symptom Inventory (TSI) ............................................................ 137
    Demographic Questionnaire ................................................................... 138
    Debriefing Form ...................................................................................... 138
  Procedure .................................................................................................... 138
    Time One .................................................................................................. 138
    Times Two and Three .............................................................................. 139

Chapter 8. Results ......................................................................................... 141
  Preliminary Analyses .................................................................................. 141
    MAP Reliability Check ........................................................................... 141
    Consistency Reliability Check .................................................................. 141
    Order of Narrative Reporting ................................................................... 142
    Themes of the Trauma Narratives ............................................................ 142
    Differences in Real and Fabricated Trauma Narratives ......................... 142
    Objective Narrative Features ................................................................... 145
    Subjective Narrative Features .................................................................. 147
    Objective and Subjective Features as Predictors of Narrative Type .......... 152
List of Tables

Table 1. Comparison of assessment criteria for criterion-based content analysis (CBCA), reality monitoring (RM), and the memory assessment procedure (MAP) .........................................................38

Table 2. Breakdown of categories of truthful (TT) and fabricated (FT) traumatic events reported .................................................................143

Table 3. Means (and standard deviations) for objective MAP features for truthful (TT) and fabricated (FT) narratives over Time 1 (T1), Time 2 (T2), and Time 3 (T3) ..................................................................................144

Table 4. Means (and standard deviations) for subjective EMS features for truthful (TT) and fabricated (FT) narratives over Time 1 (T1), Time 2 (T2), and Time 3 (T3) ..............................................................145

Table 5. Mean item and total scores (and standard deviations) for truthful (TT) and fabricated (FT) traumas across Times 1 and 2 (T1-T2) and Times 2 and 3 (T2-T3) on the CQ ........................................................................154

Table 6. Mean item and total scores (and standard deviations) for truthful (TT) and fabricated (FT) traumas over the extended Time 1 to 3 (T1-T3) interval on the CQ ................................................................................155

Table 7. Means (and standard deviations) for subjective EMS emotional features for truthful (TT) and fabricated (FT) traumas over Time 1 (T1), Time 2 (T2), and Time 3 (T3) .................................................................162

Table 8. Means, standard deviations, and severity category for truthful (TT) and fabricated (FT) traumatic events over Times 1, 2, and 3 (T1, T2, T3) for IES-R scores .............................................................................164

Table 9. Means, standard deviations, and percentages of possible PTSD for truthful (TT) and fabricated (FT) traumatic events over Times 1, 2, and 3 (T1, T2, T3) for scores on the PCL .............................................................................166

Table 10. Means and standard deviations for truthful (TT) and fabricated (FT) traumatic events over Times 1, 2, and 3 (T1, T2, T3) for TSI clinical scales ..................................................................................170

Table 11. Means, standard deviations, and percentages of endorsement for truthful (TT) and fabricated (FT) traumatic events over Times 1, 2, and 3 (T1, T2, T3) for TSI validity scales ...........................................171
List of Figures

Figure 1. Mean ratings of objective plausibility (1-7) for truthful (TT) and fabricated (FT) trauma narratives at Times 1, 2, and 3 (T1, T2, T3) (including the standard deviation from the mean indicated as vertical lines) ........................................... 147

Figure 2a. Mean ratings of self-reported vividness/clarity (1-7) for truthful (TT) and fabricated (FT) trauma narratives at Times 1, 2, and 3 (T1, T2, T3) (including the standard deviation from the mean indicated as vertical lines) ........................................... 149

Figure 2b. Mean ratings of self-reported memory quality (1-7) for truthful (TT) and fabricated (FT) trauma narratives at Times 1, 2, and 3 (T1, T2, T3) (including the standard deviation from the mean indicated as vertical lines) ........................................... 149

Figure 2c. Mean ratings of self-reported coherence (1-7) for truthful (TT) and fabricated (FT) trauma narratives at Times 1, 2, and 3 (T1, T2, T3) (including the standard deviation from the mean indicated as vertical lines) ........................................... 150

Figure 2d. Mean ratings of self-reported sensory details (1-7) for truthful (TT) and fabricated (FT) trauma narratives at Times 1, 2, and 3 (T1, T2, T3) (including the standard deviation from the mean indicated as vertical lines) ........................................... 150

Figure 2e. Mean ratings of self-reported confidence in memory accuracy (1-7) for truthful (TT) and fabricated (FT) trauma narratives at Times 1, 2, and 3 (T1, T2, T3) (including the standard deviation from the mean indicated as vertical lines) ........................................... 151

Figure 2f. Mean ratings of self-reported credibility (1-7) for truthful (TT) and fabricated (FT) trauma narratives at Times 1, 2, and 3 (T1, T2, T3) (including the standard deviation from the mean indicated as vertical lines) ........................................... 151

Figure 3a. Total consistency scores of truthful (TT) and fabricated (FT) trauma narratives over intervals between Times 1, 2, and 3 (T1-T2; T2-T3; T1-T3) (including the standard deviation from the mean indicated as vertical lines) ........................................... 157

Figure 3b. Total consistency scores of truthful (TT) and fabricated (FT) trauma narratives over the entire follow-up period between Times 1 and 3 (T1-T2; T1-T3) (including the standard deviation from the mean indicated as vertical lines) ........................................... 158

Figure 4a. Mean PCL total scores for truthful (TT) and fabricated (FT) trauma symptoms at Times 1, 2, and 3 (T1, T2, T3) (including the standard deviation from the mean indicated as vertical lines) ........................................... 167

Figure 4b. Mean PCL avoidance scores for truthful (TT) and fabricated (FT) trauma symptoms at Times 1, 2, and 3 (T1, T2, T3) (including the standard deviation from the mean indicated as vertical lines) ........................................... 167
Figure 4c. Mean PCL hyperarousal scores for truthful (TT) and fabricated (FT) trauma symptoms at Times 1, 2, and 3 (T1, T2, T3) (including the standard deviation from the mean indicated as vertical lines). .................................................... 168
Abstract

In forensic settings, decision-makers are continually challenged with evaluating the credibility of reports of victimization. The present study was designed for several purposes: (1) to assess the qualitative and phenomenological features of truthful and fabricated narratives of trauma using the Memory Assessment Procedure (MAP; Porter, Yuille, & Lehman, 1999), (2) to determine the level of consistency of both narrative types, (3) to assess symptoms of trauma associated with truthful and malingered reports, and (4) to determine the utility of all measures in discriminating the veracity of victimization accounts. Undergraduate participants were asked to describe both a truthful and fabricated traumatic experience at three time periods: Time 1 (N = 291, initial), Time 2 (N = 252, three months), and Time 3 (N = 181, six months). Measures of narrative features and trauma symptoms were administered at each phase. Results demonstrated that truthful traumas contained more detail, contextual information, emotional details, and were considered more plausible relative to fabricated traumas. The details of truthful narratives also were more consistent than fabricated narratives over an extended six month interval, although levels of consistency decreased over time for both narrative types. Fabricated symptoms were inflated and also may serve as indicators of possible deception, particularly in assessments of post-traumatic stress disorder. The findings of this dissertation indicate that the Memory Assessment Procedure has utility as a tool for deception detection in truthful and fabricated narratives, demonstrating similar levels of discrimination as established techniques such as Criteria-Based Content Analysis (CBCA) and Reality Monitoring (RM). Further research on MAP criteria and their application to statement analysis is warranted. This research is important for our basic theoretical and scientific understanding of memory and consistency of recall. Further, the current findings provide useful information for identifying possible false allegations.
List of Abbreviations Used

AA = Anxious Arousal
AI = Anger/Irritability
APA = American Psychological Association
ARJS = Aberdeen Report Judgment Scales
ATR = Atypical Responding
C1 = Criterion 1 in the CBCA Procedure (logical consistency)
C2 = Criterion 2 in the CBCA Procedure (unstructured production)
C3 = Criterion 3 in the CBCA Procedure (quantity of detail)
C5 = Criterion 5 in the CBCA Procedure (description of interactions)
C6 = Criterion 6 in the CBCA Procedure (reproduction of conversation)
C12 = Criterion 12 in the CBCA Procedure (subjective mental state)
C13 = Criterion 13 in the CBCA Procedure (attribution of perpetrator’s mental state)
CBCA = Criterion-Based Content Analysis
CSA = Child Sexual Abuse
CQ = Consistency Questionnaire
D = Depression
DA = Defensive Avoidance
DFA = Discriminant Function Analysis
DIS = Dissociation
DSB = Deviant Sexual Behaviour
DSM-II = Diagnostic and Statistical Manual of Mental Disorders, Edition 2
DSM-IV-TR = Diagnostic and Statistical Manual of Mental Disorders, Edition 4, Text Revision
EMS = Emotional Memory Survey
FT = Fabricated Trauma
IE = Intrusive Experiences
IES = Impact of Event Scale
IES-R = Impact of Event Scale – Revised
INC = Inconsistent Responding
ISR = Impaired Self-Reference
JMCQ = Judgment of Memory Characteristics Questionnaire
MANOVA = Multivariate Analysis of Variance
MAP = Memory Assessment Procedure
MCQ = Memory Characteristics Questionnaire
PAI = Personality Assessment Inventory
PCL = Post-Traumatic Stress Disorder Checklist
PDS = Posttraumatic Stress Diagnostic Scale
PTSD = Post-Traumatic Stress Disorder
REB = Research Ethics Board
RL = Response Level
RM = Reality Monitoring
SC = Sexual Concerns
SVA = Statement Validity Assessment
T1 = Time 1
T2 = Time 2
T3 = Time 3
TRB = Tension Reduction Behaviour
TSI = Trauma Symptom Inventory
TT = Truthful Trauma
UFO = Unidentified Flying Object
VCUG = Voiding Cystourethrogram Fluoroscopy
Acknowledgements

I would like to thank my supervisor, Dr. Steve Porter, for his assistance, advice, and financial resources throughout the duration of this project. I have learned many valuable lessons regarding research and academia under his tutelage, all of which have helped to shape my motivation and approach to both research and teaching in psychology. I also would like to thank my dissertation committee members, Drs. Tracy Taylor-Helmick and Patricia McMullen, for their helpful feedback, support, and willingness to be involved with this project. In addition, Drs. Sherry Stewart, Ray Klein, and Dennis Phillips have been formative in my success throughout this degree. This project also would not have been possible without funding from SSHRC awarded to Steve Porter as well as my doctoral fellowship.

There were many other people that I simply could not have accomplished this without. First, Naomi Doucette deserves a big thank you for attempting to keep me sane, while also doing massive amounts of copying, organizing, sorting, and talking me through a lot of stress. Second, I want to thank Kelly Emmett for running the pilot study for this dissertation as her honours thesis and putting up with my aimless ramblings. Third, thanks to both former and current members of the Porter Lab: Mary Ann Campbell for taking me under your wing when I first arrived in Halifax and helping me stand on my own two feet; Mike Woodworth for being a constant source of encouragement, advice, and amusement (let the giant lobsters live on!); Becky Douglas for understanding, listening, and getting me hooked on gingerbread men; and Brianna Cook for being a greatly patient honours student and friend, bringing me Café French, and helping wherever possible. Finally, to all volunteers that spent countless hours doing monotonous coding or scoring of questionnaires: Allison Eisner, Julie Nowlan, Leanne ten Brinke, Juli Kodagoda, Dan Almon, Elisha Timmons, Jennifer Heatley, Ann Marie Penny, Krista Merrill, Laura Sanschagrin, Allison Kearns, Michael Gallant, Lydie Woodfield, and Nicki Davis – your assistance was invaluable!

Finally, I would like to thank those that supported me and this project on a totally different level. Mom and Dad – I can never thank you enough. Your support, assistance, proof-reading, encouragement and love have been endless. Thanks for lending a hand whenever I stumbled, and for the successful establishment and implementation of the
“Committee for the Westernization of Kris” (CWK). You are an integral part of the “disstory”. To my friends who are dear to my heart: Tanya Leavitt – you have helped me to maintain a respectable level of motivation and insanity (let the Grindy Lizzard Beef Star Rage On!) and have provided invaluable support and humour. Danille Lazzaretto – you’ve always been “her” for me, sending me a finite number of monkeys, and helping me to believe this was all worthwhile. Allison Eisner – this was not possible without you. Not only were you a willing “minion”, but talked me through the countless hours of work and exhaustion (and random sleep-deprived discussions about moustaches). You rock my socks!! Brianna Cook – thank you so much for your help with the symptom data, dinners, yummy leftovers, and, of course, the Ace of Spades! Larry Phillips – thanks for participating in the beer-work accountability program, study skills fun (may the sock puppet live on!), the countless games of pool, and for lending a continual listening ear. Randy, Art, Doris Burrill (and Melissa Bailey and Kerri Pierce) – I love you dearly and you have challenged me to be the strong woman you believed me to be and always supported me in the small but meaningful ways that make the heart smile. Finally a thank you to the constant company of my cats Amos and Abby, even though typing around you was sometimes difficult. And as always, Enya, John Mayer, and Eva Cassidy – thanks for helping me find the Fields of Gold. This dissertation has been brought to you by the letter K.
Chapter 1. Introduction

"Now, these do either wholly contrive and invent the untruths they utter, or so alter and disguise a true story that it ends in a lie. When they disguise and often alter the same story, according to their own fancy, 'tis very hard for them, at one time or another, to escape being trapped, by reason that the real truth of the thing, having first taken possession of the memory, and being there lodged impressed by the medium of knowledge and science, it will be difficult that it should not represent itself to the imagination, and shoulder out falsehood, which cannot there have so sure and settled footing as the other; and the circumstances of the first true knowledge evermore running in their minds, will be apt to make them forget those that are illegitimate, and only, forged by their own fancy. In what they, wholly invent, forasmuch as there is no contrary impression to jostle their invention there seems to be less danger of tripping; and yet even this by reason it is a vain body and without any hold, is very apt to escape the memory, if it be not well assured."

~ Michel de Montaigne, *Essays*, Book the First, Chapter IX: Of Liars
Throughout history, false claims of victimization have occurred in legal settings (e.g., Yuille, Tymofievich, & Marxsen, 1995). Given the complexity of evaluating the credibility of a complainant’s statement, false allegations present a significant challenge for the legal system. With this recognition, psychology has a long tradition of investigating witness credibility, dating to the early 1900s with Hugo Münsterberg’s *On the Witness Stand* (1908). In this text, he argued that a psychological analysis of memory can assist in the discernment of truth in legal contexts. One source quotes him as stating: “To deny that the experimental psychologist has indeed possibilities of determining the ‘truth-telling’ process is just as absurd as to deny that the chemical expert can find out whether there is arsenic in a stomach or whether blood spots are of human or animal origin” (as cited in Hale, 1980, p. 118). Although more than a century has passed since Münsterberg’s controversial publication (which was ultimately vilified in the legal community; e.g., Goldstein, 1980), difficulties in determinations of credibility continue to persist in our legal system. In the 2001 inquiry into Thomas Sophonow’s wrongful murder conviction in Manitoba, Justice Cory highlighted this point. He concluded that (1) witnesses often provide dishonest testimony, including about their own criminal victimization and (2) injustices have resulted from errors by “honest, right-thinking eyewitnesses.” He strongly advocated the need for a better understanding of deception and memory errors among witnesses or alleged crime victims (Justice Cory, November 4, 2001). Some recent cases will illustrate this point.

In one prominent local case, the Nova Scotian government awarded various sums of money to former residents of the Shelburne School for Boys and the Youth Training Centre in Truro following claims of historical abuse by staff (see Kaufman, 2002; Porter,
Campbell, Birt, & Woodworth, 2003). Many of these claims were later found to be false allegations motivated by monetary gain. On June 6, 2006, legal actions launched by former teachers and guards against the Nova Scotia government were finally put to rest. Over $7 million dollars will be paid out to the 79 staff members who were falsely accused of abusing students to compensate for over 10 years of hardship caused by these claims (CBC News, June 6, 2006). While individuals may make false allegations of victimization for financial gain, attention-seeking and revenge also are strong motivators. For example, former alderwoman Dar Heatherington of Lethbridge, Alberta, was convicted of public mischief by sending anonymous threatening letters to herself and falsely claiming to have been abducted, sexually assaulted, and abandoned in Los Vegas (Graveland, 2004; Regina v. Heatherington, 2004a, 2004b). Similarly, the claim that 20-year-old Audrey Seiler was abducted at knifepoint on March 27, 2004, in Wisconsin has been discounted by police as being fabricated, partially due to inconsistencies in her reports of victimization. The college student has been charged in faking her own disappearance, and investigation continues into a previous claim made by Seiler that she was assaulted and left unconscious on the street by an unknown assailant in February of that year, prior to her alleged abduction (CNN, July 1, 2004). Jennifer Wilbanks, known as the “Runaway Bride”, is another well-known example of fabricated victimization. Wilbanks mysteriously disappeared just prior to her wedding in April 2005. After a massive 4-day manhunt, Wilbanks called police from a payphone in Albuquerque, New Mexico, alleging she had been abducted from Atlanta and sexually assaulted. She later recanted during police questioning and was charged with making a false statement and falsely reporting criminal behaviour (CNN, May 26, 2005). She pleaded no contest to the
charges and received probation, community service, and fines for the upheaval she caused (Associated Press, June 22, 2005). These last few cases exemplify successful credibility assessments where false allegations of victimization were discounted without lengthy delays. Unfortunately, many false claims remain undetected for years or even decades (e.g., Porter, Campbell, et al., 2003).

The consequences of such false claims of victimization include extensive resources required for the investigations, distress or trauma for falsely accused persons, and increased scepticism concerning claims by genuine victims of criminal events. Estimates of false allegations of victimization have ranged from 2-40% of all cases (e.g., Benedek & Shetky, 1985; Everson & Boat, 1989; Green, 1986; Halliday, 1988; Jones & McGraw, 1987; Kanin, 1994; Littman & Szewczyk, 1983; Mann, 1985; Sheridan & Blaauw, 2004; Spencer & Flin, 1990; Thoennes & Tjaden, 1990). For example, in their analysis of 350 abuse referrals, Anthony and Watkeys (1991) found that 63 of the 153 unsubstantiated claims (i.e., 41%) were false allegations made either by the child or the parent (the latter often during custody disputes). In another study, Penfold (1995) evaluated allegations of sexual abuse arising in approximately 2% of child custody disputes and reported fabrication rates of 8% to 16.5%. In a recent interview reported in the Chronicle Herald, head of the Nova Scotia Criminal Lawyers Association, Josh Arnold, was unable to provide an estimate of the rate of false allegations of sexual assault. However, he stated that sexual assaults and domestic violence are “very easy to fabricate” because they are private experiences (Bradley, May 9, 2006). Although one may assume that it is easy to sort fact from fiction, no ‘Pinnochio’s nose’ has been found
(e.g., DePaulo et al., 2003) and the extent of convictions stemming from false allegations remains unclear (Zaparniuk, Yuille, & Taylor, 1995).

There also has been a dramatic increase in civil claims involving post-traumatic stress disorder (PTSD) in recent years (e.g., Guriel & Fremouw, 2003). With the potential for both external gain (e.g., financial compensation) and psychological gain (e.g., attention) from reporting a false traumatic experience (e.g., Yuille et al., 1995), the validity of such claims must be carefully evaluated. According to the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR, American Psychiatric Association, 2000), the prevalence of genuine PTSD in the general population is estimated at 8%. However, research by Resnick (1997) indicated that malingering (i.e., intentionally feigning the disorder) occurs in approximately 16% of reported cases in forensic populations and 7% in non-forensic populations. Due to the subjective nature of the experience of trauma (and subsequent symptomology), PTSD symptoms may be particularly easy to mangle (Guriel & Fremouw, 2003). Some psychological disorders, including PTSD, typically are assessed based on the nature of the traumatic event reported and on post-trauma symptoms. Adding to the complexity of distinguishing true from fabricated claims of trauma and victimization is that individuals may not necessarily completely fabricate but, instead, exaggerate true symptoms (e.g., Geraerts, Jelicic, & Merckelbach, in press; Rogers, Jackson, Sewell, & Salekin, 2005). In this context, legal decision-makers require an awareness of how people report fabricated traumatic events and how to distinguish them from accounts of genuine traumatic experiences.

Despite the media focus on false allegations and the obvious need to identify malingered PTSD in civil cases, there has been little scientific consideration of fabricated
claims of victimization. Although some studies have addressed the characteristics of true and mistaken (but honest) memories for traumatic experiences, few have focused on the qualities of intentionally fabricated (deceptive) reports of trauma and victimization (e.g., Porter, Yuille, & Lehman, 1999). Some research indicates that reports of fabricated childhood experiences made by adults differ qualitatively from true experiences. Fabricated reports tend to be more detailed, contain more repeated details, and are associated with higher ratings of vividness and confidence (e.g., Porter et al., 1999).

Would fabricated reports of trauma in adulthood contain qualities similar to those of real traumatic experiences or might they have an exaggerated quality? Further, little is known about the “symptoms” exhibited by people who intentionally fabricate a traumatic experience. Do malingerers endorse different types and levels of symptoms than those who are genuinely suffering? Also relevant to victimization claims is the manner in which allegations are maintained over time, both in memory and symptom reporting. Little research has addressed the relative consistency of truthful and deceptive claims of trauma (e.g., Granhag & Strömwall, 2000a).

While studying this issue has clear importance for the basic understanding of memory (e.g., retention of emotional events over time) and deception, it also has applied forensic relevance. Although courts of law traditionally have relied heavily on the testimony of witnesses in formulating their decisions of guilt and innocence, it is now recognized that both intentional deception and mistaken recollections are all too common in the courtroom (as noted by Justice Cory in the Sophonow Inquiry). For example, in numerous cases of sexual abuse and assault, it has been established that the evidence was based on false memories (Loftus, 2003a) and intentional fabrications (Raitt & Zeedyk,
2003). This dissertation was designed to address these issues by empirically examining real and fabricated memory and symptom reports concerning a range of traumatic experiences and with a prospective, longitudinal component.

In order to establish the aims of this dissertation, it is necessary to first review the origins of views on memory in general, and traumatic memories specifically. I then review the literature on the relevant issues of credibility, consistency, and symptomology.
Chapter 2. Perspectives on Remembering

"It is singular how soon we lose the impression of what ceases to be constantly before us. A year impairs, a lustre obliterates. There is little distinct left without an effort of memory, then indeed the lights are rekindled for a moment - but who can be sure that the imagination is not the torch-bearer?" (Lord Byron, 1821-22/1979)

For several thousand years, human ability to accurately store and reliably retrieve records of our experiences has been a major subject of critical thinking and open discussion in the fields of philosophy, psychology, and law. Many scholars have made significant contributions to understanding the nature of human memory (Searleman & Herrmann, 1994). Prior to the twentieth century, the creation and recollection of mental images (i.e. memory) was primarily a subject of philosophical discussion, and not of scientific experimentation. This chapter provides a general overview of historical perspectives on remembering, and includes modern day arguments concerning the retention of trauma in memory.

Classical Spatial Perspectives of Memory

During historical periods of Greek and Roman influence, human memory was considered to be a virtual reproduction of the original sensory experience (Sutton, 1998). The concept of mind as a mental space, in which memories are retained and retrieved by searching that space, served as a broadly accepted explanation of remembering for many
centuries (e.g., Roediger, 1980). The ancient Greeks devoted much mythological and philosophical meditation to the subject of remembering. One of the primeval gods of nature was Mnemosyne, the goddess of memory, who was responsible for overseeing all mental activities (Hamilton, 1969; Lines, 1973). The Greeks essentially believed that memories were stored intact, somewhere in the mind, and that deficiencies in memory recall indicated faulty or reduced intellectual capacity (Sutton, 1998). This spatial concept of memory as a “cavern of wax” was successively advanced by Socrates, Plato, and Aristotle. Socrates believed that there is a wax tablet in the mind “which is larger in one person, smaller in another, of pure wax in one, more muddy in another, harder in some, softer in others, and in some just right” (Socrates, 369 BC/1972 AD, p. 186). Plato, however, believed that memories were stamped on the soul, and that conscious awareness of one’s past could be drawn from this reservoir of prior experiences (Knoche, 2001). Supplementing the propositions of his two predecessors, Aristotle (c. 335 BC/1908 AD) considered the formation of memory, or “sensory stimulation,” to be the result of the physical mark of perception. He suggested that “the act of perception stamps in, as it were, a sort of impression of the percept (on the soul), just as persons who make an impression with a seal” (p. 65). He proposed that those individuals who demonstrated poor memory abilities (i.e., young and old persons) had defective receiving surfaces, as if the seal were stamped on a watery surface. Memory was therefore contingent on both the quality of the stamped perception and the ability to retrieve the perceived material.

Another highly significant spatial conceptualization of memory was proposed by Augustine (354-430 AD), who argued that memory was similar to a “storehouse”, as opposed to the “wax cavern” notion of his predecessors (Herrmann & Chaffin, 1988). His
landmark treatise was titled *Augustine: Confessions and Enchiridion*, in which he discussed his own recollections of youth, and argued that all events are encoded then permanently stored in their original form. Augustine wrote of the “vast cave of memory with its numerous and mysterious recesses” which receives events and “stores them up, to be recalled and brought forth when required” (p. 112, as cited in Herrmann & Chaffin, 1988). Like Aristotle, Augustine attributed memory errors to difficulties in the retrieval process. These early spatial views of memory (wax cavern and storehouse) implied that the characteristics of recalled memories should be similar to, or at least an inverted reflection of, the actual events as originally perceived. The exact nature of an event could be inferred from one’s own memory experiences, or from listening to another person’s verbal description of his/her experience, with appropriate allowances for the mental capacities of the rememberer. Augustine believed that stressful or emotional events could be retrieved and literally communicated to another person, with little or no consideration of the human aptitude for creation of false memories. From a historical perspective, the one notable exception to the storehouse concept was Plotinus (205-270 AD), who rejected the views of Plato, Socrates, and Aristotle. In his writing of *The Enneads* (c. 244 AD/1962), Plotinus included the dissenting chapter, “Memory is not passive, but is an active power of the mind,” in which he argued that sensory perceptions are nothing like wax impressions, but that memory is a *rational* force of the soul, “the reasoning power of the universe.” Although essentially ignored for hundreds of years, his discussion marked an important philosophical transition from considering memory as a spatial imprint to a dynamic constructive entity.
Despite this exception (Plotinus), spatial concepts of human memory continued to dominate philosophical thought throughout the Renaissance period. For example, English philosopher John Locke believed, like Augustine, there to be nothing in the intellect except what had been originally experienced through the senses (Durant, 1961). According to William James (1890/1950, p. 670), a memory impression could be sufficiently emotionally arousing to almost “leave a scar upon the cerebral tissues.” And David Hume, in agreement with Aristotle’s classic view of memory, described perception as a process in which an impression is made upon the senses, after which the mind created a copy called an “idea.” Over time, the idea remained in essentially original form, but as a faint image, somewhat faded like a dust-covered portrait. The original impression and the idea differed only in their “liveliness.” Hume’s three-volume *Treatise on Human Nature* (1737/1886) contained his assertion that imagined events were inferior in their vividness and endurance as compared to memories for real events. But at the threshold of more modern constructive views of memory, Hume also recognized, in addition to the passive wax concept of memory, the power of human imagination to create memory-like images in the mind. He was one of the earliest writers to specifically compare the qualities of memory with the characteristics of imagination.

In the late nineteenth century, a school of psychological reasoning called psychoanalysis advanced the spatially-oriented position that certain childhood events become repressed or buried in the unconscious. Repression was conceived by Freud (1922) as a place in the mind which he considered to be a back room, or “ante-room”, guarded by a door-keeper. Federn (1952) summarized the psychoanalytic view of memory - that recollections of childhood events somehow remained preserved in a latent
state until re-experienced later in life. Psychoanalytic traditions were highly influential in both the psychological and biological study of the mind. For example, in the 1950s, Wilder Penfield reported that long-buried memories could be elicited by stimulation of localized brain sites, particularly in the temporal lobes (Penfield, 1952). After asserting that memory was essentially a reproduction or reinstatement of the original sensory experience, and available for retrieval, Penfield commented, “It has fallen to my lot, during explorations of the cortex, to demonstrate a mechanism in the human brain which preserves the record of the stream of thought” (Penfield, 1954, p. 47). The impact of Penfield’s research eventually weakened as problems surfaced about Penfield’s experiments (e.g., Loftus, 1993; Malcolm, 1977). For example, only a small minority (40/530) of his patients reported “memories” on stimulation of brain sites, and his methodology did not provide a reliable basis to determine the accuracy of those “memories” that were reported. And as recent as 1968, an influential paper by Atkinson and Shiffrin (1968) confirmed the “storehouse” concept with the proposition that long-term memory is a permanent storage system and a variable, or labile, retrieval system. And to conclude this discussion of the classical memory perspectives, Wickelgren (1977), in the face of mounting evidence confirming the dynamic nature of memory, concluded that: “So far as we know, time does not ‘rewrite the lines’ in our memories, it only makes what is written on some lines fade.” (p. 392).

Modern Constructive Perspectives of Memory

During the Renaissance and Enlightenment periods (16\textsuperscript{th} - 18\textsuperscript{th} Centuries), an abandonment of the spatial / passive perspective regarding memory commenced. It was
in this atmosphere of scientific exploration that a major evolution of memory theory began to emerge. During this period, memories were no longer assumed to reside somewhere in the mind or soul, or considered to be a passive receptacle or reflection of sensory inputs (Hoerl & McCormack, 2001; McNally, 2005). On the contrary, human memory came to be seen as an active, creative force governed by reason and judgement.

The principal initiators of this alternate perspective were two German philosophers – Gottfried Leibnitz (1646-1716) and Immanuel Kant (1724-1804). They advanced the idea that memory was subjective and creative, rather than passive and objective; a concept that was still on the fringe of enlightened reasoning during this period. Although he viewed memory as an act of the soul, Leibnitz (1704/1908) wrote that memory organizes future behaviour in a rational way. He observed the importance of active reasoning and interpretation, as well as the influence of social factors on memory, evidenced in his statement, “I might remember more distant things, the testimony of others might fill the gap of my remembrance” (p. 231). Leibnitz also was one of the first writers to acknowledge memory distortion: “It is true that if others conspired to deceive me (as I might even be deceived by myself, by some vision, dream or illness, believing that what I dreamed had happened to me), the appearance would be false” (p. 231). Leibnitz challenged the traditional idea that remembering was related solely to retrieval ability, and in so doing introduced discussion relating to potential distorting influences on memory, such as imagery, dreams, and source confusion.

In his work Critique of Pure Reason (1781/1966), Kant expanded upon the ideas of Leibnitz in several significant areas. He proposed that a person’s perceptions of life experiences inform him/her about the environment through individual sensations. In
trying to establish truths about themselves and nature, humans employ cognitive reasoning to interpret and organize both current and past sensory experiences. Kant rejected the passive wax and storehouse concepts established hundreds of years earlier by Aristotle and Augustine. On the contrary, he asserted that the mind was an active entity shaping and coordinating sensations into thoughts, ideas, and memories in order to create a rational theory of the environment (Durant, 1961). Kant developed a hierarchical opinion of human cognitive systems: sensation was unorganized stimulation (Empfindung), perception was organized sensations (Anschauungen), and knowledge/memory was organized and constructed perceptions (Verstand). Kant considered memory to be a highly creative dimension of human cognition, and not just a collection of reinstated perceptions. In his own words, the “productive imagination” functions to “change the manifold of intuition into an image; it must therefore first receive the impressions into its activity, which I call to apprehend” (Kant, 1781/1966, p. 112). With publication of Critique of Pure Reason, Kant established a significant theoretical framework to explain the constructive nature of memory, and to understand the mind in general.

As a result of these revised memory concepts (Leibnitz and Kant), the qualitative characteristics of memories were no longer considered to be the same as perception (reinstatement), but images of events resulting from perception, and contingent on rational cognitive processes. In this perspective, memory was related more to the person doing the remembering than to the event being remembered, and far less objective than previously held in the classical view. The quality of the memory image would, in addition to the original perception, reflect both the reasoning capacities and life experiences of the
person (Sutton, 1998). Thus, the memory image for the event itself was considered to be the object of rational reconstruction, but also the product of subjective mental activities introduced by the person(s) who observed the event.

Although essentially adhering to Aristotle’s wax cavern concept of memory (as discussed in the previous paragraphs), British philosopher David Hume also observed the power of human imagination to create memory-like images in the mind. Hume was one of the earliest writers to specifically consider the characteristics and qualities of memory and imagination. He adopted the belief that ideas of memory were “lively and strong” (as compared to imagination) from Berkeley, who had initially proposed that memories are characterized by a "steadiness, order, and coherence" not existing in the imagination (cited in Hume, 1739/1964, p. 138). Hume believed that memories (being reproductions of original events) were more logical and coherent than imagination (see Porter et al., 1999, for a modern validation of this view), whereas imagined events were subject to a high level of creativity and re-arrangement (Copleston, 1964). It appears that Hume viewed imagination as a volitional and conscious process, although he did not specifically embrace the subjects of memory distortion or false memories.

Further development of memory theories continued throughout the 18th and 19th Centuries, with one significant contribution being an essay by Samuel Dunnett in 1862 which allegedly included one of the first uses of the term “false memory” (p. 67). This writing provided the first explicit discussion of memory distortion. According to Dunnett, memory distortions had three primary origins, which parallel, to a remarkable degree, explanations given by modern researchers: (1) lied-about events that became believed by the liar, (2) a "weak will" which erroneously attributed truth to an imagined event; and (3)
mistaken details added to memory after ordinary forgetting. Although this dissertation does not concern the creation of false (mistaken) memories, Dunnett’s discussion of memory for lied about and genuine events provided the first detailed description of how people judge whether a personal memory is based on reality or imagination (known as “reality monitoring” by modern researchers). He felt that people generally have little difficulty in distinguishing fact from fiction because a factual event has “been perceived, and felt and known” (p. 57) and becomes an established truth. Dunnett’s discussion of the “honest forgetful memory” centered on the notion that most details of witnessed events are eventually forgotten by the witness. Misinformation could be added to the memories to render them more plausible, but, for the most part, honest testimony would contain a poverty of detail and little coherence, in contrast to the “lively and strong” assertion of Hume. This perspective also placed Dunnett in contradiction to the memory views of Kant and Locke, the latter two arguing that the power of reason would generally allow events to be recalled accurately.

Although the literature reveals a host of important contributions to the competing memory perspectives, the publication by Bartlett (1932) provided the reconstruction perspective with its most significant empirical development to that time. Rejecting outright the storehouse concept, Bartlett proposed the alternative view that remembering involves the reconstruction of an event through the use of pre-existing knowledge which eventually leads to the production of a coherent narrative (i.e., a social reconstruction). Bartlett tested his assumptions by exposing his participants to an old Indian legend called “The War of the Ghosts,” and then asked them to retell the story on multiple occasions. The retold stories were never completely accurate. Participants often mentioned events
that made intuitive sense or met their expectations of what should have occurred. As well, the memories seemed to evolve with each retelling. Bartlett used the notion of schemata to explain this pattern of findings, and concluded that people processed and recalled information in relation to their existing knowledge, beliefs, and social needs.

The question of how humans perceive their environment and retrieve the perceptions of past events is fundamental to many branches of psychology, philosophy, medicine, and the law. Although the wax cavern and storehouse concepts of memory advanced by our ancestors have now been refuted through diligent scientific study, the present day architectural and reconstructionist models of memory theory continue to explore the dynamic mysteries of false, fabricated, and real memories. In particular, memories of traumatic events have been the subject of intense debate and scientific inquiry (McNally, 2003a). After a definition of “traumatic” is offered, the predominant perspectives on trauma and memory will be briefly reviewed.

**Defining Trauma**

In order to study traumatic memory we must consider both the manner in which we conceptualize memory and how we decide what constitutes “trauma”. Research on trauma has considered a wide range of events from distressing to clinically traumatic. This variation must be considered in making any conclusions of the effect of trauma (Cordon, Pipe, Sayfan, Melinder, & Goodman, 2004; Pezdek & Taylor, 2002). For example, the degree to which an event is anticipated (e.g., death of loved one following a chronic illness or repeated abuse versus a car accident or natural disaster) may influence the manner in which it is retained in memory (e.g., Christianson & Engelberg, 1999;
Goodman, Quas, Batterman-Faunce, Riddlesburger, & Kuhn, 1994, 1997; Koss, Figueredo, Bell, Tharan, & Tromp, 1996; Steward, O'Connor, Acredolo, & Steward, 1996). While definitions of trauma are highly variable, trauma is also subjective in nature. That is, an event experienced by one person as extremely traumatic may be experienced by another person as mildly distressing (e.g., car accident).

One major problem with research on trauma and memory has been the narrow approach that some researchers have taken on defining trauma. For example, van der Kolk, Hopper, and Osterman (2001) argued that an event is traumatic only when it is sufficiently stressful to cause post-traumatic stress disorder (PTSD). Thus, individuals who experience an extremely stressful event but who do not develop PTSD may be excluded from research on traumatic memories. Strict adherence to the DSM-IV-TR (APA, 2000) criteria for defining traumatic events limits experiences to witnessing or experiencing actual or threatened death/injury or violation of physical integrity (A1), and responding in a manner characterized by helplessness, horror and intense fear (A2). Although Criterion A1 and A2 are important in defining trauma for the assessment of psychopathology, I argue that one does not need to meet these narrow criteria to have experienced “trauma”. In fact, many scholars also have argued that the DSM conceptualization of trauma fails to capture the true scope of distressing experiences (e.g., Carlson, 1997; Cordon et al., 2004; Foa, Steketee, & Rothbaum, 1989). An alternative approach is to adopt a more broad definition of trauma encompassing experiences that: (1) create intense fear or anxiety of threatened or actual harm (DSM-IV-TR; APA, 2000); (2) threaten one’s physical or mental well-being (Brewin, Dalgleish, & Joseph, 1996); (3) create uncertainty about the security of the world (Foa, Zinbarg, &
Rothbaum, 1992); (4) involve levels of stress that exceed one’s resources for coping (Hubbard, Realmuto, Northwood, & Masten, 1995; van der Kolk & Fisler, 1995); and (5) potentially disrupt the functioning of the individual (Cicchetti & Toth, 1997). For the purpose of this dissertation, and in line with previous research on trauma and memory (e.g., Foa & Rothbaum, 1998; Peace & Porter, 2004), I decided to include a broad spectrum of traumatic events based on subjective interpretation and a broad definition of trauma (Cordon et al., 2004; McNally, 2003a). This is necessary to accomplish the aims of this research in evaluating memory characteristics, consistency, symptomology and fabrications across a range of traumatic events. Further, a broad definition of trauma allows for a better evaluation of the types of truthful and fabricated events that individuals may report in everyday, clinical, and forensic settings.

Models of Traumatic Memory

In recent years, researchers have extended these philosophical views (described above) to the direct study of memories for traumatic events (McNally, 2005). Some empirical studies have provided evidence that traumatic memories are processed and retained in a different manner relative to other types of emotional experiences (e.g., Alexander et al., 2005; Porter & Peace, in press). On the other hand, other research has demonstrated marked deficiencies in memory (see McNally, 2003a). As a result, there remains considerable debate about the nature of traumatic memories. Several theoretical perspectives have been proffered to explain the way in which trauma affects memory (e.g., Shobe & Kihlstrom, 1997).
Traumatic Memory Argument.

Reminiscent of the “storehouse” view advocated by Augustine and his predecessors, proponents of the traumatic memory argument have asserted that traumatic events are associated with memory impairment (van der Kolk, 1996). According to this perspective, memories for traumatic experiences are encoded differently from non-traumatic events, and can be completely buried (i.e., repressed) or dissociated from normal recollective experience (e.g., Herman, 1992; Herman & Schatzow, 1987; Terr, 1994). Derived out of psychoanalytic traditions, one major premise is that trauma is so painful that integration of the experience into thought and memory would overwhelm available coping mechanisms (e.g., Freud, 1922). As a result, trauma may be parcelled off from conscious awareness, with only emotional and sensory fragments (if any) of the experiences remaining (Reiser, 1994; van der Kolk, 1996). For example, Janet (1889) argued that “when people experience ‘vehement emotions’ their minds may become incapable of matching their frightening experiences with their existing cognitive schemes” (as cited in van der Kolk et al., 2001, p. 24). Instead, traumatic memories are disintegrated from personal awareness due to the activation of defense mechanisms (Mollon, 1998). Psychiatry as a whole advocates this perspective, as seen by the inclusion of dissociative amnesia in the DSM-IV-TR (APA, 2000).

Trauma Equivalency Argument.

An alternate perspective to the idea that trauma impairs memory is the trauma equivalency argument. Related to the arguments of philosophers Leibnitz and Kant, recollection is an active and constructive process and subject to distorting influences of cognitive and social factors. As such, memory for trauma is subject to the same
processes. The basic tenet of the trauma equivalency argument is that traumatic events are not “special” but that they show progressive deterioration over time (similar to non-traumatic events) (e.g., Berntsen, 2001; Epstein & Bottoms, 2002; Shobe & Kihlstrom, 1997). Cordon et al. (2004) argued that trauma is processed according to the same cognitive mechanisms that influence memory for everyday or positive events. This argument suggests that while traumatic memories may be well-retained, in general, decay occurs due to ordinary forgetting and distorting influences (Laney & Loftus, 2005; Pezdek & Taylor, 2002). Accordingly, the conclusion that autobiographical memory is malleable and fallible (e.g., Loftus, 2003b) extends to traumatic memories. Proponents of the trauma equivalency view have argued that cognitive and event factors account for variations in the characteristics and retention of traumatic versus non-traumatic memories (reviewed below). As such, it is not the nature of trauma per se that influences memory but cognitive processes (Kihlstrom, 2004; Schacter, 1999).

Trauma Superiority Argument.

Contrary to the perspectives highlighted above, proponents of the trauma superiority argument suggest that traumatic memories indeed may be “special” (e.g., Alexander et al., 2005; Peace & Porter, 2004). Further to arguments by William James that emotionally evocative events could “leave a scar upon the cerebral tissues” (James, 1890/1950, p. 670), the trauma superiority argument suggests that traumatic events are remembered better than other types of emotional or autobiographical events. Even Breuer and Freud (1893) acknowledged that traumatic memories do not deteriorate relative to all other memory types in their monograph On the Nature of Hysteric Phenomena. While these authors went on to claim that trauma was securely retained in the unconscious, a
wealth of research has found that traumatic events to be "scars of memory" in that they are, quite simply, unforgettable (e.g., Porter & Peace, in press). Trauma is argued to facilitate recall according to this perspective, and leave a "lasting imprint" in memory (Grinker & Spiegel, 1945). Although this perspective may be interpreted as reverting back to early spatial views on memory, proponents of this view would acknowledge the reconstructive nature of memory. Specifically, traumatic memories may not be completely insulated from memory distortion, especially when misleading information and suggestive influences are present (e.g., Laney & Loftus, 2005; McNally, 2005; Porter et al., 1999). However, recollections of trauma can be highly reliable over time relative to memories for non-traumatic events (e.g., Alexander et al., 2005; Christianson, 1992; Porter & Peace, in press). There remains considerable debate about the "special" processing of trauma relative to other emotional and mundane events (see McNally, 2003).

Summary

Several divergent perspectives (as reviewed) exist regarding the nature of memory (in general) and the retention of traumatic events. This debate continues to remain controversial given the difficulties in conducting empirically valid research in applied populations of severely traumatized individuals (McNally, 2005). However, real-world research addressing some of these concerns has suggested that memories for traumatic events are, in general, remembered better over time than other types of experiences (e.g., Bernstein, 2002; Peace & Porter, 2004), regardless of short or lengthy delays in retention period (e.g., Merritt, Ornstein, & Spicker, 1994; Peterson & Parsons, 2005; Porter &
Peace, in press; Wagenaar & Groeneweg, 1990; Yuille & Cutshall, 1986). However, extant research remains unclear as to the characteristics of genuine traumatic experiences relative to fabricated or imagined victimization. William James (1890/1950) argued that:

"… the re-recollected past and the imaginary past may be much the same. … there is nothing unique in the object of memory … the object of memory is only an object imagined in the past … to which the emotion of belief adheres" (p. 652, as cited in Libby, 2003).

As suggested by James, it may be possible that genuine and deceptive claims are characteristically similar. If so, are there any features of memory to discern fact from fiction, and does recall change over time as a function of trauma veracity? The next two sections of this dissertation address the assessment of credibility and consistency of memory.
Chapter 3. Credibility and Memory

"Oh what a tangled web we weave, When first we practise to deceive!"


Having discussed the perspectives on remembering and the retention of trauma, I next turn to the central issue of this dissertation: how does intentional deception influence how recollections of trauma are reported? How do fabricated trauma reports change over time relative to genuine claims? Lipian, Mills, and Brantman (2004) suggested that the features of memory may reveal clues as to the veracity of a claim. This chapter reviews historical and current conceptualizations of deception and false allegations before moving on to a discussion of credibility assessment approaches. In particular, analyses of narrative content are explored as they form the basis of my dissertation research.

Throughout history, thinkers have described the manifestations of deception in society, and often offered strong views on their immorality. For example, in *De Mendacio*, St. Augustine offered a taxonomy of eight kinds of lies, all with several subtypes, and all immoral since the tongue was divinely intended to profess the truth (see Herrmann & Chaffin, 1988). Similarly, Kant (1788/1966) argued that telling the truth is the "perfect duty" of humans and is a basic necessity to the functioning of society. In his view, regardless of motivation, lies have no justification and are not defensible by any value system. Sixteenth century thinker de Montaigne saw lying as moral indiscretion and a vice to which people become enslaved (*Essays*, Chapter IX: Of Liars, 1580/1946). To the contrary, Nietzsche (1844-1900) argued that lying could be advantageous and used to protect self-interests. In *Human, All Too Human* (1879/1984), he wrote that we
refrain from lying only because of the difficulty in maintaining the lie and not due to moral obligations imposed on us; “for lies demand imagination, dissembling, and memory” (Para. 54). It is this latter view that highlights why people may lie (i.e., revenge, attention, monetary compensation) and the difficulties associated with “telling a good story”, which is the focus of this dissertation.

Defining Deception

The manner in which lies are constructed by the deceiver and presented to another person is of critical importance in this dissertation. Various definitions of lying have been offered. In an early definition, Augustine argued in On Lying (395/1956) that “a lie consists in speaking a falsehood with the intention of deceiving”. A modern definition of lying proposed by Vrij (2000) considered the speaker, receiver, and their interaction: “a successful or unsuccessful deliberate attempt, without forewarning, to create in another a belief which the communicator considers to be untrue” (p. 6). This definition acknowledges that lying goes beyond the intent to deceive and must consider the liars’ belief that the information is incorrect (Memon, Vrij, & Bull, 2003b). For the purpose of this study, lies are defined broadly and refer to verbal behaviour that is “intended to conceal, misrepresent, or distort truth or information for the purpose of misleading others” (Porter, 1998, p. 169). The types of lies examined in this dissertation are not the everyday deceit (i.e., white lies) that we all engage in on some level (e.g., “Oh, I’m doing good thanks”; DePaulo, Kashy, Kirkendol, Wyer, & Epstein, 1996; Ekman, 2001; King, 2006). Rather, I examined deceptive behaviour that has serious consequences in legal and clinical settings – false allegations of traumatic events.
False Allegations

One form of deception that has been witnessed throughout history is the false allegation of criminal or otherwise unacceptable behaviour. Such false allegations are exemplified below with reference to cases involving a single claim by one against another to more complex, wide-reaching “contagious” claims involving numerous complainants and accused persons. While by no means an encyclopaedic listing of such cases, those provided are intended to give representative examples that lead into my discussion of the modern phenomenon.

Historical False Allegations.

It appears that the social climate in fifteenth century Italy encouraged many questionable accusations of crimes or moral indiscretions. For example, in Florence, allegations were made anonymously by placing one’s written story in a tamburo (wooden box) located in front of the Palazzo Vecchio. Although such allegations typically were unfounded, they served as the primary evidence in many legal cases. In 1476, Leonardo da Vinci himself, along with four others, was accused of engaging in a homosexual affair via the tamburo. Although later cleared of the charges, many such allegations ended in convictions during this time (Nuland, 2005). One of the best-known historical examples of false allegations stemming from mass hysteria was the “witch trials” of Europe and America during 1400-1700s. Among the best-documented cases were the 1692 Salem trials in Massachusetts, which began with the “strange” behaviour of several young girls who were alleged to be under the influence of a witch. Upon pressure from town and church authorities, they made allegations against three women for engaging in witchcraft. Three allegations soon turned into over 400 accusations (Hiltunen & Doty, 2002).
Hiltunen (1996) conducted a discourse analysis of legal documents collected during the trials for 26 cases. He noted that the false claims contained vivid information and had a sense of “authenticity” in their linguistic detail. On the other hand, the accused denials of witchcraft were disbelieved based on the characteristics of their responses during interrogations (often under threat of or actual torture). Defendants were asked leading questions, and their denials were considered to be “inconsistent” – lending to the assessment that they were being deceptive (Hultinen, 1996). These perceived inconsistencies in the rebuttals of the accused women led to unabashed belief in the false allegations of four young girls and subsequent discrediting of anything said by defendants (Hiltunen & Doty, 2002). Similarly, Sjöberg (1997, 2002) discussed an interesting set of historical false allegations involving the witch panic in a small parish in Rättvik, Sweden, in the seventeenth century. In 1671, Nils Larsson alleged that he had been repeatedly abducted, taken to Blåkulla (a traditional village of folklore), and was the victim of perverted sexual practices and satanic rituals. In total, 60 children testified to mass satanic activity and convictions were obtained, resulting in a death sentence in some cases. However, oddly, the claims made by Nils Larsson himself were disbelieved. Sjöberg (2002) reports they were “too fantastical to be true” (p. 134) and that it was the features of his narrative that allowed investigators to discredit his claims. One problem with Larsson’s claims was that they were viewed as being too colourful and extensive, and contained many descriptions of interactions. Interestingly, these characteristics (i.e., abundance of detail, spontaneous reproduction, coherence, contextual imbedding) are often considered to be indicative of truthfulness in modern approaches to credibility assessment (e.g., Sjöberg, 2002; Vrij, 2005). This case demonstrates that the specific
narrative features of false claims may appear credible, but the story as a whole does not contain the "ring of truth". This appears to be one of the few cases during the widespread witch panic across Europe and America in which a thorough inquiry was conducted prior to burning people at the stake. Modern claims of mass satanic ritual abuse and child sex rings are reminiscent of the days of the witch trials (e.g., Bottoms, Shaver, & Goodman, 1996; Qin, Goodman, Bottoms, & Shaver, 1998).

*Historical Approaches to Credibility Assessment.*

There have been numerous approaches to detecting deceit throughout history. Since biblical times, methods of detecting deception have varied widely across the world (Korn, 1997). The underlying premise of ancient measures was that liars exhibited physical or physiological changes that "proved" them to be lying. For example, it was believed that individuals experienced salivary inhibition ("dry mouth") when under the stress of being deceptive (Ford, 2006). In Arabia, multiple suspects of a crime were required to lick a hot iron and the one who was most severely burned was considered guilty. In ancient China accused parties had to chew on grains of rice during questioning and spit these out to determine if they were dry (guilty) or moist (innocent). During the Inquisition, the British also employed deception detection measures involving a "trial piece" of bread and cheese that suspects were required to swallow. If particles of food remained in the mouth or throat, guilt was assumed (Troville, 1939). These ancient methods based on salivary inhibition and physiological responses are still evidenced in some parts of the world today (Ford, 2006; Kleinmuntz & Szucko, 1984).

Physical characteristics and behaviour also were important considerations throughout the history of credibility assessment. For example, an Egyptian scroll from
900 BC described signs of distress (i.e., running hands through hair, facial
discolouration), evasiveness, and rubbing one’s big toe in the sand to be indicators of
deceptive behaviour (Hocking, Bauchner, Kaminski, & Miller, 1979; Trovillo, 1939).
Detection of witchcraft during the Inquisition involved examination for “stigmata” or
physical marks on the body. In Great Britain, credibility assessments were made during
“trials by ordeal”. Suspected witches were bound at their hands and feet and thrown into
a lake of water. Based on the “sink or float” principle (i.e., immorality and impurity were
buoyant), witches would float whereas innocent suspects would sink into the water (and
often drown during the “trial”)(see Kramer & Sprenger’s *The Malleus Maleficarum* in
1486/1979). In the early 1900s, a “truth serum” consisting of a cocktail of barbiturates
(e.g., sodium pentothal, sodium amytal, scopolamine) was developed under the
assumption that barbiturates compelled individuals to provide truthful and accurate
information (Piper, 1993). Although there is little empirical evidence supporting the use
of “truth serums”, some investigative agencies continue to use barbiturates in eliciting
confessions (Bartol & Bartol, 1999, 2004).

*Modern Legal Perspectives on Allegations.*

Historically, then, allegations often were accepted as genuine based on belief in
the morality of individuals – why would any upstanding citizen make claims up,
especially when there is no gain to be received? In addition, methods for detecting
deception were ambiguous and often resulted in false convictions. Today, some scholars
tenaciously hold on to the opinion that people rarely fabricate claims of trauma or
criminal victimization (e.g., Brown, Schefflin, & Hammond, 1998; Terr, 1994). The
courts (in general) also have demonstrated a bias towards believing allegations based on
“repressed memories” (see Porter, Peace, Douglas, & Doucette, in press). However, the judiciary has acknowledged that unfounded claims of any sort should be approached with scepticism. Lord Justice Salmon in Regina v. Henry and Manning (1969) stated that with respect to allegations of rape, “Human experience has shown that in these courts girls and women do sometimes tell an entirely false story which is very easy to fabricate, but extremely difficult to refute” (as cited in Raitt & Zeedyk, 2003, p. 463). Several Canadian cases have demonstrated that judges recognized the need to assess the credibility of claims of victimization, especially in “he said, she said” cases (see Porter, Campbell, et al., 2003). For example, allegations of sexual assault were correctly identified as false in Regina v. Lukasik (1982), Regina v. Hudon (1996), and Regina v. Ambrose (2000), and the alleged victims were convicted of perjury and/or public mischief. In Regina v. Heatherington (2004a), Alberta Court Judge Caffaro cited Heatherington’s inconsistent and vague statements as indicators of her deceptive behaviour. The modern legal approach to false allegations also has recommended that judges not only look for cues to deception, but also signs of truthfulness. In Regina v. Norman (2005), the claimant’s (R.C.) allegation of sexual assault was determined to be truthful even though she had previously made a false allegation of sexual assault against a former boyfriend. In his decision, Justice De Filippis indicated that, besides physical evidence, several factors influenced his decision to accept the allegation as genuine:

“The amount of detail provided by the complainant about the assaults is a significant indicator of reliability. The assaults were not described in vague or generic terms. For example, the complainant recounted how the
defendant sprayed her with a water bottle not once but three times. She pointed out that during one of these incidents he took a drink of water and spat it at her. She recalled clearly that when she struck her nose after falling out of bed, the defendant said, sarcastically, “baby, you’re so clumsy” and pushed her head into the wall. Similarly, a lying witness need not, and would not, have pointed out that the defendant was looking for a cigarette during one of the assaults or to have testified that he asked whether she wanted her other breast bruised just before he ejaculated” (para. 28)

Modern legal approaches to false allegations support the need for deception detection methods in the courtroom. We have long since surpassed the trials by ordeal of medieval England, dry rice tests of Ancient China, and the hot iron tests of Arabia (Ford, 2006; Wrightsman & Porter, 2006). However, modern approaches to deception detection may still be fallible (Vrij, 2004). These can generally be categorized into three domains: physiological, behavioural, and verbal (e.g., DePaulo et al., 2003; Lykken, 1998; Memon et al., 2003), the latter of which is under investigation in this dissertation. Several methods of verbal or written statement analysis have been derived through empirical research (Masip, Sporer, Garrido, & Herrero, 2005). The predominant procedures are Statement Validity Assessment (SVA), which includes Criterion-Based Content Analysis (CBCA; Steller & Köhnken, 1989), and an extension of the Reality Monitoring (RM; Johnson & Raye, 1981) approach. CBCA originated from analysis of legal cases in both Germany and Sweden (e.g., Undeutsch, 1989), whereas RM has been derived from
empirical research on memory in the United States (e.g., Johnson, Foley, Suengas, & Raye, 1988). Both approaches are conceptually similar in the premise that experienced and fabricated/imagined accounts can be discriminated on the basis of their features (Schooler, Gerhard, & Loftus, 1986). The following section will review these approaches and discuss how these guided the development of the methodology and assessment measures used in this dissertation.

Statement Validity Analysis

One widely used approach to credibility assessment is Statement Validity Assessment (SVA), a practical tool designed in Germany to evaluate children's allegations of sexual abuse (Bartol & Bartol, 1999)\(^1\). Although originally created to assess children's statements, this technique also has been extended to adult populations (e.g., Köhnken, Schimossek, Aschermann, & Höfer, 1995). Further, this approach has garnered substantial empirical support for its utility in adult populations of both witnesses and suspects (e.g., Akehurst, Köhnken, & Höfer, 2001; Landry & Brigham, 1992; Porter & Yuille, 1995, 1996; Ruby & Brigham, 1998; Sporer, 1997; Vrij, Edward, & Bull, 2001a, 2001b; Vrij, Akehurst, Soukara, & Bull, 2002; Zaparniuk et al., 1995). Although formally conceptualized in the 1970s, research on statements of autobiographical events was rare until the early 1990s (see Vrij, 2005, for review).

Historical Foundations.

The roots of SVA are steeped in German testimonial traditions and research (see Mittermaier, 1834). Informed by Binet's (1900) research on suggestibility in children and

\(^1\) This approach is also known as Statement Reliability Analysis (SRA; Undeutsch, 1989).
the implications of variations of memory reporting as a function of the questions posed (e.g., leading questions), German psychologist William Stern began to conduct an abundance of research on the reliability of testimony (Stern, 1902, 1903, 1904; Yarmey, 1984). The experimental paradigm applied by Stern involved free recall followed by questions about the details of a witnessed scene (presented as a picture to the subject). Upon examination of the accuracy of memory for scenes, he reported that "perfectly correct remembrance is not the rule but the exception" (Stern, 1902, p. 327). The implications of these experiments on eyewitness testimony were evident, and research efforts became focused on forensic settings (Vrij & Akhurst, 1998). Specifically, attention was directed to cases of child sexual abuse (CSA) due to the paucity of evidence often available (other than the child's statement) and uncertainty regarding the reliability of children as eyewitnesses. In fact, Undeutsch (1989) argued that witness statements are the most important evidence in determining case outcomes. This shift in focus to the psychology of testimony and memory early in the twentieth century came to be known as the *Aussage* period (Bartol & Bartol, 1999). CSA remained the focus of much research on memory reliability and testimony as children were believed to have difficulty distinguishing reality from imagination. During the period between 1900 and 1930, German courts became increasingly sceptical about accepting CSA allegations as genuine as psychological experts were questioning the reliability of memory (Undeutsch, 1967).

While the courts remained resistant to psychological experts for many years (e.g., Wigmore, 1909), recognition that psychological research could inform legal processes pertaining to credibility burgeoned in the mid-1950s. Previously rebuked arguments made by Münsterberg in *On the Witness Stand* (1908) were now being validated. In 1954,
the Supreme Court of the Federal Republic of Germany decided that psychological or psychiatric testimony on the reliability of witnesses' accounts would be permitted in court, especially for "he said, she said" cases reliant upon memory evidence (e.g., Bartol & Bartol, 1999; Porter et al., 2003). Since this time, psychologists in Germany have testified in well over 40,000 cases involving statement credibility, and the impetus for empirical research on the subject was renewed (e.g., Undeutsch, 1982, 1984; Honts, 1994). The framework that guided research efforts on the development of a formalized system of statement analysis was primarily developed by psychologist Udo Undeutsch. The foundation rested on what is known as the Undeutsch Hypothesis (1967; see also Undeutsch, 1982, 1984, 1989). This states that memories for experienced events will differ from those that have not been experienced (e.g., imagined or fabricated reports) in both content and quality. Theoretical support has been derived for this hypothesis (e.g., Köhnken, 1989, 1996, 1999, 2002; Steller, 1989), and it forms the basis of the SVA approach (see below). Many scholars around the world have devoted considerable efforts to the study of statement analysis using this approach (see Vrij, 2005, for review).

Components of SVA.

The SVA procedure consists of three parts: (1) a structured interview, (2) criterion-based content analysis (CBCA), and (3) a validity checklist (Vrij, 2005). During the interview stage, the interviewer builds rapport with the witness and then a detailed memory report of the experience in question (e.g., sexual abuse) is elicited using free recall and direct questioning (Raskin & Esplin, 1991a, 1991b). Statements obtained during the interview(s) are compared for the presence of criteria on the CBCA argued to be indicative of "truth" (reviewed in detail below). The final aspect of the procedure is a
validity checklist, which includes evaluation of the appropriateness of language, interview characteristics (e.g., leading/suggestive), motives to report, and an evaluation of consistency of statements (e.g., Steller & Köhnken, 1989; Undeutsch, 1989; Yuille & Farr, 1987). The latter component involves a comparison of consistency between statements made across and within witnesses (Granhag & Strömwall, 2002). Thus, SVA is the overall credibility assessment procedure that includes rapport-building, the interview setting, and demeanour, whereas CBCA is more specific to addressing the detailed content characteristics of a statement (Sporer, 1997; Steller, 1989; Vrij, 2005). Inclusion of all three components allows investigators to make credibility determinations as follows: incredible, probably incredible, indeterminate, probably credible, and credible (Bradford, 1994). It has been stated that the underlying assumption of the SVA technique is that it is not sufficient to evaluate the general “truthfulness” of a person without examining the credibility of the specific allegation (Bradford, 1994). After all, reminiscent of Aesop’s fable The Boy Who Cried Wolf, generalized liars may be telling the truth and truthful people have been known to lie (Handford, 1954).

The application of content analysis techniques to credibility decisions (upon multiple retellings) is of primary focus in this study. Credibility criteria used in CBCA were developed from research and writings by legal scholars and psychological experts and evolved over more than a century (e.g., Arntzen, 1983; Leonhardt, 1930; Mittermaier, 1834; Trankell, 1972; Undeutsch, 1967). After much refinement, nineteen criteria (also referred to as “reality criteria”) were identified in the late 1980s and have been used in the systematic assessment of statements (Köhnen & Steller, 1988; Steller &

---

2 Refer to Chapter 4 in this dissertation for a detailed discussion of the consistency of memory.
Köhnen, 1989). Each criterion is argued to be present to a greater extent in truth-telling situations relative to deceptive statements (see Table 1) (e.g., Vrij, 2005; Vrij, Akehurst, Soukara, & Bull, 2004a; Zaparniuk et al., 1995). Further, an application of CBCA techniques has been identified as the core part of the SVA procedure (e.g., Vrij et al., 2002) and is heavily relied upon in credibility assessment decisions (e.g., Gumpert & Lindblad, 2001). At a basic level, scores on all criteria are summed, and high CBCA scores are considered indicative of a credible report. However, it has been emphasized that while high scores suggest truthfulness, the absence of these criterion should not invariably suggest deception (Bradford, 1994).

Researchers and clinicians applying CBCA techniques outside the context of child sexual abuse claims have found some of the criteria to be irrelevant (Vrij, 2000). Marxsen, Yuille, and Nisbet (1995) identified that all 19 CBCA criteria need not be applied (nor are applicable) to all cases. For example, criterion 10 (accurately reported details misunderstood), 17 (self-deprecation), and 18 (pardoning the perpetrator) appear to involve elements that are common in CSA claims but may not be associated with adolescent or adult reports of other types of traumatic victimization (e.g., Porter & Yuille, 1996; Zaparniuk et al., 1995). Overall, Marxsen et al. (1995) argued that the first five criteria and that two of the remaining 14 should be present in a credible statement. Thus, some have argued that CBCA assessments essentially boil down to five core criteria (details, spontaneous reproduction, amount of detail, contextual imbedding, and interactive descriptions) and the remainder do not add much more predictive validity in veracity determinations (Sporer, 1997). In general, CBCA criteria have been condensed into five main categories of assessment: (1) general characteristics (e.g., quantity of
details, logical structure); (2) specific contents (e.g., reproduction of conversation); (3) content peculiarities (e.g., unusual details); (4) motivation-related contents (e.g., pardoning perpetrator); and (5) details characteristic of the offence (based on criminological findings) (Sporer, 1997; Yuille & Farr, 1987). Some of these (such as 4 and 5) are more specific to the context of child sexual abuse cases where the criteria were originally applied and used extensively. Conversely, others argue that all criteria are critically important in the assessment of credibility and should be combined for a well-rounded statement analysis (Undeutsch, 1989; Steller & Köhnken, 1989). This is important as Köhnken (1999) has argued that CBCA criteria are theoretically related to cognitive and motivational factors. For example, Criteria 1-13 (see Table 1) are considered too difficult to successfully falsify due to the cognitive complexity of the task (e.g., providing a lot of detail consistently over time). Criteria 14-18 (see Table 1) are more likely to occur in truths for motivational reasons in which deceivers engage in impression management and appear to over-compensate with how they believe truthful individuals act (e.g., not admitting lack of memory, doubts, or correcting themselves). Accordingly, each criterion is important in assessing the veracity of a claim (Köhnken, 2002). However, a detailed evaluation of the discriminatory power of each criterion is necessary to ultimately determine what should be relied upon in SVA (Vrij, 2005).

Therefore, SVA (and CBCA in particular) is a well-known technique in credibility assessment when dealing with written statements or accounts of an event. Ultimately, “the aim of SVA is to elucidate the difference in the quality of statements” (Bradford, 1994, p. 95). Some researchers have argued that this approach is a highly useful tool for distinguishing between truthful and fabricated claims (e.g., Honts, 1994),
Table 1.
Comparison of assessment criteria for criterion-based content analysis (CBCA), reality monitoring (RM), and the memory assessment procedure (MAP)\(^3\).

<table>
<thead>
<tr>
<th>CBCA</th>
<th>RM (MCQ/JMCQ(^4))</th>
<th>MAP(^5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Coherence</td>
<td>1. Clarity/vividness</td>
<td>1. Amount of detail &amp; number of words</td>
</tr>
<tr>
<td>2. Spontaneous reproduction</td>
<td>2. Sensory information</td>
<td>2. Emotional components (self and other)</td>
</tr>
<tr>
<td>5. Descriptions of</td>
<td>5. Reconstructability</td>
<td>5. Relevancy</td>
</tr>
<tr>
<td>interactions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>conversation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Unexpected complications</td>
<td>7. Realism/plausibility</td>
<td>7. Memory perspective</td>
</tr>
<tr>
<td>10. Accurate reported details</td>
<td>10. Familiarity</td>
<td>10. Thinking about event</td>
</tr>
<tr>
<td>misunderstood</td>
<td></td>
<td></td>
</tr>
<tr>
<td>associations</td>
<td>perspective</td>
<td></td>
</tr>
</tbody>
</table>

\(^3\) See Appendix Q for description of CBCA, RM, and MAP criteria.

\(^4\) The Judgment of Memory Characteristics Questionnaire (JMCQ; Sporer & Küepper, 1995) primarily assesses the first eight broad domains. The remaining criteria were drawn from the original MCQ (Johnson et al., 1988).

\(^5\) The MAP involves both objective (items 1-6) and subjective (items 7-18) assessment components.
whereas others have reported that more research is required before any conclusive claims can be made (e.g., Ruby & Brigham, 1997, 1998; Wells & Loftus, 1991).

*Empirical Findings.*

The first scientific studies on the use of CBCA in determining veracity were conducted in the late 1980s and early 1990s (e.g., Esplin, Boychuk, & Raskin, 1988; Landry & Brigham, 1992; Raskin & Esplin, 1991a, 1991b). For example, Steller, Werttershausen, and Wolf (1988) examined memory characteristics of truthful and
fabricated events of an emotional nature (e.g., animal attacks, medical procedures, getting lost) in 88 child victims of CSA. They found that truthful events were characteristically different from deceptive reports; truthful statements contained more detail (including peripheral details), contextual imbedding, reproduction of conversation, and were more coherent (see also Vrij, 2005). Using the transcripts generated in this study, Steller (1989) conducted follow up research that demonstrated that statement veracity could be determined using CBCA criteria, especially if judges were trained in this assessment technique. Raskin and Esplin (1991a, 1991b) reported a study evaluating transcripts of confirmed \( n = 20 \) and unconfirmed \( n = 20 \) cases of CSA and found that all cases were correctly classified based on the SVA procedure. Although the 100% classification rate has not been replicated and this study was criticized for methodological issues (e.g., only one rater, unconfirmed cases may not be deceptive), it prompted a wealth of empirical research on the classification potential of SVA in general and CBCA criteria in particular.

Research since this time has demonstrated that the premise of CBCA (i.e., Undeutsch Hypothesis) is not limited to cases of sexual abuse, witnesses and victims, or children (Vrij et al., 2002). Landry and Brigham (1992) conducted a lab simulation study wherein several adult participants prepared 6 true and 6 fabricated claims and described these on a videotape after two days of preparation. Undergraduate participants then made ratings on the basis of either videotape or transcript only conditions, with respect to 14 of the CBCA criteria. The researchers found mixed results (47% overall discrimination), in that truths were generally verified at a level above chance while judges generally performed poorly at correctly identifying lies. In addition, judges appeared to be more successful when assessing credibility on the basis of videotapes (thus having access to
both verbal and non-verbal cues to deception) rather than statements alone. However, this study was criticized based on the unknown veracity of events used, the relatively short training of judges (45 minutes), and the short length of the statements (see Vrij, 2005). An important study by Zaparniuk et al. (1995) assisted in addressing this shortcoming in the empirical literature and prompted further interest in the necessity of determining reliable methods of credibility assessment. In this study, undergraduate participants were assigned to either view a videotape of a robbery or a minor theft (true condition), or hear a verbal description of the event (false condition). All participants were asked to recall the event shortly afterward, with individuals in the false condition being instructed to report their memory as if they had actually witnessed the event. Statements were then assessed using CBCA criteria and rules on which to base decisions (i.e., the presence or absence of a certain set of defined criteria). Overall, use of CBCA was able to successfully discriminate truthful from false statements in the majority of cases, with an overall rate of 78% correct identification. Specific analysis of this technique revealed that consideration of spontaneous reproduction (Criterion 2) was the most reliable in differentiating statements. However, findings of this research were limited by a small sample size, difficulties in inter-coder agreement, and using naïve coders to score CBCA criteria. Other studies have reported that the amount and richness of detail in a memory report is frequently used as an indicator of deception, such that those rich in detail are believed to be more credible (e.g., Bell & Loftus, 1989; Vrij, 2000). In fact, Vrij (2000) reported that statement analysis using CBCA has reliably found the “detail criterion” to discriminate between truthful and deceptive reports. More detailed and consistent reports are argued to arise in truthful claims due to difficulties in embellishing a fabricated report
with details generated solely from the imagination and maintaining this over time (Arntzen, 1983).

Research also has provided support for the application of CBCA to varied types of statements. For example, Houran and Porter (1998) examined the statement made by Jim Ragsdale, the only direct eyewitness testament to the alleged UFO crash in Roswell, New Mexico, in 1947. Using SVA and fact pattern analysis, they concluded that the statement was not credible due to the lack of core criteria used to assess veracity. In their study on truthful and deceptive statements made by suspects and witnesses in criminal events (i.e., thefts), Porter and Yuille (1996) found truths were rated higher in amount of detail reported, coherence, and admissions of lack of memory, and could be reliably differentiated from deceptive reports with a classification rate of 78.3%. However, it is important to note that the majority of CBCA criteria did not significantly discriminate between reports. Some researchers have reported that CBCA is more effective for evaluating witness statements than suspect statements, with sometimes only small differences in CBCA scores being noted for suspects (e.g., Porter & Yuille, 1995; Vrij et al., 2002; Zaparniuk et al., 1995). However, this finding is likely due to differences in interview style with witnesses (information gathering/open-ended) as compared to suspects (interrogative/specific-questions) (Vrij, 2000). One of the only studies to examine truthful and fabricated claims of highly traumatic (and directly experienced) events was conducted by Parker and Brown (2000). They used CBCA and validity checklists associated with the SVA procedure to determine differential characteristics in genuine (56%) versus deceptive (37%) allegations of rape (n = 43). They predicted that

---

6 The veracity of three statements (7%) could not be determined.
CBCA scores would be higher in corroborated claims of rape relative to unfounded claims. Further, validity criteria were hypothesized to be more prevalent in unfounded claims (i.e., indicative of deception). Overall, the results strongly supported these hypotheses. Criteria that showed the highest discriminative power were unstructured production (C2), quantity of detail (C3), description of interactions (C5), reproduction of conversation (C6), subjective mental state (C12), and attribution of perpetrator’s mental state (C13). The breadth of sexual acts reported in genuine statements was one discerning feature as these were not present in confirmed false allegations of rape. However, reports of severity or type of injury did not significantly differ between claims (Parker & Brown, 2000). Further, statement analysis techniques were more accurate in correctly discriminating true from false rape allegations than judgments made by police officers (Parker & Brown, 2000).

CBCA analyses have also been applied to false and fabricated memory paradigms (e.g., Erdmann, Volbert, & Böhm, 2004; Porter et al., 1999). Erdmann et al. (2004) used suggestive questioning techniques with children over the course of 6 interviews after which memories for genuine positive events and fictitious-suggested negative events were assessed using CBCA criteria. Little difference between memories was evidenced, with the exception of truthful statements, which were more detailed and logically structured. Judges generally were poor at discriminating the veracity of the reports (Erdmann et al., 2004). Although these findings are likely influenced by the adoption of the fictitious event by some children as having occurred in reality (76%), this research demonstrates the difficulties that can arise in discerning truthful from fictional accounts.
even using standardized criteria originally generated for assessing children's statements (e.g., Steller & Köhnken, 1989).

Overall, accuracy rates for the correct classification of truthful and deceptive statements have widely varied using statement analysis techniques. The lowest rate of discriminability using CBCA was 47% (Landry & Brigham, 1992) and highest have ranged between 93% (Köhnken et al., 1995) and 100% (Raskin & Esplin, 1991b). Qualitative reviews of 16 (Vrij, 2000) and 37 (Vrij, 2005) studies using CBCA techniques have yielded several conclusions. In general, CBCA criteria are met more frequently when individuals provide truthful statements as compared to deceptive accounts of an event. There also are some variations within CBCA criteria as to which are more effective predictors of truthfulness than others (e.g., amount of detail, unstructured production, coherence). In their meta-analysis on cues to deception, DePaulo et al. (2003) found significant effect sizes across all CBCA studies on criteria such as amount of detail, plausibility, logical structure, repeated details, spontaneous corrections, and admitted lack of memory. Other CBCA criteria failed to be meaningful predictors of veracity. In addition, several studies have reported that the characteristics of truthful and deceptive statements may vary as a function of direct versus indirect experience, motivation, and coaching (e.g., Köhnken, 2002; Vrij, 2005).

Factors Influencing CBCA Scores.

Although the discriminatory power of CBCA has demonstrated some potential, it also is important to recognize that a variety of factors can influence CBCA scores decreasing the utility of the tool (Köhnken, 1999). For example, Akehurst et al. (2001) conducted a study examining the role of direct participation or witnessing via video a
photography session (truthful conditions) and reading about the session and claiming they were there (fabricated condition). Overall, they noted that fabricated claims could be distinguished from truthful claims when they were directly experienced, but not when only observed. Merely having witnessed an event (on video) and reporting truthfully about it produced lower CBCA scores than directly participating in the event. The authors argued that these results have “implications for the real world when children and adults use information from television programmes and videos on which to base false allegations” (Akehurst et al., 2001, p. 65). Characteristics that reliably discriminated between truthful and fabricated claims (i.e., were more prevalent in truths) were: reproduction of conversation, amount of detail, details of interactions, descriptions of mental state, and logical consistency (Akehurst et al., 2001). Familiarity of an event is also relevant to how individuals describe events. For example, Pezdek et al. (2004) found that CBCA scores were inflated when children reported on painful medical procedures (i.e., the VCUG) that were familiar to them rather than unfamiliar. In a follow-up study, Blandon-Gitlin, Pezdek, Rogers, and Brodie (2005) had children report on real and fabricated events that were either familiar or unfamiliar (i.e., sewing a button on a shirt). Their results replicated those of the previous study – CBCA scores were higher for familiar experiences and were not influenced by the veracity of the event (Blandon-Gitlin et al., 2005). Although the target event in the latter study is far removed from traumatic victimization, the implications of the findings are important to consider when evaluating

---

7 The relative familiar condition had experienced the VCUG more than once whereas the relative unfamiliar had only had this painful procedure once. VCUG = Voiding Cystourethrogram Fluoroscopy.
statements. That is, familiarity could be evidenced by a history of past trauma – making truthful and fabricated claims less discernable via assessment of memory characteristics.

CBCA criteria also may vary according to the level of motivation of the prevaricator (e.g., Köhnken, 1996, 1999, 2002). Honest individuals are less concerned about managing the impression they are giving to other people as compared to telling the truth. Conversely, deceptive individuals engage in impression management to make themselves appear more credible. This becomes translated into deceivers behaving according to how they “think” honest people would, fitting their statements into a form of “stereotypical” report. However, research has noted that truthful individuals often provide information that is “contrary-to-stereotype” such as admitting they can’t remember aspects of the event, having a degree of uncertainty about some components, and making spontaneous corrections to their report (Ruby & Brigham, 1998). Thus, the absence of such factors can be seen when one is motivated to come across as sincere and honest when they are being deceptive. Endorsement of behaviours that are believed to be associated with truthfulness (i.e., truthfulness stereotype) often results in deceptive claims that are highly detailed and complete, with little logical inconsistency (Ruby & Brigham, 1997, 1998; Vrij et al., 2002). Research has confirmed these verbal and written differences that emerge when individuals are motivated to appear credible (e.g., Porter & Yuille, 1995, 1996; Porter et al., 1999).

Further, the role of motivation extends to seeking out knowledge of what makes a statement more “believable” to produce a more credible, but deceptive, report (Vrij et al., 2002). Past studies evaluating the influence of having information about CBCA have indicated that people can indeed be trained to give credible but deceptive statements that
remain undetected (Vrij et al., 2002; Vrij, Akehurst, Soukara, & Bull, 2004b; Vrij, Kneller, & Mann, 2000). Vrij, Kneller, et al. (2000) examined how information or coaching on CBCA criteria altered fabricated claims. Their results indicated that CBCA scores were comparable across truthful and fabricated accounts, and that an expert trained in CBCA could not accurately distinguish between truthful accounts and informed liars (Vrij, Kneller, et al., 2000). Similarly, Caso, Vrij, Mann, and DeLeo (2006) used a within subjects design with participants lying or telling the truth (counterbalanced) about an object in a backpack. Deceptive statements were associated with less detail, fewer interactions, and were unlikely to contain admissions of uncertainty or memory deficiencies when questioned by an interviewer. However, when coached on CBCA criteria, they found that participants' employed these verbal countermeasures to some extent in their deceptive claims, reporting more references to their mental state and reproductions of speech (Caso et al., 2006). It has been argued that informing people (or training/coaching) about CBCA criteria “nullifies the discriminative power of CBCA” (Vrij et al., 2002, p. 277).

**Criticisms of SVA.**

Several criticisms of SVA (and CBCA in particular) have been proffered by scholars (e.g., Ruby & Brigham, 1997). Some have argued that SVA is more of an art than an empirically validated technique (e.g., Davies, 1994). Variations arising from experiential, motivational, and coaching factors (as discussed above) also have been considered limitations of the approach (Vrij, 2005).

Apart from these, one of the primary criticisms of CBCA is the lack of a sound theoretical framework for the approach. Sporer (1997) argued that although the
Undeutsch hypothesis predicted differences with respect to genuine and deceptive claims (i.e., how they differ), there is no theoretical basis as to why these reports should differ. There also are few specifications as to when differences in statements would be observed, with the exception of the length of the statements (e.g., at least one page; Merckelbach, 2004; Raskin & Steller, 1989). However, as noted by Sporer (1997), certain criteria used to assess statements may be dependent on the type and complexity of event (e.g., Blandon-Gitlin et al., 2005). Expansion of the criteria and application to a broad scope of credibility assessment domains (other than child sexual abuse cases) requires a theoretical basis.

This also is directly related to the content of the CBCA criteria. Some have argued that SVA is biased in that it only seeks truth-verification and has no criteria specific to lying (e.g., Lucas & McKenzie, 1999). Even Undeutsch (1989) himself stated that “But a useful technique for the evaluation of eyewitness testimony ought to work in both directions: it should be as useful for the tracing of possible errors or falsehood as for the verification of a truthful and reliable account” (p. 103). Further, the absence of CBCA criteria should not be indicative of deception (e.g., Bradford, 1994). Due to unspecified cut-off scores for labelling an account as definitively truthful or fabricated, and uncertainty regarding the relative weighting of some criteria (i.e., are some more indicative of truthfulness than others?), scholars have suggested that CBCA should not be the only measure used in credibility assessments (e.g., Bradford, 1994; Davies, 1994; Pezdek et al., 2004). Deficient operational and behavioural definitions of the criteria are also considered to be problematic in the application of CBCA to the criminal justice
system as well as sometimes only moderate levels of inter-rater agreement (e.g., Anson, Golding, & Gully, 1993).

Although it has utility, many criticisms have been levelled at statement analysis based on CBCA with regard to its formulation and implementation. Therefore, caution is advised in the application of CBCA to forensic practice (e.g., Ruby & Brigham, 1997; Vrij, 2000; Vrij, 2005). CBCA was derived from practice in the courts and relies on bottom-up reasoning due to its lack of theoretical underpinnings (see Masip et al., 2005; Sporer, 1997). Conversely, the reality monitoring (RM) approach is theory driven (top-down) and its application has been extended to the domain of credibility assessment.

*Reality Monitoring*

The Reality Monitoring (RM) approach (Johnson & Raye, 1981) has been proffered as an alternative content (or narrative) analysis approach for credibility assessment (e.g., Sporer, 1997). This method also evaluates differences between truthful and false verbal responses. Not unlike the *Undeutsch Hypothesis*, proponents of the RM approach specify that qualitative differences exist between experienced versus imagined (e.g., fabricated, false) events. However, RM provides a rationale as to why genuine versus deceptive experiences differ. Numerous studies have used this approach in evaluating memories for autobiographical experiences (e.g., Alonso-Quecuty, 1992, 1996; Alonso-Quecuty, Hernandez-Fernaud, & Campos, 1997; Höfer, Akehurst, Metzger, 1996; Manzanero & Diges, 1996; Roberts, Lamb, Zale, & Randall, 1998; Vrij, Akehurst, et al., 2004a; Vrij, Edward, Roberts, & Bull, 2000; Vrij, Harden, Terry, Edward, & Bull,
2001). Only recently has attention turned towards the application of RM to forensic settings (e.g., Porter et al., 1999; Sporer, 1997; Vrij, 2005).

*Historical Foundations of RM.*

Long before modern perspectives on reality monitoring, observations of differences in “memories” for genuine and imaginary experiences were offered. Aristotle (c. 335 BC/1908 AD) believed that perceptions, or sensory stimulations, formed the basis of memory. He argued that “the act of perception stamps in, as it were, a sort of impression of the percept (on the soul), just as persons who make an impression with a seal” (p. 65). As a result, perceived (i.e., genuine) experiences made an impression upon the mind according to (primarily) the sensory characteristics of the event. Further, Plato in his work *The Republic* discussed the idea of reality and imagination being akin to a divided line in our perceptual world (360 BC/1998). He argued that we have “relatively clear knowledge of something that is more real and ‘true’ when we attend to ordinary perceptual objects like rocks and trees; by comparison, if we merely attend to their shadows and reflections, we have relatively obscure opinion of something not quite real” (Plato, Book VI). Based on these views, it could be argued that experiences of internal origin (i.e., imagined events) would not bear the same physical mark of perception, and would fail to leave a distinct “stamp” upon the soul of memory as they are but a ‘shadow’. In fact, sixteenth century philosopher David Hume (1737/1886) was one of the first to consider and elaborate the ideas of Aristotle and Plato in comparing the qualities of memory and imagination. While he endorsed a “spatial” conception of memory, Hume acknowledged that the mind was sufficiently powerful to generate images akin to
memories for perceived events. He argued that events experienced in reality left a durable mark upon memory distinct from products of the imagination:

"'Tis evident, at first sight, that the ideas of the memory are much more lively and strong than those of the imagination, and the former faculty paints its objects in more distinct colours, than any which are employed by the latter. When we remember any past event, the idea of it flows in upon the mind in a forcible manner; whereas, in the imagination, the perception is faint and languid, and cannot, without difficulty, be preserved by the mind steady and uniform for any considerable time" (Hume, 1737/1886, p. 173).

Hume adopted his belief that the ideas of memory were "lively and strong" (relative to imagination) from Berkeley, who also had argued that memories are characterized by a "steadiness, order, and coherence" that did not exist in the imagination (cited in Hume, 1739/1964, p. 138). Hume believed that memories (as reproductions of original events) were more coherent and logical than imagination (see Porter et al., 1999 for a modern validation of this view), whereas imagined events were subject to a high degree of creativity and re-arrangement (Copleston, 1964).

Collectively, these early views suggested that the qualities of memory and imagination may differ; memory reflects events as originally perceived whereas imagination is a creative reconstruction of reality that leaves a permeable stamp on the mind. Shaped by these philosophical foundations on the functioning of memory and later
development of cognitive psychology, the process by which real (perceived) and
imagined events are discriminated and potentially confused in memory formed the basis
for the reality monitoring framework (Johnson & Raye, 1981).

The Reality Monitoring Perspective.

Reality monitoring is referred to as the strategies employed by individuals to
discriminate memories from external (perceived/experienced) or internal (imagined)
Sources, that is, reality from imagination (Johnson & Raye, 1981). Conceptually, RM is
Based upon the characteristics of our memories for events, and subsequent reality
Judgments (Johnson et al., 1988). Accordingly, memories of genuinely experienced
Events should be rich in perceptual details (e.g., colours, sounds), contextual details (e.g.,
time and place), and affective information (e.g., feelings during event). Conversely,
Imagined events should contain more information about the cognitive processes involved
In creating the imagery (e.g., sensory, perceptual, or reflective processes) (Arbuthnott,
Geelen, & Kealy, 2002; Johnson, 1983; Larsson & Granhag, 2005). Cognitive operations
Are argued to occur more frequently in imagined events as a result of deliberate imaging,
Searching for pieces of memory, intention, planning, and making conclusive statements
(Johnson & Raye, 1998). As such, the origin of a “memory” generation is reflected in its
Characteristics, with internally generated (imagined) events differing from externally
generated (perceived) events (Johnson & Raye, 1981). Thus, a memory rich in spatial and
Visual details but lacking in reflective construction and cognitive operations (i.e., records
Of organization, elaboration and retrieval such as “I think I remember it being on a
Tuesday”) would be considered to be externally generated and therefore real (see also
Johnson, Hashtroudi, & Lindsay, 1993). Further, the inclusion of additional information
in memory based on knowledge/reasoning of past experiences may assist in determining that an event was perceived rather than imagined (Johnson et al., 1988).

In order to empirically examine their model, "reality" criteria were developed and measured using the original 39-item self-report Memory Characteristics Questionnaire (MCQ) (Johnson et al., 1988). This questionnaire surveyed a variety of subjective properties of autobiographical experiences assessing the perceptual, contextual, affective, and cognitive characteristics of memory (see Table 1). Factor analysis of the MCQ has yielded eight subscales of memory characteristics: clarity/vividness, sensory experiences, spatial information, time information, emotions and feelings, reconstructability of the story, realism, and cognitive operations (Sporer, 1997). This modified version of the MCQ was called the Judgment of Memory Characteristics Questionnaire (JMCQ; Sporer & Kühpper, 1995) and used to objectively assess the features of memory. Larsson and Granhag (2005) offer definitions of how some of the RM criteria may be used: "visual (e.g., "I saw a frog"), audio (e.g., "I heard that he sang a song"), sensory (e.g., "The ice cream was cold"), affective (e.g., "I was happy when we played"), spatial (e.g., "The boy put his toy on the table") and temporal information (e.g., "He sang before he left"), as well as cognitive operations (e.g., "I remember thinking he was pretty")." (p. 50).

Overall, the MCQ and subsequent modifications was used to determine the properties of memory that are relied upon in making reality monitoring decisions (e.g., Suengas & Johnson, 1988).

*Empirical Support.*

A study by Johnson et al. (1988) was one of the first to test the RM paradigm using autobiographical events and not stimuli based on words, pictures, or actions (e.g.,
Foley & Johnson, 1985). Using a within-subjects design, participants recalled perceived events such as a trip to the dentist, social occasion, or a library visit, as well as a fantasy or dream they had previously experienced. Results indicated that perceived events differed from imagined events on a variety of measures: they were associated with higher levels of sensory details (i.e., visual, sound, smell), realism, spatial arrangement of environment, and time and place details (i.e., location, setting, year, season, day, hour). In addition, real events included surrounding details of memories preceding and following the event being described (e.g., contextual embedding). Conversely, imagined events were rated as being more intense, consequential (i.e., had more implications at the time), complex, thought about more often, and less realistic. In a second study, Johnson et al. (1988) asked participants to indicate what factors influenced their knowledge or belief that the event happened in reality or was imagined. Responses were divided on the basis of characteristics about the memory itself (e.g., perceptual details), supporting memories or information (e.g., pictures of a trip), and general knowledge-based reasoning (e.g., it must be fantasy because I never went to Spain). As suspected, judgments of genuine experiences were based on details of the memory itself and the surrounding context, whereas imagined events were associated with cognitive reasoning about the likelihood of the event.

The RM model continues to be supported, in general, by recent research using self-discrimination paradigms (e.g., do you think your memory is real or you just imagined it?) (see Masip et al., 2005). However, some studies have suggested that the distinctiveness or type of event may impede reality monitoring processes. Kelly, Carroll, and Mazzoni (2002) reported that when making reality monitoring discriminations,
people relied on metamemory processes (e.g., "feelings of knowing") rather than ratios of real to imagined events presented. They also noted that distinctiveness interfered with RM judgments, such that a bias towards the real is evidenced when distinctive imagined events are presented (Kelly et al., 2002). That is, if an experience is associated with unusual or idiosyncratic features, participants would be more likely to think "that must be real because it stands out in my mind" whereas this is not associated with the qualities of real or imagined events. This is supported by literature that has reported that negative events may be more resistant to difficulties in reality monitoring relative to positive experiences (e.g., Kemp, Burt, & Sheen, 2003). Genuine and imagined experiences of negative content (relative to positive events) were easier to distinguish, possibly because they are more distinctive in nature. In a recent study, Kealy, Kuiper, and Klein (2006) had participants generate made-up and real autobiographical memories from childhood or adolescence (< 16 yrs), and subsequently rate their own memory characteristics. Participants' self-ratings of their memories indicated that real memories were more accurate, typical, and likely to be true than made-up events. These findings were noted, in particular, with ratings for pleasant relative to unpleasant events (Kealy et al., 2006). This may have resulted from a difficulty in creating plausible or typical stories of negative experiences, or an inherent belief that it is "not right" to make up stories about negative events. As a result, hesitancy to imagine negative events may have reduced self-ratings of accuracy and typicality of these events. Further, it could be argued that negative events may not be as typical as positive social experiences in that negativity is highly subjective and dependent upon the perceiver. For example, different people may experience different degrees of negativity at being embarrassed in a social setting, whereas feelings
of positivity at a 13th birthday party may be much more "typical". Finally, it is highly likely these findings were evidenced due to the internal source monitoring of participants; higher ratings of likeliness and truth were associated with real events because participants were aware these had occurred. Kealy et al. (2006) also found that real events were associated with greater visual, emotional, and contextual (time and place) details relative to made-up events. While the latter findings reinforce previous studies that have noted differences between real and imagined events (e.g., Johnson et al., 1988; Pezdek, Finger, & Hodge, 1997; Suengas & Johnson, 1988), there are several limitations to this study. First, it only focused on recall of events that were experienced prior to age 16, limiting the range of experiences recalled. Second, self-ratings of memory characteristics are useful but don't provide any measure of objective assessment of memory properties. Third, participants were not motivated to "make up" a realistic event. The design of this dissertation addressed these limitations by querying experiences of adult trauma, using subjective and objective measures, and providing instructions and some motivation to provide deceptive but believable reports (e.g., these will later be assessed by students, police officers, lawyers, and judges to try to see if they can tell the difference).

Despite limitations of some studies, reality monitoring is an alternate framework for conceptualizing differences between genuine and fabricated or imagined events. The approach was originally designed to evaluate the processes people use to discriminate between their own memories of genuine experiences versus those only imagined (Johnson & Raye, 1981; Suengas & Johnson, 1988). This framework has been extended to encompass source monitoring as well as applied to a diverse range of events (e.g., Porter et al., 1999). Further, and more important to the present research, researchers have
begun to investigate its application in assessing the reality of *other* people’s recollective experiences (e.g., Masip et al., 2005; Sporer, 1997). Imagined events are internally generated representations that may intentionally or unintentionally bear similarities with genuine memories for externally experienced events. As a result, imagined events can contain compelling “memory” details that make them appear more viable – such as distinct imagery and contextual (i.e., time and place) information (e.g., Johnson et al., 1988). This dissertation is focused on reality monitoring as applied to the assessment of others’ memories for genuine or fabricated events and is the focus of the remaining literature review on the subject. This model has been well-supported through empirical study (see Masip et al., 2005, for review).

**Reality Monitoring in Forensic Contexts**

Several studies have extended the reality monitoring approach using imagined events to apply to studies of false memories and deception (e.g., Porter et al., 1999; Sporer, 1997; Vrij, Evans, Akehurst, & Mann, 2004). Vrij, Akehurst, et al. (2004a) argued that this framework could be applied to truthful (experienced) and fabricated (imagined) events. The focus of RM in forensic contexts often doesn’t involve assessment of one’s own memories, but extending evaluations to other people’s memories of truthful and fabricated events (Barnier, Sharman, McKay, Sporer, 2005). This also has been termed interpersonal source monitoring (Johnson, Bush, & Mitchell, 1998), and may be a useful technique to address the reliability of a memory or witness statement in criminal and civil contexts (e.g., Larsson & Granhag, 2005)
Extending the RM approach to deception detection, truthful statements are those resulting from events that have been experienced whereas deceptive claims are internally generated fabrications. Therefore, truths are based on memory for events that were perceived whereas lies have no basis in reality and exist only in the imagination (Alonso-Quecuty, 1995). As such, the principles that apply to perceived and imagined memories should also be applicable to truthful and fabricated statements (e.g., variations in sensory, contextual, semantic, and cognitive processes). Some have further differentiated imagined and deceptive accounts, which could invariably influence outcomes pertaining to memory characteristics (Barnier et al., 2005). Imagined events have been associated with the creation of false memories (i.e., created but believed to be true) under conditions such as "memory work" and suggestive questioning (e.g., Hyman & Pentland, 1996; Porter et al., 1999; Read & Lindsay, 2000). The importance of the belief in deceptive events has been noted even back to the times of Nietzsche (1879/1984), who argued that a great deceiver engages in the following:

"In the actual act of deception, among all the preparations, the horror in the voice, expression, gestures, amid the striking scenery, the belief in themselves overcomes them. It is this that speaks so miraculously and convincingly to the onlookers" (Human, All Too Human; Section II: On the History of Moral Feelings, Para. 52)

However, deceptive accounts are associated with an awareness that a lie is being told, and are rarely accompanied by the belief that deceptive events happened in reality (e.g.,
Garry, Manning, Loftus, & Sherman, 1996). Raitt and Zeedyk (2003) noted that regardless of genuine, false, or intentionally fabricated, one way to discredit victims is to call into question their memory for events.

Research has demonstrated that imagined events arising from suggestion (false memories) are qualitatively different from fabricated experiences (e.g., Porter et al., 1999). Specifically, deceptive accounts were more detailed, coherent, and vivid than false (imagined) memories. It has been suggested that internal source monitoring processes allow us to discriminate both of these internally generated experiences (Barnier et al., 2005). However, comparisons of these events with memories for genuine experiences may yield differing results depending on whether a memory is solely imagined, false, or deceptive (see Porter et al., 1999). These two internal experiences may be further distinguished on the basis of motivation and perception (Ford, 2006). For example, a lie is told with the intent to relay false information (known to be false) to convince the perceiver it is true (e.g., I try to convince others that completing my dissertation required very minimal effort due to my superior intellect). Conversely, someone reporting a false memory for an event is providing false information they believe to be true without the intent to deceive (e.g., I recall laying in my crib watching our roof get torn off in the “terrible tornado” of Regina in 1980 and believed this to be true for a long time). For the purposes of this study, fabricated events were intentionally deceptive accounts generated by participants. As such, the theoretical framework of RM (suggesting characteristic differences between memories of internal versus external origin) is relevant to consider.

---

8 This false memory was later debunked by confirmation from my parents. In addition, rationale reconsideration made me realize I was too old to be still sleeping in a crib in 1980, and that I doubt my parents would have left me upstairs alone in my bedroom during a tornado while they sought security in the basement crawlspace.
However, due to the relatively recent application of RM to forensic psychology, standardized criteria and methods of applying RM do not yet exist (Masip et al., 2005; Vrij & Mann, 2004).

In one of the first studies to apply RM to deception detection, Alonso-Quecuty (1992) evaluated the characteristics of truthful and fabricated statements about a criminal event using a within-subjects design. Participants \((n = 22)\) viewed a video of a criminal act and then were required to prepare truthful and deceptive accounts of what they had witnessed either immediately following the event or with a 10 minute preparation time. The results demonstrated that when statements were generated immediately, deceptive claims evidenced more cognitive operations whereas truthful claims were richer in sensory and contextual detail (as would be predicted by RM according to Johnson & Raye, 1981). However, when a brief preparation time was permitted, this pattern of findings was reversed. Deceptive statements were rated as higher in sensory, contextual, and cognitive details relative to truthful accounts. Further, the length of the statements varied between immediate and delayed conditions; truthful were longer in the former, deceptive in the latter. The author suggested that the level of sensory and contextual detail may have been dependent on the number of words in a report (Alonso-Quecuty, 1992). That is, longer reports contain more sensory and contextual details because they are more detailed in general. Recent research has confirmed these findings. Larsson and Granhag (2005) found that RM was able to discriminate observed and imagined events on the basis of spatial, temporal, visual, and affective information (higher in observed) in children who watched a film versus imagined the film. However, the criterion of cognitive operations was not greater in imagined events as predicted by RM theory.
(Johnson & Raye, 1981). Further, observed events were more detailed overall then those internally generated. When this was taken into account in the analyses, the RM framework decreased in its discriminative power (Larsson & Granhag, 2005). In a follow-up to her original work, Alonso-Quecuty (1993, n = 68) found that sensory and contextual information was again present to a greater degree in truthful statements. This time, however, there was no reversal across preparation conditions. Preparation resulted in embellishment of the detail types in all statements relative to no preparation. Delay in reporting was also assessed in this study (Day 1/Day 2); this was associated with reduced sensory information for deceptive statements and an amplification of contextual details for both accounts.

Additional research by Alonso-Quecuty and colleagues has found mixed support for the RM approach using live staged events (versus video or audio-tapes) in child and adult witnesses (Alonso-Quecuty, 1995; Alonso-Quecuty et al., 1997; Masip et al., 2005). For example, when evaluating the influence of statement repetition, Hernandez-Fernaud and Alonso-Quecuty (1997, n = 73) found overall that across three statements, contextual and sensory information increased for both truthful and deceptive accounts but this effect was more pronounced in truthful statements. While this research has provided insight into the application of RM in making credibility judgments, there are several limitations. The manner in which RM components (e.g., sensory, contextual, internal thoughts/feelings) were operationalized was not specified, there were small sample sizes and events were not personally salient or directly experienced (e.g., video/audio conditions) (Masip et al., 2005).
A second line of RM research has been conducted by Sporer and colleagues using personally experienced and relevant events. For example, Sporer and Küpper (1995) had participants (n = 100) provide a report of an autobiographical event as well as a self-generated event, with a one-week interval between the two accounts. This study was the first to expand upon and group RM criteria into eight dimensions for assessment: sensory experiences, clarity/vividness, spatial/location details, temporal/time details, emotions/feelings, realism, reconstructability, and cognitive operations. Based on these criteria, the reported rates of correct classification were 70% for deceptive accounts and 68% for truthful statements. Ratings conducted by naïve coders relying on intuition rather than RM criteria were not significantly different from chance. Vividness and contextual details were found to be higher in truthful statements, but this effect was only evident after a time delay of one week. Overall truthful claims were associated with more temporal and realistic information and less sensory information relative to deceptive statements (Sporer & Küpper, 1995). Similar findings were reported by Sporer (1997), where classification analysis found discrimination rates of 75% accuracy for truthful events and 67.5% accuracy for deceptive statements. Biland, Py, and Rimboud (1999) found a lack of results with respect to sensory and contextual information as predicted by the RM approach, but did reveal verbal hedges and self-references to be associated with truthful (relative to deceptive) statements. Only sound and temporal details were found to be greater in truthful claims across studies conducted by Vrij, Edward, et al. (2001a, 2001b) and Granhag, Strömwall, and Olsson (2001).

Thus far, research using the RM approach has noted perceptual and contextual differences between real and imagined events (e.g., Porter et al., 1999; Vrij, Akehurst, et
al., 2004b). However, the proposed deceptive marker of increased cognitive operations has not been consistently evidenced (e.g., Höfer et al., 1996; Sporer, 1997; Vrij, Kneller, et al., 2000). Vrij, Akehurst, et al. (2004a) noted that cognitive operations may not be revealed when participants are asked to lie about factual information or events in which they were not personally involved. They found that when lying about personally relevant experiences, deceptive statements did contain more indications of cognitive operations. Comparisons of studies applying RM to deception detection also may yield contradictory results because of differences in experimental paradigms, sample sizes, operational definitions of RM criteria, scoring, and statistical analyses (Masip et al., 2005). In addition, different variables may affect variations in RM criteria, such as: (1) mode of presentation – studies that have used directly experienced events appeared to show more support for some RM predictions; (2) preparation and delay – mixed findings have been yielded with varying delays (minutes/days) with some reporting more contextual and sensory information in deceptive accounts after a delay (e.g., Alonso-Quecuty, 1992) whereas others finding more contextual information in truthful reports (e.g., Sporer, 1997); (3) repeated statements – some studies report no change in characteristics with repetition (e.g., Granhag & Strömwall, 2000b, 2001b) whereas others reported increases in sensory, contextual, and cognitive operations over time (more so for deceptive accounts; Alonso-Quecuty & Hernandez-Fernaud, 1997). For example, Granhag et al. (2001) analyzed statements made by children who watched a magician show versus those who just “pretended” they had over a series of interviews. The authors found a disproportionately high rate of correct classification for truthful (91%) and fabricated (81%) claims based on statements made in the initial interview (overall 85%). Although
the rates were still high (73% true, 85% false, 79% overall) for discrimination at the time of the second interview, it is interesting to note that there is a trend for truthful and fabricated claims to become less distinct over time (Granhag et al., 2001). Whether this operates as a function of deterioration in the memory itself, or inflation in deceptive claims with repetition that serves to consolidate details (e.g., Suengas & Johnson, 1988), or a combination of both is yet unclear.

Thus, findings using this approach have not been consistent (Masip et al., 2005). The limitation of previous studies arises, in part, from a failure to consider experiences involving directly experienced and/or personally salient events (with a few notable exceptions; e.g., Porter & Yuille, 1996; Sporer, 1997). Most studies have used mundane, neutral, or mildly negative events as the subject matter of the statements. While it is true that we may fabricate frequently in everyday life situations and interactions (e.g., DePaulo et al., 2003; Memon et al., 2003), fabrications involving highly emotionally charged and traumatic events may contain distinctly different qualities from truthful memories of such experiences. Mimicking emotion associated with traumatic events may be more difficult since we presumably experience trauma far less commonly than other events. The nature of memory for genuine claims of trauma is still hotly contested (e.g., McNally, 2003), lending to an unclear picture about the retention of deceptive traumas relative to genuine claims.

Despite these contradictions, the discriminative power of the RM approach is promising. RM techniques, in general, are useful in discriminating real from imagined events in statements made by children and adults (e.g., Larsson & Granhag, 2005; Vrij, 2000). Masip et al. (2005) noted that criteria such as sensory details (i.e., visual and
auditory), contextual and time information, and realism were good discriminators. However, it is cautioned that "there is no single piece of information that invariably marks a memory as an accurate reflection of the past" (Johnson & Raye, 1998, p. 137). As such, it is useful to examine the relative abilities of criteria used in both the RM and CBCA approaches (reviewed previously) in their ability to correctly identify liars and truth-tellers.

A Comparison of CBCA and RM Approaches

The difference between CBCA and RM approaches is that RM provides a theoretical rationale as to why truthful and generated memories differ. We experience real events through our sensory and perceptual processes, and these are argued to be reflected in memories of such events (Johnson & Raye, 1981; Sporer, 1997). Namely, real experiences are associated with more perceptual and contextual information. That is, they are rich in visual, auditory, olfactory, gustatory, and physical sensations, and include spatial details (i.e., location of the event, relationships between objects and people involved) and temporal details (i.e., order, sequencing, and duration of events). Conversely, fabricated events are conjured from the imagination and are more likely to contain cognitive representations of the event, such as thoughts or rationales (Johnson et al., 1993; Johnson & Raye, 1998; Johnson & Suengas, 1989). A further difference between CBCA and RM is that RM offers criteria for assessing not only truths but also falsehoods (Vrij, Akehurst, et al., 2004b).

Many studies in both forensic and non-forensic contexts have applied CBCA and RM criteria in credibility assessment protocols in order to compare the relative
effectiveness of approaches (e.g., Vrij, 2005). Some research has garnered support for CBCA techniques demonstrating heightened efficacy relative to RM. For example, Santtila, Alkiora, Ekholm, and Niemi (1999) examined truthful and deceptive statements of negative events (i.e., animal attack, injection) in children and used logistic regression analysis to determine rates of discriminability of CBCA and RM criteria. Overall, CBCA was able to distinguish truthful (69.1%) and deceptive (63.8%) accounts better than RM criteria (61.7% and 66.2%, respectively). A study by Porter and Yuille (1996) evaluated statements of perpetrators and innocent suspects in a staged “crime” and found that only a few criteria (i.e., detail, lack of memory, logical structure) derived from the CBCA approach demonstrated utility in differentiating truthful and deceptive statements.

Conversely, Granhag, Strömwall, and Landstrom (2006) compared CBCA and RM memory assessment procedures across real and imagined events as a function of the number of recalls. Children experienced an interaction with a stranger or only imagined doing so (i.e., a man showing them a masquerade mask in his vehicle), and then were questioned either one or four times over a 2 week period. They found that repeated recall was associated with more detailed reports relative to one recall session, although events lessened in unstructured production (i.e., became more “script-like”) and affective details over time. Truthful reports were associated with conversational details, lack of memory, admitting doubts, sensory and affective details, and greater realism. Conversely, fabricated stories contained more unstructured production and unusual details. Similarly, Strömwall, Bengtsson, Leander, and Granhag (2004) studied the effects of repeated experience on memory for real and imagined events in children. After experiencing either one or four health examinations, statements about the events were elicited and later coded
using CBCA and RM. Their results are similar to those found with respect to repeated recall (Granhag et al., 2006). Overall, RM served to better discriminate between truthful and imagined events, even after repeated recalls of an experience. Vrij, Akehurst, et al. (2004b) found that both CBCA and RM techniques successfully discriminated between real and deceptive statements, such that genuine experiences are associated with higher scores on both measures. They also found that cognitive operations (indicative of deception in the RM framework) are more prevalent when individuals are lying about a personally experienced event. Further, RM was better at identifying liars versus truth-tellers relative to CBCA. Finally, self-monitoring or impression management was positively associated with CBCA scores. That is, people who engage in more socially desirable behaviours are better at successfully escaping detection when they are lying (Vrij, Akehurst, et al., 2004b).

In addition, use of RM in tandem with CBCA has produced noteworthy correct classification scores to date: 82.5% truthful, 75% fabricated (see Masip et al., 2005, and Vrij, 2005, for review). Similar attempts to classify truthful and fabricated claims according to either CBCA or RM criteria have yielded lower estimates of correct classification. For example, Vrij, Edward, et al. (2000) reported a rate of 64.1% for fabricated and 70.6% for genuine events based on several RM sources of information used to differentiate reports (i.e., visual, sound, spatial, temporal, affective, cognitive). However, use of a combination of these approaches has revealed promising results. Vrij, Akehurst, et al. (2004a) studied the effectiveness of CBCA and RM in discriminating truthful and fabricated statements. The methodology used in this study was similar to that reported in their previous research (e.g., Vrij et al., 2002; Vrij, Akehurst, et al., 2004b).
Participants played a game with the experimenter and were interrupted by a confederate who came in and wiped some important information off the blackboard. Participants were later interviewed and instructed to provide truthful or deceptive accounts. Further, some participants in the fabrication condition were given information on CBCA criteria and advised to incorporate these characteristics into their story about the event. Their results indicated that CBCA and RM were both able to discriminate truthful and deceptive claims based on their characteristics, with RM performing slightly better. In addition, knowledge of CBCA criteria lessened the discriminative power of the technique – truthful and deceptive claims were much more similar with coaching (Vrij, Akehurst, et al., 2004a).

Collectively, it appears as though both CBCA and RM criteria have demonstrated utility in the detection of deception in statements. However, inconsistent results limit their application to forensic settings and more research is necessary to determine the contribution of each in credibility assessment (e.g., Masip et al., 2005). Other studies have combined criteria from both CBCA and RM approaches to create a more global assessment of memory characteristics. The present dissertation employed an approach (discussed below) based on the theoretical and empirical foundations of both criteria-based content analysis (Steller & Köhnken, 1989) and reality monitoring (Johnson & Raye, 1981).

*Holistic Approach to Memory Assessment*

Previous studies have used several methods for assessing the verbal content of statements, involving components such as consistency, logic, and analysis of details (e.g.,
Newman, Pennebaker, Berry, & Richards, 2003; Sapir, 1987; Shuy, 1998; Steller & Köhnken, 1989; Vrij, Edward, et al., 2000). In general, approximately 70% of truths and lies can be distinguished on the basis of verbal and content analysis measures such as CBCA, RM and other approaches (Vrij, Evans, et al., 2004). The two primary techniques reviewed above have become the focus of efforts to create a combined approach to assessing the characteristics of truthful and fabricated claims.

For example, several scholars have applied a narrative structure approach to memory assessment that is reminiscent of CBCA and RM criteria (e.g., Canter, Grieve, Nicol, & Benneworth, 2003). Labov (1972) provided a detailed structure of narrative stories that is particularly relevant to accounts of autobiographical events. In general, he suggested that narratives are composed of six components: (1) abstract – initial summary of story; (2) orientation – contextual details including time, place, people, actions; (3) complicating action – the central or main part of the story and associated behavioural reactions; (4) evaluation – assessment of the importance of event and attitude; (5) outcome of the complicating action; and (6) coda – linking the narrative to the present time or situation (Labov, 1972). Stein and Glenn (1979) elaborated on this model dividing the components of narratives into setting and episodic facets. The setting facet involved the following: (1) abstract – summary of preceding actions and tone setting, and (2) setting – details about the physical, social, and time context of the event. With respect to the episode, it was argued that the typical narrative contains descriptions of the: (3) initiating event – key event of the narrative or disruption; (4) attempt – actions resulting from the initiating event; (5) consequence – the results of the attempt; and (6) reaction – evaluation of feelings about the event and what the likely consequences are (Stein &
Glenn, 1979). So why is this important? Plausibility judgments are frequently based on whether memories follow this type of narrative structure or contain the components listed above (Canter et al., 2003). In turn, assessments of plausibility directly map on to evaluation of the credibility of a claim. Simply put, when events follow these suggested sequences, they are likely to be perceived as genuine, and when they deviate they are more likely to be judged as dishonest. Using the narrative structure model (Stein & Glenn, 1979), Canter et al. (2003) manipulated statements made by criminal suspects to adhere to the narrative structure or not (e.g., sequence of events disordered), and also whether a criminal anchor (e.g., once a thief, always a thief) was present in reports. Their pilot study results indicate that claims following the “typical” narrative structure were considered more coherent and plausible relative to unstructured statements. Further, statements without a criminal anchor were deemed more believable. In a follow-up study, the authors of this research created a measure of plausibility including perceptions such as how coherent, consistent, credible, and logical the report was. Using the same stimulus materials as before, they found only a main effect of anchoring – the presence of a criminal anchor reduced the perceived plausibility of reports. Interestingly, no effects were found for structured versus unstructured narratives and plausibility, although there was evidence of a trend for unstructured accounts to be considered more plausible. This effect was accounted for when the separate effects of statements of burglaries versus homicides were factored out, such that the results were in the anticipated direction (Canter et al., 2003). Of most importance to the present study is the content of these sequences that draw upon more formalized criteria used in CBCA and RM assessments (e.g., sensory and contextual details, logical consistency, reproduction of conversation,
temporal order, affective information, to name a few). Pezdek and Taylor (2000) furthered these ideas and provided a theoretical basis for differences between truthful and fabricated claims. They proposed that claims differ in the amount of episodic and schematic descriptions provided. In this case, fabricated stories would be more associated with "story-like" gist information, where claims fit a general schema for "a negative or traumatic event". Although truthful claims would also consist of schematic descriptions of events (e.g., a "script" memory for robbery where a gun is pulled and fired at the ceiling to get the attention of everyone in the vicinity), these are argued to also contain more specific details related to the incident itself (episodic details). This idea is similar to CBCA and RM approaches, and factors into the current conceptualization of plausibility and memory characteristics as assessed in this study. Although tests of the narrative structure model yielded mixed findings, this approach can inform the development of assessment tools measuring statement plausibility and veracity.

Closely related to the present study, recent research conducted by Barnier et al. (2005) used the Memory Characteristics Questionnaire (MCQ; Johnson et al., 1988) and the Aberdeen Report Judgment Scales (ARJS; Sporer, 2004) to investigate memory differences between real, imagined, and deceptive reports. The ARJS is a combination of CBCA and RM criteria into a more global measure of memory, with assessment of 13 scales of information such as realism/coherence and time information. Participants in their study generated truthful, imagined, and deceptive accounts of autobiographical childhood events that were either positive or negative in nature. Imagined reports were statements that were made up or fabricated but only imagined, whereas deceptive reports were fabricated claims made up with the intention of convincing someone else the
experience was real. The authors predicted that fabricated reports would not differ substantially from genuine claims due to the motivation inherent in deceptive claims that they be believable. Conversely, imagined events may be subject to an “inflation” of details and thus appear more realistic (e.g., Pezdek & Taylor, 2000). Using the MCQ, Barnier et al. (2005) found: (1) truthful reports were rated as more clear/vivid than deceptive claims; (2) deceptive accounts contained fewer thoughts/feelings relative to truthful and imagined memories; and (3) negative memories were richer in sensory and contextual detail compared to positive memories. Assessments based on the ARJS criteria revealed: (1) memory processes were more prevalent in genuine relative to deceptive accounts; (2) deceptive accounts contained less emotion and thought details relative to imagined events; and (3) higher personal significance, thoughts, and spatial details were reported in negative memories relative to positive events. These results are consistent with previous research that has found qualitative differences between memory reports (e.g., Porter, Peace, & Emmett, in press; Porter et al., 1999; Sporer, 1997). In addition, Barnier et al. (2005) noted that when accounts were rated by human coders, judges performed better at correctly classifying deceptive claims (73%) relative to imagined (61%) and genuine events (52%). However, negative memories appeared more difficult to correctly assess relative to positive experiences. This likely resulted from a positive response bias; people are hesitant to judge a negative event as being “made up” (see Peace, Porter, & Almon, 2006). Barnier et al. (2005) interpreted their findings as supporting an imagination inflation framework. Imagined events were more characteristically similar to genuine experiences; therefore the process of imagining enhanced the sensory and perceptual details making them appear more real. Barnier et al.
(2005) refer to additional research (in preparation) they have conducted on truthful and fabricated emotional, unemotional, and traumatic memories – where deceptive accounts may be convincingly real based on the qualitative characteristics of memory. Use of the ARJS shows some promise but future research is necessary to discern its discriminative value when used in conjunction with the MCQ. The measures by Barnier et al. (2005) are similar to the approach taken in the present study. However, the MCQ and ARJS still rely primarily on subjective memory ratings made by participants for discriminative analyses.

Another issue with respect to past techniques has been the lack of (or inconsistent) criteria relied upon for making determinations of “truth” (Parker & Brown, 2000). Some studies have used few criteria whereas others have used an abundance of measures. For example, allegations of rape were assessed by MacLean (1979) on the basis of four criteria: were the components of the claim internally consistent; did the claim have “the ring of truth”; did police and medical reports substantiate the claim; and did the victims admit their allegations were false. McDowell (1992) measured 57 items in the attempt to develop a credibility assessment tool specific to sexual assault. These were heavily weighted on two factors: the timing of the report (immediate versus delayed) and the ability of the victim to describe specific details of sexual victimization. However, this measure was limited in its utility as “there was little conceptual articulation underpinning the choice or weighting of criteria or any principles governing cut-off criteria” (Parker & Brown, 2000, p. 239). There has also been little standardization in assessment measures and specific guidelines on what “validates” or “discredits” a statement. Researchers not only have debated the number of criteria that should be assessed (e.g., Ruby & Brigham, 1997; Zaparniuk et al., 1995), but also what the content of these should be (e.g., Horowitz
et al., 1997; Porter et al., 1999). Using CBCA techniques, Yuille (1988) suggested a statement could be verified if the core five criteria, as well as two others, were present. Alternately, Raskin and Esplin (1991a) argued that logical consistency (C1) and quantity of detail (C3) are the two essential criteria to be considered in determinations of veracity. Very few studies have applied cut-off levels indicating degrees of truthfulness (e.g., Jones, 1994; Parker & Brown, 2000; Zaparniuk et al., 1995).

Despite the wealth of research on statement analysis techniques, few well-rounded and thorough approaches to assessing memory characteristics have been derived. The assessment procedure used in the present study was developed by Porter et al. (1999) and modified for use in this dissertation with consideration of the need to balance reliability and utility of the measure (Woolnough & MacLeod, 2001).

*Memory Assessment Procedure.*

The Memory Assessment Procedure (or MAP) was devised by Porter et al. (1999) based on a combination of reality monitoring and statement analysis criteria. Using elements of the Memory Characteristics Questionnaire (MCQ; Johnson et al., 1988), as well as from Criteria-Based Content Analysis (CBCA; Steller & Köhnken, 1989), the MAP provides a global assessment of memory characteristics. This includes both phenomenological characteristics (subjectively rated by participants) and qualitative characteristics (objectively assessed by trained coders) (see Table 1). The original procedure involved 12 factors rated subjectively (5) or objectively (7), including: vividness, confidence, sensory components, amount of detail, repeated details, and coherence (see Porter, 1998, and Porter et al., 1999, for a detailed discussion of memory criteria). Testing an architectural (reconstructive) model of remembering, participants (n
were asked to recall several negative childhood experiences in a series of interviews involving suggestion and guided imagery techniques. Participants reported a genuine childhood event, were repeatedly interviewed about a suggested event that had not been experienced (as confirmed by the participants’ parents), and a fabricated event. One of the primary findings of this research was that 56% of participants developed partial or complete false memories for experiences that had never occurred after engaging in “memory work”. More relevant to the present study, Porter et al. (1999) also noted differences in the characteristics of memories for real, false (created/implanted), and fabricated experiences using the MAP. Specifically, reports of genuine experiences were rated as more detailed, moderately coherent, vivid, clear, and were associated with participants admitting gaps in their memories. False memories were rated as coherent, but were less vivid and held with less confidence. Fabricated reports were associated with an exaggerated ‘over-the-top’ quality: they were associated with high stress ratings, many repeated details, high levels of vividness and clarity, but very few reports of lapses in ‘memory’. In addition, while participants reported “participant” perspectives in both fabricated and implanted memory reports (i.e., the memory was from their own point of view), real traumatic experiences were associated with an “observer” perspective in memory (i.e., they reported seeing themselves in the memories)(Porter et al., 1999).

Overall, these findings suggest that memory characteristics for truthful, falsely recalled and deceptive events are characteristically different and the MAP is sensitive to these variations.

This method of assessing memory has been used in additional research evaluating the characteristics of memories for traumatic (e.g., sexual and non-sexual traumas,
homicide and violent crimes) and positive emotional experiences (e.g., Peace & Porter, 2004; Peace & Porter, 2006; Porter & Peace, in press; Porter, Woodworth, Doucette, & Peace, 2006). Most directly related to my dissertation is a pilot study conducted by Porter, Peace, and Emmett (in press). For the purposes of that study, I modified the MAP to include an assessment of plausibility (objective) and perceived credibility (subjective) items, based on narrative structural models, statement analysis research, and its relevance to the legal system. Previous research has found plausibility to be related to perceptions of credibility, and the accurate detection of deception (e.g., Hiltunen, 1996; Kraut, 1978; Landström, Granhag, & Hartwig, 2005; Strömwall & Granhag, 2003a). In our pilot study, participants were asked to generate both a truthful and fabricated report of a negative or traumatic emotional experience (within their level of comfort). Results indicated that fabricated memory reports (compared to truthful reports) were associated with higher levels of subjective emotionality (e.g., emotional intensity, feeling detached), fewer time and place details, shorter reports, greater endorsement of post-traumatic stress disorder (PTSD) symptoms, and higher scores on measures of traumatic stress (Porter, Peace, & Emmett, in press). Memory analysis using the MAP revealed characteristic differences between claims.

One purpose of this dissertation was to further extend this approach to evaluate the ability of the MAP to successfully discriminate between truthful and deceptive claims of trauma⁹. Further, I applied this approach to assess the changes in memory characteristics over time (see next chapter). No studies have examined the characteristics and consistency of real and fabricated traumatic memory reports over an extended period

⁹ The MAP was also used in conjunction with measures of trauma symptomology to evaluate factors that assist in judging the veracity of a claim.
of time. Recent studies have demonstrated that traumatic memories remain relatively consistent over time as compared to positive emotional events (e.g., Peace & Porter, 2004; Porter & Peace, in press), but little is known about how fabricated events are maintained in memory relative to negative memories. As stated by Vrij and Mann (2001), it is easier to tell the truth than fabricate a convincing and plausible lie that must be maintained over time.

Empirical evaluation of these issues is of critical importance in forensic contexts, where legal professionals are continually faced with the challenge of determining the credibility of reports of victimization, at many different levels within the criminal justice system and across varied durations.

Summary

Much research using the RM approach has looked at differences in memory based on self-reported ratings of characteristics (e.g., Johnson & Raye, 1981; Johnson et al., 1988; Suengas & Johnson, 1988). Alternatively, past research on credibility assessment (e.g., SVA and CBCA) has relied on objective measures to assess the veracity of a claim (e.g., Merckelbach, 2004; Vrij, 2005). The aim of the present dissertation was to combine approaches to studying memory and credibility by evaluating the characteristics of truthful and fabricated claims using both subjective and objective measures of assessment (MAP; Porter et al., 1999). Based on this approach, features of a verbal or written statement can be used to differentiate between truthful and fabricated “tales of trauma”. Comparison of truthful and fabricated claims using a within-subjects design allows for an assessment of how we would realistically provide deceptive reports. It is unlikely that we
generate fabricated experiences without any basis in our own personally experienced realities. This is a naturalistic experience, such that Loftus (2003b) calls memory a “synthesis of experiences”.

Related to the examination of characteristics present in truthful and deceptive statements is how they are retained in memory over time. Witnesses to a crime often provide multiple statements about the event and are re-interviewed on several occasions prior to testimony. Only recently has research on credibility assessment begun to address the effect that multiple interviews or recall attempts has on memory for genuine and imagined or fabricated experiences (e.g., Granhag & Strömwall, 2001b; Granhag & Strömwall, 2002; Granhag et al., 2006). Brewer, Potter, Fisher, Bond, and Luszcz (1999) reported that inconsistencies in statements made by witnesses were perceived as being indicative of memory inaccuracy and potential fabrications, thus less likely to be believed by potential jurors. Thus, the concept of “truth” in the criminal justice system is shaped by our perceptions and the indicators that we rely on to make assessments of credibility (Brewer et al., 1999). Consistency between testimony in a criminal trial and previous statements appears to be one of the prime indicators of whether eyewitness memory is accurate (Fisher & Cutler, 1995). In general, consistency is used as a cue to justify decisions that a person is being deceptive but also a rationale for why someone is judged to be truthful (Granhag & Strömwall, 2001b, 2001c). As a result, it is highly relevant to examine consistency of both truthful and fabricated claims of trauma. Research on consistency and memory retention is the focus of the next chapter.
Chapter 4. Consistency and Memory

Montaigne: “He who is not sure of his memory should not undertake the trade of lying.” (Essays, Of Liars)

There is much wisdom in the old adage “The wheels of justice grind slowly.” Typically, in the Canadian justice system, there are lengthy delays between a complainant’s initial allegation of a crime and his/her testimony in the court (e.g., Euale & Turtle, 1999; Porter, Campbell, et al., 2003). How might a report of a false traumatic victimization change over time relative to a report based on a true experience? How well can a complainant making a false allegation maintain this story over months or years in an attempt to appear credible? In making predictions about the consistency of true and fabricated reports over time, it must be kept in mind that they are based on very different phenomenological experiences. On one hand, a true report is based on a traumatic memory or a re-experiencing of the sensory and emotional aspects of an event typically requiring no conscious attempt at “getting the story straight.” On the other, false reports are based on cognitive inventions, or lies, with which the liar must actively maintain a coherent narrative. In the latter case, “getting the story straight” is imperative. In the following sections, I will consider each of these individually before making predictions of the relative consistency of true and fabricated reports of trauma over time. Importantly, this issue has not been addressed in previous research.

The manner in which the passage of time influences memory is one of the oldest questions of psychological study (see Roediger, 1996, and Schacter, 1995, for historical reviews). Interest in the nature of memory in forensic contexts also emerged early in the history of psychology (e.g., see Yuille et al., 1995), in recognition of its tremendous
consequences. Before exploring the research in this area, it is important to differentiate the accuracy/reliability and consistency of memory over time, as these are sometimes used interchangeably in the literature. The accuracy of a memory reflects the level of similarity between the factual details of an event and one’s recollection for it. While it is usually not possible to measure accuracy in forensic settings because of the absence of a recording of the event (ground truth), it can be estimated by comparing the details that most likely occurred (e.g., from police reports) with those reported by victims or witnesses (e.g., Orbach & Lamb, 1999). For example, Schelach and Nachson (2001) estimated the accuracy level of the memories of survivors of the Auschwitz concentration camp by comparing the details they provided with well-documented information. In general, they estimated that accuracy was approximately 60% after fifty years and was enhanced for emotional events relative to neutral or routine experiences. Consistency, on the other hand refers to the level of concordance between the details of a memory report given on two or more retellings, regardless of the actual reliability of the memory. In general, it would be expected that the major details of highly accurate memories would remain generally consistent over time. On the other hand, highly consistent memory reports might not be at all reliable. For example, in therapy a person could develop a false memory for a childhood event that might be persistent and consistent over many years but whose accuracy is 0%. Additionally, in theory, someone could intentionally fabricate an event (e.g., being attacked in the park), memorize the details, and produce a highly consistent memory report over multiple retellings. The consistency might be 95% while the accuracy of the report is 0%. Thus, we might say that a reasonable amount of consistency is necessary but not sufficient for a reliable, accurate report, while reliability
is neither necessary or sufficient for a consistent report. Unfortunately, the two are often not distinguished in the literature (Woolnough & MacLeod, 2001). For the purpose of this dissertation, traumatic memory reports were compared strictly in terms of consistency across three recall sessions, as there was no basis (other than the participants' self-reported information) for establishing accuracy. In addition, the design of this research incorporated the use of fabricated memory reports. It is impossible to establish the “accuracy” of a deceptive claim as they have no origin in reality.

Consistency and Credibility

Arguably, the perceived credibility of a memory is more dependent on the level of consistency in the reports than the actual accuracy of the report (e.g., Magner, 2000; Talarico & Rubin, 2003). For example, a witness or victim to a crime may report additional details upon further interviewing and omit details previously mentioned (reducing the objective consistency of consecutive statements), but these details may not be inaccurate in comparison with the ground truth of the event (e.g., Brewer et al., 1999). In their research on a robbery/homicide, Yuille and Cutshall (1986) found that over 60% of details from memories elicited during research interviews represented newly mentioned (but accurate) material when compared to police interviews. That is, witnesses evidenced variations in the reports that they made about the events across two time periods; with 81.56% of details that were provided during both interviews reported consistently (whether right, wrong, or unclassifiable). Magner (2000) argued that similarity between prior statements about an event reflects consistency in reporting, but does little to substantiate the evidence of a case. This distinction is particularly relevant in
applied and legal settings where access to the objective facts of a traumatic experience often is limited\(^\text{10}\).

Research in forensic and experimental contexts has confirmed a heavy reliance on statement consistency as a determinant of credibility (e.g., Brewer & Burke, 2002; Granhag & Strömwall, 1999; Granhag & Strömwall, 2000a; Granhag & Strömwall, 2004; Seelau & Wells, 1995). For example, Greuel (1992) asked police officers to indicate what cues they believed were associated with deception, specifically focusing on cases of sexual assault. Results indicated that statement inconsistency was the highest rated cue associated with deception. Research manipulating the consistency of statements in mock-juror scenarios has repeatedly demonstrated that exposure to inconsistent statements leads to greater perceptions of inaccuracy, lowered confidence in the witnesses testimony, and drastic reductions in the perceived credibility of the witness (e.g., Berman & Cutler, 1996; Berman, Narby, & Cutler, 1995; Goodman, Batterman-Faunce, Schaaf, & Kenney, 2002; Lieppe & Romanczyk, 1989; Semmler & Brewer, 2002). For example, Brewer et al. (1999, Study 1) studied potential jurors’ reactions to eyewitness inconsistencies, nervousness, ignored questions, gaze aversion, levels of confidence, exaggerations, coherence of report, and added details. All participants rated the extent to which each of these items was suggestive of inaccuracy. While their findings indicated that all such behaviours led to juror suspicion about the accuracy of the account, inconsistency between statements was considered the strongest indicator of inaccuracy (Brewer et al.,

---

\(^{10}\) Establishing the ground truth of traumatic experiences is further complicated by the fact that victimization is often a “private” experience, and may not have been witnessed by others or reported (e.g., sexual assault, domestic violence).
1999). In fact, witnesses are often discredited on the basis of inconsistencies (e.g., Glissan, 1991; Salhani, 1991; Stuesser, 1993; Wellman, 1986).

An additional consideration in the perception of credibility is how individuals determine what is consistent or not. Although consistency serves as a “cue to deception”, it is unclear how this is assessed. For example, the police officers in Greuel’s (1992) study all saw the same testimony in a rape case. However, half of the officers rated the statements as consistent whereas half found them to be inconsistent (see also Granhag & Strömwall, 2000b, 2001a). This may result from variations in the way testimonial consistency is construed (e.g., contradictions, omissions, added details) (Brewer & Burke, 2002). Despite variations in the way individuals perceive the degree of consistency, consistent statements are viewed as credible whereas inconsistent statements are discredited – known as the consistency heuristic (Granhag & Strömwall, 2000a). In fact, these researchers consider our reliance on the consistency heuristic alarming given the lack of empirical research addressing the consistency of truthful and deceptive statements (Granhag & Strömwall, 2001b).

Although forensic psychologists have played an active role in educating the legal system on memory and testimonial biases (e.g., Wells et al., 1998), memory continues to be poorly understood in this context. Judges and juries often hold beliefs that memories of traumatic events will be accurately and consistently retained over time with little variation (e.g., Granhag & Strömwall, 2001b). In turn, if victims or witnesses (or perpetrators) alter their accounts of an incident, their memory is perceived to be unreliable and their testimony discounted. Strategies for legal cross-examination are often focused on statement consistency and are well-documented in legal texts (e.g., Salhani,
1991; Stuesser, 1993). For example, Glissan (1991) stated that an inconsistent report can destroy the credibility of a witness and advised lawyers to seek out such alterations in statements.

Procedures (e.g., SVA) employed by legal professionals in some jurisdictions have recognized that consistency between statements and between witnesses must be considered when assessing the credibility of a statement (e.g., Granhag & Strömwall, 2002; Steller & Köhnken, 1989; Undeutsch, 1989). Unfortunately, extant research on consistency has failed to provide “consistent” conclusions regarding memory retention over time. Further, empirical studies of consistency and credibility are lacking.

*Consistency of Memories*

Studies on memory consistency have generally focused on several areas of investigation, such as recollections of childhood experiences or flashbulb and personally traumatic events in adulthood. These are briefly reviewed to elucidate the different perspectives on traumatic memory consistency.

*Children’s Long-Term Recall.*

Much research on consistency has focused on children’s recall for stressful or traumatic events (Cordon et al., 2004; Peterson, 2002; Pezdek & Taylor, 2002). In general, results support the notion that children’s recollections of trauma can be highly reliable over time following the approximately two years of infantile amnesia (e.g., Howe, Courage, & Peterson, 1996; Peterson & Rideout, 1998) and when interviewed in non-misleading contexts (e.g., Goodman et al., 2003; Jones & Pipe, 2002; Peterson, 1996; Peterson & Parsons, 2005; Pillemer, Picariello, & Pruett, 1994; Pipe et al., 1997;
Quas et al., 1999; Terr, 1979, 1983). For example, Peterson and Parsons (2005) found children's recollections of serious injuries were detailed and mostly accurate for children who were 2-years old at the time of the event (77.7% accurate) relative to memory distortions (i.e., intrusions) evident in recollections reported by 1-year old children (40.7% accurate). In another study, Peterson and Whalen (2001) examined children's recollections of a traumatic childhood injury (and corresponding treatment in a hospital emergency room) immediately following the event and five years later. They found a remarkably high level of memory accuracy after five years; more so for central (84.45%) and peripheral (81.77%) details of the injury relative to recollections of the hospital treatment (77.06% and 69.97%, respectively). Therefore, although child participants can exhibit generally consistent recollections of trauma over the long term, memories were subject to some distortion over time depending on their age and the content (or personal salience) of the event (Peterson & Whalen, 2001). In addition, researchers have found that reports of trauma by young children may become more detailed over long periods (e.g., Fivush, McDermott-Sales, Goldberg, Bahrick, & Parker, 2004; Peterson & Whalen, 2001). Several explanations have been proffered for this increase in detail, ranging from improvements in narrative/language skill, rehearsal, knowledge, and incorporation of other source information into memory (Cordon et al., 2004; Peterson & Parsons, 2005). These factors also may influence recall of traumatic memories over extended durations in adult samples. Although such studies address long-term retention, they rely upon child samples and often assess accuracy rather than direct measures of consistency.
*Consistency in Adults.*

Research involving the consistency of memory in adult populations has been conducted (to a large extent) using “flashbulb” events, such as natural disasters, tragedies, or other shocking and unexpected events of a public nature. Studies on flashbulb memories have examined events such as the Challenger explosion (e.g., Bohannon, 1988; Neisser & Harsch, 1992; Terr et al., 1996), the Kennedy assassination (e.g., Brown & Kulik, 1977; Schacter, 1995), the O.J. Simpson murder trial verdict (e.g., Schmolck, Buffalo, & Squire, 2000; Winningham, Hyman, & Dinnel, 2000), and the terrorist attacks of September 11, 2001 (e.g., Altmann, 2003; Kvavilashvili, Mirani, Schlagman, & Kornbrot, 2003; Lee & Brown, 2003; Pezdek, 2003; Schmidt, 2004; Smith, Bibi, & Sheard, 2003).

According to the original conceptualization of flashbulb memories (Brown & Kulik, 1977), these events are retained like a photograph in the mind and believed to be impermeable to change. The fixed and indelible nature of these memories operated according to a ‘binding hypothesis'; emotionally intense responses to events activate a mechanism that fuses our emotional responses with salient features of the experience, like contextual details, and these are maintained better over time (MacKay & Ahmetzhanov, 2005). However, some have criticized this idea noting that flashbulb memories have been associated with minor to major inconsistencies over time (e.g., Neisser & Harsh, 1992), and that these may vary according to the type of event. Talarico and Rubin (2003) investigated whether flashbulb memories of the terrorist attacks on September 11th were maintained more consistently in memory than everyday memories (e.g., parties, sports events, studying). Participants were tested immediately following the
event, and then either 1-week, 6-weeks, or 32-weeks following the attacks (between subjects). Overall, they found that both traumatic (flashbulb) and everyday memories similarly decreased in number of consistent details reported over time (as assessed by responses to a series of open-ended questions such as “When did you first hear the news?”)\(^{11}\). The number of inconsistent details also increased over the interval; participants’ frequently changed their responses when asked about “distinctive” details about the event for both September 11\(^{th}\) and everyday memories. However, flashbulb memories were associated with higher ratings of recollection (i.e., “I feel as though I am reliving the event”), belief in the accuracy of the memory, and vividness that were maintained over time relative to everyday memories (Talarico & Rubin, 2003). Consistency also has been examined by investigating the retention of different types of details about a flashbulb event. For example, Christianson and Engelberg (1999) examined memory for flashbulb details and emotional reactions surrounding the Estonia Ferry disaster in Sweden. Flashbulb details of the original event (e.g., how they heard about it, where they were upon hearing the news, who they were with at the time) were recalled more consistently (ranged from 48-77% consistent) relative to participants’ emotional state and reactions to the event (ranged from 31-59% consistent). Collectively, these and other studies (e.g., Finkenauer, Luminet, Gisle, El-Ahmadi, & Filipopot, 1998; Nachson & Zelig, 2003) indicate that elements of flashbulb traumas may be retained differently in memory, and may not be as “fixed and indelible” as previously believed (e.g., Brown & Kulik, 1977).

\(^{11}\) Specific consistency mean scores for each interval cannot be reported as these were not listed in the empirical paper.
Research also has examined consistency for highly emotional and directly experienced events (e.g., Southwick, Morgan, Nicolaou, & Charney, 1997; Yuille & Cutshall, 1986). Southwick et al. (1997) conducted a prospective examination of memory for combat related trauma in 59 veterans of Operation Desert Storm one month and two years after their return from war. Results indicated that 88% of veterans provided inconsistent responses to one of nineteen memory items on a combat-trauma questionnaire (e.g., “extreme threat to your personal safety”, “seeing others killed or wounded”), and 61% changed responses to two or more items. Inconsistencies resulted from both addition of information (e.g., reporting an event at the two year interval not previously reported) and errors of omission (e.g., not reporting an event after the two year interval that had been previously reported). On the other hand, King et al. (2000) found that retrospective reports of combat experiences by veterans remained moderately consistent over a two-year period. Although these studies examined consistency of memory for directly experienced traumatic events, the results should be interpreted with caution. Consistency was measured by comparing YES/NO responses on a standardized list of combat-related experiences across two time intervals (e.g., did you see another get wounded or killed), rather than eliciting free recall responses to traumatic experiences (e.g., King et al., 2000; Southwick et al., 1997).

Previous studies have reported levels of consistency in traumatic memories ranging from 48% to 85% (e.g., Christianson, 1989; Christianson & Engelberg, 1999; Peace & Porter, 2004; Porter & Peace, in press; Weaver, 1993; Yuille & Cutshall, 1986). Longer intervals have been associated with greater impairment and a reduction in memory consistency (e.g., Neisser & Harsch, 1992; Schmolck et al., 2000). However,
studies reporting drastic decays in consistency seem to rely on memories for flashbulb
types of events that are not directly experienced, personally salient or “traumatic” per se.
Studies that have looked at the long term recall of traumatic experiences have generally
reported high levels of memory reliability. Weaver (1993) tested memories of veterans of
Operation Desert Storm at one day, three month, and one year intervals. He noted
consistency of memory for war experiences decreased between the first and second
testing period, and then remained stable up to one year following traumatic events. In a
long-term study on the retention of CSA, Alexander et al. (2005) investigated how
trauma severity influenced memory and accuracy over the course of 12-21 years.
Participants from a previous study on CSA and criminal prosecutions (Goodman et al.,
1992) were re-contacted and re-interviewed (N = 94) 10-16 years following the original
study. Participants’ were asked open-ended questions (e.g., “How old were you when the
CSA first happened?”) and responses were compared with information previously
obtained from participants and court transcripts. Overall, a mean level of 72% accurate
information was reported over time, with 14% errors of commission and 14% errors of
omission. Further, they found several predictors of memory accuracy and retention.
Trauma severity (i.e., number of PTSD criteria met) was positively related to an
enhanced memory for CSA, relative to those whose levels of trauma were less severe.
The authors acknowledge that their results combine measures of accuracy and
consistency depending on the availability in the information in previous self-reports or
court records (Alexander et al., 2005). Regardless, this prospective study has important
implications for the study of long-term recall of trauma.
Other scholars have found a linear decline in recall consistency of negative experiences (e.g., Schmolck et al., 2000). Some have argued this is to be expected based on the rates of normal forgetting (e.g., Ebbinghaus, 1885/1964; Penrod, Loftus, & Winkler, 1982). However, this may be dependent on the types of events reported. For example, Peace and Porter (2004) investigated consistency for traumatic and positive emotional experiences in 49 participants from the community over a three to four month interval. They found that 82.5% of information was consistent between Time 1 (initial interview) and Time 2 (3.5 months later) for traumatic memories, relative to 64.7% for positive experiences. Follow-up with this same sample three to five years later demonstrated that consistency in positive memories had dropped to 42.8% whereas details about traumatic memories were 74.2% consistent (Porter & Peace, in press). Thus, information about traumatic events (including autobiographical and event details) was more consistent over time than positive event memories, which declined at a more rapid rate. Traumatic events can be persistent in memory and resistant to the effects of dramatic decay over time, differing from the Ebbinghaus forgetting curve (Yuille & Cutshall, 1986). It is unknown whether this relationship will extend to genuine versus fabricated claims of trauma. Although memories for traumatic events may be well retained, they are rarely (if at all) perfectly consistent over time. Laboratory and field studies of testimonial consistency have reported high rates of consistent recall but never at a level above 95% (e.g., Brewer et al., 1999; Fisher, Geiselman, & Amador, 1989; Yuille & Cutshall, 1986).

To address the shortcomings of previous research, this dissertation used a design with both retrospective and prospective elements to allow for a detailed comparison of memories for truthful and fabricated traumas (including objective and subjective
characteristics) at three time intervals (initial, three months, six months). It has been argued that prospective designs are essential to improve the understanding of memory consistency over time (e.g., Erickson, Wolfe, King, King, & Sharkansky, 2001; King et al., 2000). As discussed, empirical research on the consistency of traumatic memory is mixed; the extent to which previous findings apply to memory for details of truthful and fabricated narratives of trauma remains unclear. Consideration of additional factors that influence long-term retention of personally experienced events (e.g., Anderson, Cohen, & Taylor, 2000) may help to elucidate the relationship between credibility and consistency.

Factors Influencing Consistency

Research on consistency of memory has failed to address various factors that may influence the stability of memory over time and across multiple recalls (Anderson et al., 2000). As a result, studies should recognize conditions that facilitate consistency or distortion (e.g., Spanos, Burgess, Burgess, Samuels, & Blois, 1999). Conway (1992) argued that memories are accessed through an organizational framework, and that they remain stable over time because of their interconnections with knowledge from other memories. However, the specific items retrieved at each successive recall may vary depending on the context of recall. Conceivably, methods of measurement, personal salience, and levels of rehearsal may influence levels of consistency evidenced in the recall of traumatic events over time.

Measurement of Consistency.

It is clear that the methods by which memory reports are elicited and consistency is measured influence outcomes in various studies. In fact, researchers have argued that
marked inaccuracies and inconsistencies may be an artifact of psychological testing (Loftus, 1979; Undeutsch, 1989). For example, many historical and present day researchers have assessed the reliability of memory based on the following paradigm: (1) present a picture, scene, or videotape for subjects to view for a short period; and (2) follow this with questions about the details of the event (e.g., see Vrij, 2005). The problem in this approach is that sometimes researchers ask an overabundance of questions concerning details from large to minute about the viewed event. Witnesses may change their responses to questions (whether by omission or addition) in successive interviews also according to the form of question (Anderson et al., 2000; Scrivner & Safer, 1988). For example, research has demonstrated that responses to closed YES/NO questions tend to result in inconsistencies over extended intervals (e.g., 1-2 years), as well as a reduction in accuracy (e.g., King et al., 2000; Poole & White, 1995). As mentioned, in their study of war veterans, Southwick et al. (1997) reported a rate of 88% inconsistent responding between Time 1 (months) and Time 2 (2 years). This exceedingly high rate of traumatic memory inconsistency likely resulted from changes to YES/NO questionnaire responses (i.e., item endorsements), rather than genuine distortions in memory over time (see previous section). Padilla-Walker and Poole (2002) found that recognition and cued recall tests were positively related to reports of difficulties in recalling prior information (i.e., forgot it all along), whereas free recall of sentences promoted greater consistency. Open-ended questions also have been associated with reliable responding over time (e.g., Memon & Vartoukian, 1996; Yuille & Cutshall, 1989). Regardless, it is unlikely that a witness will report details that are 100% consistent over extended time intervals (e.g., Peters, 1972; as cited in Undeutsch, 1989).
Related to the type of questions used when assessing consistency (e.g., specific, open-ended, free recall) is the scoring procedures used to determine what is “consistent” versus “inconsistent”. Researchers widely vary on the rigidity of their scoring procedures, from stringent (i.e., exact wording = consistent) to more lenient criteria (i.e., same general information or gist = consistent) (e.g., Christianson & Engelberg, 1999; Talarico & Rubin, 2003; Yuille & Cutshall, 1986). As a result, variations in reported levels of consistency depend upon the manner in which consistency was assessed. While some have concluded that memory is inaccurate and inconsistent, experimental design should be considered in the interpretation of results. The present dissertation measured consistency according to responses to both free-recall and open-ended questions about traumatic victimization to examine true levels of memory consistency, rather than forced choice or recognition forms of testing. Further, the scoring procedure was successfully used in previous research (e.g., Peace & Porter, 2004; Porter & Peace, in press) and offers a balance between overly strict and overly lenient criteria (see Methods for further discussion).

Memory Consolidation.

In addition, the amount of time elapsed between the original experience and when a memory report is first elicited also may influence memory consistency. According to the consolidation hypothesis, memory changes the most in immediate aftermath of a traumatic event (e.g., days or weeks) due to heightened cognitive and emotional processing during that period (Winningham et al., 2000). After this point, memory may be consolidated such that narratives are fairly consistent over time (Kvavilashvili et al., 2003; Winningham et al., 2000). Therefore, the time at which memory is initially elicited

93
following an event may influence how consistent recall appears. In a test of this hypothesis, Winningham et al. (2000) found that consistency of memories for the verdict in the O.J. Simpson murder trial varied as a function of the initial testing period. Participants were first tested either immediately (within five hours) or one week (delayed) following the event (Time 1), and again eight weeks later (Time 2). Those in the immediate condition were less consistent during subsequent memory tests relative to participants in the delayed condition. In a case study with a victim of sexual assault, Christianson and Nilsson (1989) found that her memory immediately following the experience was poor and consisted mostly of potent sensory images. Sixteen weeks later, she was able to produce a complete and detailed verbal narrative of the experience.

Although it appears her memory was enhanced over time (i.e., hypermnesia, see Scrivner & Safer, 1988), less variation would have been evidenced had there been a delay (e.g., weeks) in the initial report of the attack due to memory consolidation. These, and other studies (e.g., Smith et al., 2003), support the memory consolidation hypothesis (Winningham et al., 2000).

Other studies have found little effect of initial memory testing. In child participants, Tizzard-Drover and Peterson (2004) found that providing memory reports immediately after a traumatic experience (i.e., emergency medical treatment in children) helped to consolidate memory over a one year period on measures of completeness and accuracy. Lee and Brown (2003) reported that consistency (66%) over a seven month interval was not dependent on the timing of the first report (up to 24 hours after the event versus 10 days). In both of these studies, the first recall time was relatively soon after the traumatic/negative experience. While these interviews may have helped to maintain
memory over time, lower levels of accuracy or consistency (e.g., less than 70%) may result from continued processing of the event in subsequent weeks (Winningham et al., 2000). In this dissertation, truthful traumatic events were not elicited immediately following the original experience (due to ethical constraints), whereas fabricated traumas were generated in the experimental session with little preparation. As a result, truthful traumas should be maintained more consistently relative to fabricated traumas due to memory consolidation.

*Personal Salience.*

Variations in event recall also have been found to occur as a function of the degree of participants' personal involvement in the event, and the degree of personal and emotional salience for participants. The role of salience is critical as much of traumatic memory research has been criticized in terms of ecological validity (e.g., negative scenes or videos) and generalizability to events such as personal victimization. This issue was recognized a century ago when Stern (1910) and Whipple (1913) argued that real-world events may be recalled in a different manner than those created in laboratory settings. Several approaches to memory research have been used to remedy this problem and investigate the role of personal salience. For example, Yuille and Cutshall (1984) studied eyewitness memory using a paradigm involving witnesses seeing either a realistic “staged” robbery or a videotape of the same event. They found that amount of detail recalled was positively related to personal salience (i.e., witnessing the live event); direct involvement in the event was associated with more detailed reports relative to witnessing the event on videotape. A case study of a real criminal event (i.e., robbery and homicide)
by the same authors revealed that witnesses central to an event recalled the experience in more detail that those that were more peripherally involved (Yuille & Cutshall, 1986).

Similarly, studies of flashbulb memories also support the idea that recall differs as a function of the degree of involvement and salience of an event. Neisser et al. (1996) compared memories for the earthquakes during 1989 in California among a group of moderate (Berkeley), severe (Santa Cruz), and remotely (Atlanta) involved participants. Those who were more directly involved in the event retained better and more accurate memories of the experience even after a 1.5 year delay. Er (2003) studied memories for the Marmara earthquake in Turkey across groups who had direct involvement versus those who had only heard about the event. They found that memory accuracy and consistency was the highest in groups who were directly victimized by the experience over a 1-year period. Comparisons of memories for the death of Princess Diana and the events of September 11th also demonstrated that personal significance facilitates memory retention: British participants (higher personal/emotional salience) recalled both of these experiences more consistently over time than Italian participants (lower salience) (Kvavilashvili et al., 2003). Such findings could be explained by the narrative-rehearsal hypothesis (Neisser et al., 1996): direct experience of a traumatic event is linked to more rehearsal of the autobiographical (i.e., personal circumstance) details of the experience due to social interactions with people (i.e., asking how you are coping, etc.) and individual thought processes in rationalizing “why” it happened. In turn, the more frequent an event is rehearsed, the more durable the memory.

Related to the idea that general memory mechanisms can account for more consistent traumatic memory recall, Pezdek (2003) proposed that traumatic experiences
may be retained in two forms of memory: event memory (i.e., memory for the details of the experience itself) and autobiographical memory (i.e., personal circumstances and feelings about the event). Specifically, higher levels of emotional significance will be attached to whatever aspect of the event that is the most vivid to the individual. In turn, emotional salience will lead to greater rehearsal and accurate retention over time. In order to test this idea, Pezdek (2003) studied memories for the September 11th terrorist attacks in three groups that varied in their proximity to the event. It was speculated that those more personally affected by the event (i.e., New York residents) would attach emotional significance to details of the event, whereas further in proximity (i.e., California and Hawaii residents) would attach emotional significance to their autobiographical circumstances. In general, her findings support the notion that emotion moderates memory retention such that enhanced recall is evidenced for the information that is considered vivid and emotional to the individual (Pezdek, 2003). The importance of emotional salience in recall was confirmed in a Canadian sample of memories for September 11th (Smith et al., 2003). They found that autobiographical details were more consistently remembered (65.3%) relative to event details (30.9%). It was argued that Canadians were more removed from the event and attached emotional significance to the circumstances of hearing the news rather than specific details of the event (Smith et al., 2003). Additional research using different events and paradigms supports this distinction (e.g., Peterson, 1999; Peterson & Whalen, 2001; Tekcan, Ece, Gülgöz, & Er, 2003).

Extant research supports the notion that consistency may vary as a function of personal and emotional significance attached to traumatic or flashbulb experiences (e.g., Er, 2003; Neisser et al., 1996; Pezdek, 2003; Qin et al., 2003). Related to the present
study, heightened emotionality and personal salience may be more prevalent in truthful relative to deceptive claims, and may differentially affect retention through rehearsal.

*Rehearsal and Retellings.*

Consistency also changes over time according to the frequency and nature of rehearsal. One perspective is that rehearsal serves to consolidate autobiographical experiences into memory (e.g., Altmann, 2003; Cohen & Faulkner, 1988; Winningham et al., 2000), especially when spaced over time (e.g., Bahrick, 1979; Rubin, 1996; Rubin, Schrauf, & Greenberg, 2003). In turn, lack of rehearsal also may erode memory with the passage of time (Van Oorsouw & Merckelbach, 2004). Past research has demonstrated that dates and some details of autobiographical events become more durable over time with repeated memory testing and rehearsal (e.g., Allen, Mahler, & Estes, 1969; Hogan & Kintsch, 1971; Linton, 1975, 1978). Rehearsal is believed to increase the availability of the experience in memory, allowing for details to be readily accessible (Rubin & Kozin, 1984).

Other studies have found a relationship between frequency of rehearsal and source confusion or memory distortion (e.g., Porter et al., 1999). For example, Suengas and Johnson (1988) evaluated the role of rehearsal frequency on the qualitative differences between real versus imagined events. Participants recalled a series of real and imagined events (e.g., eating cookies, wrapping parcels), and were assigned to rehearsal conditions of either six or twelve times per event over a 24 hour time period (with a no rehearsal control condition). Overall, details of imagined events did not decrease significantly over successive recalls and, in some cases, were inflated relative to those reported for real events. This effect was only evident in the rehearsal conditions relative
to the no rehearsal control group (Suengas & Johnson, 1988). These findings suggest that rehearsal may lessen differences between real and imagined events, making it more difficult to determine the veracity of memory reports (e.g., Johnson & Suengas, 1989). On the other hand, studies have demonstrated that memory differences between real and imagined events over time are maintained, and the discrimination of veracity remains possible over the course of successive recalls. For example, Lampinen, Odegard and Bullington (2003) studied memory for performed (real) and imagined (false) actions over multiple sessions. They found that real (performed) memories were associated with more sensory, emotional, cognitive, and kinesthetic detail relative to false (imagined) memories, even after repeated imaginings or retellings. These results are contrary to the idea that repeated imagining or rehearsal over time may limit the detectable differences between memory types. However, the data reported by Lampinen et al. (2003) did note a non-significant trend for imagined events to assimilate features of performed/real experiences over time. Rehearsal of imagined events also may be associated with embellishment and memory distortion. As a result, the consistency of imagined events would decrease over time even though narratives might be richly detailed. Furthering this idea, Mandler (1994) proposed that repeated recall attempts lead to memory elaboration as more extensive memory networks are formed (i.e., integration). In turn, there is an expansion of cues available to elicit more information during successive recall attempts. Memory elaboration by rehearsal can apply to both real and imagined events. However, it is argued that fabricated events may be more susceptible to inflation when rehearsed multiple times as it is more difficult to "keep the story straight".
Based on the above review, it is possible that rehearsal (or retesting) would serve to consolidate memory for both truthful and fabricated experiences, but this effect may not be evidenced until Time 2 and Time 3 of the design (given sufficient opportunity to rehearse). Genuine claims of trauma are already consolidated in memory and likely have been thought and talked about with others. Therefore, real traumas will be more consistent and show less decrease over time. On the other hand, fabricated traumas generated for the purpose of this study would be lacking in prior rehearsal, and would likely be less consistent over time as a result of distortion, inflation, and “telling a good story”. The manner in which deceptive claims of trauma are retained over time (with or without prior rehearsal) is a critical issue for the criminal justice system.

*Perspectives on Consistency*

While past research appears to have reached a consensus that inconsistencies undermine perceived credibility, the paucity of empirical data on the retention of truthful and deceptive claims of trauma has led to discrepant viewpoints on how memory recall will vary over time (e.g., Erdmann et al., 2004; Granhag et al., 2006; Masip et al., 2005).

One perspective is that fabricated or deceptive claims will be *less consistent* over time relative to genuine experiences, especially in terms of the salient or central aspects of events (e.g., persons, actions) (e.g., Greuel, 1992; Undeutsch, 1989). Wagenaar, van Koppen, and Crombag (1993) evaluated criteria of truthful and fabricated claims and argued that truth can be inferred from consistency in consecutive statements. Deceptive statements are less likely to contain relevant and detailed information reported in a consistent manner. In an interview with the *National* magazine, Dr. Stephen Porter noted
that “allowing time for memory to fade makes it more difficult for people to keep their stories straight” (Dotto, 2004). In a series of studies on memory for lying, Polage (2004) found that deceptive items were less believable over a five week interval relative to truthful items – what she calls “fabrication deflation”. Research by Mazzoni and Memon (2003) provides an explanation of why this may occur. Repeated imagination and rehearsal may inflate the details of events that were never experienced such that details of deceptive experiences would not be consistent over time due to imagination inflation. Some studies suggest that the more time you give someone to fine-tune a lie, the more “fantastic” it may become in successive retellings (Johnson et al., 1993). This is in line with research that has reported fabricated claims to have an exaggerated or “over the top” quality (e.g., Porter, Peace, & Emmett, in press) – elaborations and exaggerations would tend to lead to inconsistencies across reports.

Alternatively, other researchers argue that truthful and deceptive statements may be equally consistent over time and change at similar rates across recall (e.g., Granhag & Strömwall, 2000a; Granhag, Strömwall & Jonsson, 2003). This rests on the idea that in order for fabricated claims to be successful, the liar must have a good memory and engages in more rehearsal of deceptive details. Because of the cognitive load of such a task (e.g., Vrij, 2004), he/she may evidence more fluctuation or a greater drop in consistency between initial recalls but this will level off as the story gets solidified over successive retellings. With genuine experiences, the malleability of memory is associated with providing more, less, or changed details over time (e.g., Loftus, 2003b; Turtle & Yuille, 1994). Thus, consistency for truthful experiences will also decline over time. For example, Granhag and Strömwall (2002) evaluated verbal and nonverbal cues to
deception over successive interrogations. Participants were exposed to a realistic mock-crime scenario and instructed to provide truthful or deceptive reports of what they witnessed. They were interrogated about their memories on three occasions over the span of 11 days. Overall, deceptive claims were associated with significantly fewer words (although not details), and this decreased across interrogations. Further, levels of consistency for truthful and fabricated claims were relatively equivalent for information reported across interrogations (Granhaag & Strömwall, 2002). Therefore, it is possible that consistency may change at a similar rate for both event types. Lipian et al. (2004) argued that although the words of a statement may change, the story may not. Therefore, it is feasible that truthful and fabricated claims may decrease in levels of consistency while being able to "keep the story straight". Related to this is the possibility that genuine and deceptive reports may become more similar over time due to a "synthesis of experiences" (Loftus, 2003b), influencing the comparative consistency between reports (Erdmann et al., 2004).

The third perspective is that deceptive statements are maintained more consistently than reports based on genuine experience. Canadian lawyers interviewed in the National commented that event recall may appear scripted when people are lying, especially if the same narrative is provided on multiple occasions (Dotto, 2004). This is argued to result from the fact that liars are focused on remembering what they have said in previous statements, whereas truth-tellers base their reports on their memory for the event and may show more variation (Granhaag & Strömwall, 2000a). For example, Granhaag and Strömwall (2002) stated that rehearsal strategies (and the frequency of rehearsal) employed by liars may promote more consistent responding which differs from
increased variation that can result from repeated recall of a genuine memory. Variations in consistency and recall in genuine memory reports relative to deceptive claims has been recognized by proponents of CBCA techniques. In an analysis of real and false allegations of rape, Parker and Brown (2000) found that genuine victims reported a range of emotions and an unstructured narrative. In contrast, false allegations appeared manufactured, statements were highly structured and scripted over time, and descriptions lacked the “ring of truth” present in truthful allegations. It could be argued that deceptive claims are often based on a schema of traumatic victimization and could appear more stable over time due to the repeated activation of a “trauma-script” (e.g., Anderson et al., 2000).

In light of these discrepant viewpoints and the lack of conclusive empirical research on the retention of deceptive statements, it is difficult to predict the outcomes of the present study. Overall it is anticipated that memories for both types of trauma will decrease in consistency over time, but that the rate of decay for fabricated claims will be greater. Recall of “fiction” is likely to show more room for error or fallibility over time relative to events based in reality.

Summary

Past research reviewed has generally found rates of consistency ranging from 48-85%. The implication of these figures is that honest witnesses are still likely to report some inconsistent details. Further, at the lower end of the spectrum (48%), approximately half of the report has changed over time so what becomes the cut off point to invalidate a statement? Research has found that judges, lawyers, and police believe that truthful

103
statements are more consistent and more detailed over time than deceptive reports, but are generally unable to identify what constitutes consistency (e.g., Strömwall & Granhag, 2003b). While there is little dispute about details (e.g., Vrij, 2005), there remains mixed findings with respect to consistency (e.g., Granhag & Strömwall, 2002). Collectively, these studies are relevant with respect to the present dissertation as they can be extended to apply to those events that we have genuinely experienced relative to those we claim to have experienced (as identified throughout Chapters 3 and 4).

Despite varying perspectives on the long-term retention of traumatic experiences, emotional or psychological factors are also important to consider. These variables not only influence mental health in the aftermath of trauma, but can influence memory for traumatic events. In addition, the extent and manner in which trauma symptomology can be successfully malingered is a critical factor in the assessment of credibility.
Chapter 5. Trauma Symptoms and Malingering

Quote: “I am having a flashback … all the way to the bank” (Eden, 2001)

Symptoms of traumatic stress experienced in the aftermath of trauma can have profound effects upon an individuals’ life (Carlson, 1997). However, these symptoms are subjective and may be easy to “fake” (Iverson & Lange, 2006; Rogers et al., 2005). A further consideration in the discrimination of truthful and fabricated trauma is the manner in which the traumatic stressor itself is experienced and reported (e.g., Koch, Douglas, Nicholls, & O’Neill, 2006; McGuire, 2002). Fabricated reports of symptomology may be associated with distinctly different characteristics relative to genuine symptoms, and these may be maintained differently over time. This chapter reviews literature on symptoms of trauma as well as characteristics of malingered symptomology.

*Traumatic Stress*

Psychological and emotional responses to aversive events are defining features that constitute an event being “traumatic” relative to a negative or stressful experience. The *DSM-IV-TR* (APA, 2000) definition of trauma involves an intense emotional response such as fear, horror, or helplessness to an event that threatens the physical and mental integrity of an individual (refer back to Chapter 2 for definitions of trauma). Typically, it is argued that symptoms of trauma operate according to a dose-response linear relationship (e.g., Herman, 1992; Schiraldi, 2000; Sprang, 1999). That is, as the severity of trauma increases, the array and level of symptoms of traumatic stress also increase (e.g., Maercker, Beauducel, & Schützwohl, 2000; McCarroll, Ursano, & Fullerton, 1993; Steel, Silove, Phan, & Bauman, 2002). For example, Runtz and Roche
(1999) used the Trauma Symptom Inventory (TSI; Brière, 1995) to assess trauma symptoms resulting from physical and sexual child abuse in a sample of Canadian university women. Their results indicated that TSI scores were elevated for both forms of abuse (relative to other stressful or traumatic life events), and that the severity of abuse had a linear relationship with the severity of symptom scores. Specifically, CSA was related to intense emotional responses, elevated scores on most subscales, and significant destructive behaviours (e.g., risk-taking, self-harm, suicidal actions). Physical abuse was associated with slightly lower (but still clinically significant) levels of anxiety, dissociation, anger/irritability, and impaired self references (Runtz & Roche, 1999). In their research with car accident victims, Frommberger et al. (1998) found that levels of continuing post-traumatic symptoms, and memory intrusions in particular, were directly proportional to their immediate traumatic stress reactions following the event. Other studies have found high levels of traumatic stress to be associated with a broad range of traumatic events, regardless of perceived severity (e.g., Bowman, 1999). For example, Peace and Porter (2004, June) found comparable levels of traumatic stress across individuals who had been criminally victimized relative to those who had experienced the death or suicide of a loved one. It is evident that responses to trauma are highly subjective in nature; therefore, the dose-response relationship between trauma and symptom severity does not always capture the variability in levels of traumatic stress experienced following an event and whether symptoms persist (e.g., Bowman, 1999; Creamer, McFarlane, & Burgess, 2005). In addition, claims of chronic levels of psychological distress are common among malingers (i.e., individuals reporting false or exaggerated symptoms for some gain) as traumatic stress is difficult to "disprove"
(Iverson & Lange, 2006). In order to effectively discern trauma symptomology reported by individuals who are malingering, it is necessary to understand the characteristics of genuine suffering and persistent traumatic distress.

*Post-Traumatic Stress Disorder (PTSD).*

One psychological outcome following traumatic experiences is post-traumatic stress disorder (PTSD), affecting 9-15% of the general population (*DSM-IV-TR*, APA, 2000). Most individuals undoubtedly experience some form of trauma during their lives (e.g., death of a loved one). Studies have estimated that the lifetime prevalence of exposure to any type of trauma is at least 69% (e.g., Resick, 2001; Resnick, Kilpatrick, Dansky, Saunders, & Best, 1993). However, only a small minority develop PTSD as a direct result of their traumatic experience (e.g., Kamphuis & Emmelkamp, 2005; Olff, Langeland, & Gersons, 2005; Rosen, 1995). While the prevalence rate listed in the *DSM-IV-TR* applies to the general population, PTSD is typically manifested at higher rates (ranging from 3-58%) in certain “at risk” traumatized populations such as combat veterans, victims exposed to mass violence/disaster, and rape/assault victims (APA, 2000). For example, research on female victims of sexual assault has reported PTSD prevalence rates of 34-80% (e.g., Breslau, Davis, Andreski, & Peterson, 1991; Kessler, Sonnega, Bromet, Hughes, & Nelson, 1995; Solomon & Davidson, 1997; Yehuda, 1999).

In response to the devastating psychological outcomes experienced by veterans of the Vietnam War, researchers began to carefully evaluate the constellation of symptoms they exhibited. This became characterised as PTSD and was first included in the *DSM-II* in 1980. However, it was quickly realized that Vietnam veterans were not the only individuals afflicted with this disorder (Flouri, 2005). With continued study, the
diagnostic criteria for PTSD have been refined in subsequent editions of the *DSM* and are now applicable to a range of traumatic experiences (refer to discussion on defining trauma in Chapter 3) (Flouri, 2005; Gold, Marx, Soler-Baillo, & Sloan, 2005; Schiraldi, 2000). Based on this body of research, the *DSM-IV-TR* (APA, 2000) has laid out the following diagnostic criteria reflecting the hallmark characteristics of post-traumatic responses (in general) and specified levels of pathological disturbances. Three core clusters of symptoms exemplify the defining features of PTSD. The first is *re-experiencing* of the traumatic event, involving recurrent and intrusive thoughts, dreams, and flashbacks of the event. Intense distress and physiological reactivity to trauma-related stimuli also are indicators of re-experiencing (Criteria B). The second cluster involves active *avoidance* of thoughts and feelings, people and places, and memories of the traumatic event. Avoidance symptoms also are associated with a numbing of general responsiveness to daily activities and life plans, feelings of detachment, and restricted or blunted affect (Criteria C). *Hyperarousal* is characterized by disturbances in sleep and concentration, irritability/anger, hypervigilance and an exaggerated startle response (Criteria D) (APA, 2000; Bury & Bagby, 2002; Schiraldi, 2000). Clinical levels of PTSD are evidenced when individuals persistently experience at least one sign of re-experiencing, three or more signs of avoidance, and two or more symptoms of hyperarousal for more than a month. The disorder is classified as *acute* if the duration of symptoms is less than three months, or *chronic* if symptoms persist for longer than three months. While approximately 50% of individuals suffering symptoms of PTSD recover within three months following trauma, many others experience persisting symptoms of PTSD for twelve months or longer (*DSM-IV-TR*; APA, 2000).
Various factors have been reported as predictors of the development of symptoms of traumatic stress or PTSD, such as the severity of trauma (e.g., Jeavons, 2000; McGruder-Johnson, Davidson, Gleaves, Stock, & Finch, 2000) and the proximity to the event (e.g., Blanchard, Rowell, Kuhn, Rogers, & Wittrock, 2005). A literature review by Andreotti, Bucci, Paloscia, and Marozza (1999) noted associations between childhood trauma, and particularly CSA, with the development of later PTSD, somatoform, mood, and dissociative disorders. In addition, proximity and level of involvement in the event appears to directly impact psychological symptoms experienced. Fivush, Edwards, and Mennuti-Washburn (2003) analyzed responses in diaries participants used during a writing exercise about emotions experienced in the aftermath of September 11th. Although the authors did not directly elicit memory reports for participants’ autobiographical experience of the terrorist attacks, they found that individuals more directly involved in the events (e.g., knew someone who had been killed) had higher emotional reactions that were maintained over a 6 month period (e.g., see Pezdek, 2003). Schlenger et al. (2002) assessed a large national sample of participants across different metropolitan areas after the September 11th terrorist attacks and found rates of probable PTSD in 11.2% of New Yorkers relative to 4.3% residents of other states. Longitudinal research on PTSD with those directly involved or injured in the September 11th attacks has reported prevalence rates of 14-37% (Grieger, Fullerton, & Ursano, 2003; Resnick, Galea, Kilpatrick, & Vlahov, 2004). Thus, proximity to and involvement in trauma may increase the likelihood of traumatic distress.

In general, extant research indicates that symptomology generally diminishes over time with the exception of cases of chronic PTSD (e.g., Eid, Johnsen, & Thayer, 2001;
Resick, 2001). Intrusions and avoidance may be associated with symptom/memory consolidation or amplification (i.e., reporting more symptoms), such that these remain relatively stable or increase with time (e.g., Fivush et al., 2004; King et al., 2000). Given these findings, comparing truthful and fabricated symptoms of traumatic stress may provide a useful method of differentiating claims. Self-reported trauma symptomology also may become critically important when corroborating evidence of trauma does not exist (e.g., recovered child sexual abuse, claims of psychological damage). In such cases, clinicians often rely on measures of traumatic stress to distinguish genuine individuals from those who are malingering for some type of internal or external gain (e.g., attention, compensation) (see Koch et al., 2006). Especially when litigation is associated with claims of psychological distress, symptoms often are assessed on multiple occasions over lengthy delays (Rosen, 2004). Therefore, it is also relevant to consider how truthful and fabricated symptomology are reported over time. Are these symptoms consistently maintained, or do fabricated reports of trauma symptomology contain more variation or inflation over time relative to truthful symptoms? Knowledge concerning the possible differences in truthful and fabricated trauma symptoms is critical as judges and juries often experience great difficulty in determining the credibility of testimony regarding traumatic events (e.g., Porter, Campbell, et al., 2003).

**Malingering Trauma**

False claims of injuries and psychological distress have been of interest to clinicians for well over a decade (Rosen, 2004). On the other hand, empirical study of fabricated symptomology has been neglected until recently (Geraerts et al., in press;
Koch et al., 2006). In clinical contexts, *malingering* can be defined as the intentional production of symptoms or illness motivated by some type of external incentive (e.g., APA, 2000; Koch et al., 2006; Rogers & Cruise, 2000). Specific to this dissertation, malingering refers to the fabrication of mental health symptoms associated with traumatic victimization (e.g., Meyer & Deitsch, 1996).

Research has shown that malingering often is used in clinical and forensic settings to evade responsibility, lend support to a defense strategy (e.g., intoxication), or raise doubts about an individual’s involvement in a crime (e.g., Calhoun, Earnst, Tucker, Kirby, & Beckham, 2000; Elhai, Gold, Sellers, & Dorfman, 2001; Gutheil & Sutherland, 1999; McGuire, 1999; Porter, Birt, Yuille & Herve, 2001; Rogers et al., 2005). The fabrication of mental illness as a defense is a well-known phenomena, but less attention has been directed towards victims of trauma. Many legal cases exist based on the principle of “psychological damages” resulting from a traumatic incident (Iverson & Lange, 2006). For example, Anna Ayala was charged and arrested in connection with the “finger in the chili” hoax at a Wendy’s Restaurant in San Jose, California in April 2005. Prior to her arrest, she had launched a lawsuit against the restaurant chain suing for psychological damages based on a false allegation of trauma and distress (Dornin, 2005). Rogers (1997) argued that such claims occur according to an adaptational model of malingering in which individuals evaluate the costs (e.g., likelihood of being caught) and benefits (e.g., escaping penalty, revenge) of malingering. If the benefits outweigh the costs, it would be considered advantageous to fabricate symptomology. Malingers are often playing a “strategic role” in order to obtain some objective (van IJzendoorn & Schuengel, 1996). Even genuine victims of stressful and traumatic events may be tempted
to exaggerate or fabricate psychological distress in order to attain some external goal. Binder and Rohling (1996) found that patients seeking compensation for a traumatic injury reported significantly more symptoms relative to patients not involved in legal processes, independent of the severity of trauma experienced.

As a result, the fabrication of psychological symptoms is a legitimate concern when evaluating reports of distress associated with traumatic experiences, particularly in cases of civil or criminal litigation (e.g., Burges & McMillan, 2001; DeClue, 2002; Lees-Haley, 1989). As recommended by the DSM-IV-TR (APA, 2000), evaluators should carefully appraise the pattern of symptom reporting when a medico-legal context is present (e.g., compensation seeking), and substantial discrepancies between subjective versus objective assessments of distress exist. Although the current study was among the first to examine false symptoms of victimization over time, several studies have attempted to discern reliable cues for detecting malingered symptoms of trauma in general or personality profiles (e.g., Early, 1990; Holden & Evoy, 2005; McGuire, 1999; Resnick, 1997).

**Prevalence of Malingered PTSD and Traumatic Stress.**

PTSD symptomology is highly subjective and clinical diagnosis relies primarily upon patients’ self reports of trauma symptomology (Elhai et al., 2001). As a result, PTSD is relatively easy to mangle (e.g., Burges & McMillan, 2001; Guriel & Fremouw, 2003; Kozaric-Kovacic & Borovecki, 2005; Sparr & Atkinson, 1986). Although some researchers suggest malingering is a rare phenomenon (e.g., Gerson, 2002), others argue that malingering PTSD symptomology occurs more often than we would like to believe.
(e.g., Calhoun et al., 2000; Candel & Merckelbach, 2004; Lees-Haley, 1992; McGuire, 1999; Resnick, 1997).

As identified previously in this chapter, only 9-15% of people develop PTSD in response to traumatic events according to strict *DSM-IV-TR* criteria (APA, 2000). Rates of alleged PTSD when litigation is involved are disproportionately higher (Binder & Rohling, 1996). For example, Rosen (1995) found that 19/22 (86%) victims of the Aleutian Enterprise marine disaster reported PTSD symptomology. This inflated rate of PTSD was unlikely and incompatible with estimates in the clinical literature with regard to exposure to a wide variety of different types of traumatic events\(^\text{12}\). Upon further consideration of the findings, Rosen (1995) found that many victims were involved in civil law suits and had claimed workers compensation. This was associated with coaching from lawyers who educated their clients about the symptoms of PTSD, such that individuals could successfully feign the disorder to add merit to their litigation. While definitive prevalence rates for malingered PTSD are not known, Lees-Haley (1997) estimated that as many as 20-30% of individuals involved in litigation to seek monetary compensation could be malingerers. Resnick (1997) found that 40% of individuals who reported being completely disabled as a result of PTSD and who were receiving compensatory benefits showed no sign of disability in reality. Other researchers have argued that falsely reported psychological distress is much more common than these estimates (e.g., Gerson, 2002). The high prevalence of fabricated reports of trauma symptoms is disconcerting and warrants further investigation. Can we differentiate malingered reports of traumatic stress from those who are genuinely suffering?

\(^{12}\) In fact, Kessler et al. (1995) estimated the rate of PTSD following natural disasters to be less than 5%.
Malingered Profiles of Trauma.

Researchers have investigated malingering and PTSD in different populations, including motor vehicle accident survivors (e.g., Blanchard & Hickling, 2004; Hickling, Blanchard, Mundy, & Galovski, 2002), survivors of maritime disasters (e.g., Eid, Johnsen, & Thayer, 2001; Rosen, 1995; Rosen, 2004), workplace accident victims (e.g., Bury & Bagby, 2002), PTSD-naïve individuals (e.g., Burges & McMillan, 2001), child sexual abuse survivors (e.g., Elhai et al., 2001), personal injury plaintiffs (e.g., Lees-Haley, 1992; Lees-Haley, 1997), and combat veterans (e.g., Calhoun et al., 2000; DeViva & Bloem, 2003; Frueh et al., 2005; Frueh & Kinder, 1994; Liljequist, Kinder, & Schinka, 1998; McCaffrey & Bellamy-Campbell, 1989; Van Atta, 2005). For example, Frueh et al. (2005) reported that approximately 5% of individuals seeking treatment and compensation for combat-related PTSD in their sample had made false claims of being a Vietnam veteran or having war exposure. When compared to veterans with confirmed combat experiences on measures of PTSD, depression, and other clinical variables, there were no significant between group differences.

Additional studies have demonstrated that levels of traumatic stress can be successfully fabricated and may be difficult to discern from genuine claims (e.g., Calhoun et al., 2000; Hickling et al., 2002; Porter, Peace, & Emmett, in press). For example, Lees-Haley and Dunn (1994) found that 86% of naïve undergraduates could fabricate a PTSD diagnosis. Similarly, Burges and McMillan (2001) found that 94% of undergraduates could successfully simulate PTSD. These findings indicate that fabricated symptoms of trauma may be difficult to detect relative to truthful symptoms. Lees-Haley (1990) found that scores on the Impact of Event Scale (IES; Horowitz, Wilner, &
Alvarez, 1979) were highly similar between fabricated and genuine reports. Further, McGuire (2002) compared a trauma group (all of whom had an active compensation claim) with undergraduate students who had not experienced significant trauma in the past year (no trauma group). Students were instructed to complete the IES as though they were in the midst of a compensation claim and were experiencing emotional problems as the result of a traumatic experience. The results demonstrated that it was difficult to differentiate between the trauma and no-trauma groups. The only distinguishing feature was an elevated score on the intrusion subscale in the malingers (McGuire, 2002). DeClue (2002) reported a case of a forensic psychological assessment with a 20 year old male. The client’s scores on psychological testing yielded valid profiles and generally were not above cut off levels on measures indicating inconsistent, infrequent, or unusual responding and impression management.

Despite the lack of symptom differences in the above research, other studies have found variations in truthful and fabricated trauma symptoms. Liljequist et al. (1998) compared PTSD symptomology (on the Personality Assessment Inventory; PAI; Morey, 1991) among genuine PTSD sufferers (i.e., war veterans) and a student control sample that was instructed to mangle. They reported a trend where malingers scored higher than veterans with PTSD on six different clinical scales assessed: depression, schizophrenia, anxiety, somatic complaints, paranoia, and antisocial features. Malingers also scored lower on the alcohol problems subscale relative to the PTSD group, and demonstrated highly elevated scores on the Negative Impression Management validity subscale (pointing to the utility of validity scales in assessment measures). Overall, malingers exaggerated and inflated their reports of psychopathology in an attempt to
appear genuinely disordered (Liljequist et al., 1998). Lees-Haley (1984) reported that malingers tend to include contradictory and exaggerated symptoms, as well as “academically correct” symptom descriptions, in their psychological profiles. In addition, this author found that memory reports regarding illness and criminal events contained a rehearsed quality, with excessive logical consistency, obtrusiveness, and reports of delusions relative to non-malingered reports (Lee-Haley, 1984). Similarly, Lanyon (1996) reported that deceptive individuals deliberately endorse items indicating virtue or extreme adjustment, engage in self-enhancement, random or stereotypical responding, and symptom over-endorsement.

Most closely related to the present study was research conducted by Porter, Peace, and Emmett (in press). In this study, undergraduate participants ($n = 126$) provided truthful and fabricated symptoms of trauma. Across all measures of traumatic stress, fabricated symptoms were exaggerated and reported at significantly higher levels than genuine symptoms of trauma. The level of potential PTSD related to truthful traumas in this sample was 28.6%, relative to 50.8% for fabricated traumas. In addition, 82.5% of fabricated profiles on the Trauma Symptom Inventory (Brière, 1995) were classified as valid according to the validity scales. Interestingly, this “over the top” quality to symptom profiles was not manifest in higher impression management scores, suggesting a level of sophistication in deceptive behaviour (Porter, Peace, & Emmett, in press).

Symptom profiles for PTSD are subjective and potentially easy to fake, especially when knowledgeable about the disorder (e.g., Bury & Bagby, 2002; Eldridge, 1991; Sparr & Atkinson, 1986). Burges and McMillan (2001) suggested this is especially so when participants are given standardized self-report checklists, where both trained and
naïve malingerers can successfully feign PTSD criteria. Hickling, Taylor, Blanchard, and Devineni (1999) examined the effect of knowledge about traumatic symptomology on the ability to successfully fake psychological symptoms surrounding a motor vehicle accident. Half of their participants had training in the diagnostic criteria for PTSD, with the other half remaining naïve to this information. All participants were asked to simulate how they thought an accident survivor would respond on psychological tests of PTSD and trauma symptoms. Their results indicated that those who were trained reported higher levels of symptomology on all psychological symptoms. In addition, when compared to genuine profiles from accident victims, trained malingerers reported a different pattern of depressive, anxiety, and intrusion/avoidance scores (i.e., greater symptom endorsement). Similarly, Marshman (2001) studied knowledgeable versus naïve malingerers on levels of PTSD symptoms and traumatic stress. The results indicated that 84% of participants who had knowledge about the specific symptoms of PTSD were able to fake the disorder without detection on the Trauma Symptom Inventory (TSI; Brière, 1995) and the Posttraumatic Stress Diagnostic Scale (PDS; Foa, 1996), compared to 18% of the naïve malingerers. In contrast to these findings, Edens, Otto, and Dwyer (1998) used the TSI (Brière, 1995) in a sample of 155 college students to determine whether past histories of trauma (i.e., knowledge) enabled participants to successfully malinger trauma symptoms on this inventory. All participants were asked to feign traumatic stress as if they had been recently victimized. Results indicated that participants who had experienced prior traumas were no more successful at avoiding detection of malingering than those with no such history, and that the validity scales provided a reliable way to discern between true and malingered symptoms of trauma. Using a different methodology, Buckley, Galovski,
Blanchard, and Hickling (2003) had actors feign post-traumatic stress disorder and compared their responding on the emotional stroop task to outcomes of genuine PTSD patients. All of the actors were coached by psychologists on the symptoms of PTSD. Their results indicated that individuals who were dissimulating PTSD demonstrated similar overall response latencies on the stroop task (e.g., delays) relative to genuine patients. However, the actors were unable to replicate the difference in delayed responding when confronted with neutral words relative to PTSD-threat related words in the same fashion that genuine PTSD patients demonstrated (i.e., increased delay for threat-related words). Much research suggests that individuals who are educated in traumatic symptomatology may have an increased ability to evade detection when malingering, and can provide symptom reports that appear to be genuine (e.g., Bury & Bagby, 2002). It is reasonable to assume that even “naïve” undergraduate participants (frequently students in psychology classes) have some knowledge of symptoms of traumatic stress (e.g., Burges & McMillan, 2001).

*Detecting Malingering.*

In light of these findings, researchers have investigated the ability of various psychometric tools and tests to accurately detect fabricated symptoms of traumatic stress and PTSD (Rogers et al., 2005). These include Symptom Validity Testing (e.g., Nies, 2005), the Personality Assessment Inventory (e.g., Calhoun et al., 2000; Liljequist et al., 1998), the Rorschach Inkblot Test (e.g., Frueh & Kinder, 1994), both versions of the Minnesota Multiphasic Personality Inventory (e.g., Elhai et al., 2001; Fairbank, McCaffrey, & Keane, 1985; Lees-Haley, 1989; Lees-Haley, 1992; Lees-Haley, 1997; McCaffrey & Bellamy-Campbell, 1989), and the Trauma Symptom Inventory (e.g.,
Edens et al., 1998; Elhai et al., 2005; Guriel et al., 2004). While some of these psychometric tests have shown promise as tools to detect malingered PTSD (i.e., those with validity scales), research has yielded mixed results. Candel and Merckelbach (2004) suggest that screening for malingering using tools such as the Structured Inventory of Malingered Symptomatology (see Merckelbach & Smith, 2003; Smith & Burger, 1997) should be conducted especially in cases of suspect reporting or where external incentives exist with respect to symptoms of traumatic stress. To date, there is no ‘gold standard’ psychometric tool that is universally recognized as being able to detect malingered symptoms of PTSD (see Guriel & Fremouw, 2003, for a comprehensive review).

For the purpose of this dissertation, traumatic symptoms and PTSD were assessed using three different measures. The Trauma Symptom Inventory (TSI; Brière, 1995) was used to assess various clinical indicators of traumatic stress across a range of subscales (e.g., dissociation, depression). Further, this measure includes validity scales that were used to evaluate the veracity of the profile. The Impact of Event Scale – Revised (IES-R; Weiss & Marmar, 1997) was chosen as a measure of traumatic distress that includes evaluation of intrusive, avoidant and hyperarousal experiences. The IES-R has been correlated with direct measures of PTSD, and also is sensitive to picking up generalized non-clinical levels of traumatic stress in community samples (e.g., Creamer, Bell, & Failla, 2003). Finally, the Post-Traumatic Stress Disorder Checklist (PCL; Weathers, Litz, Huska, & Keane, 1994) was used as a screening tool to assess self-reported levels of PTSD symptomatology and is widely used in this capacity (e.g., Dobie et al., 2006). Scores on each of these measures for both truthful and fabricated traumas were elicited to determine if different patterns exist as a function of trauma type.
Summary

Fabricated symptomology is problematic as clinicians are generally poor at discriminating accurate versus inaccurate representations of symptoms (Rosen, 2004). Thus, research is needed to clarify profiles of those who are attempting to fabricate psychological disorders. Are there discernable characteristics in fabricated symptomology even if standardized measures have labelled them as "valid"? Taken together, the studies reviewed above provide evidence that malingered symptoms of trauma may be qualitatively different from genuine reports on certain criteria. However, no studies (to my knowledge) have examined truthful and fabricated symptoms of trauma on multiple measures and over an extended time interval. In the present study, trauma symptom profiles were examined to determine if real and falsified accounts of trauma can be differentiated based on the patterns of traumatic stress reported. In forensic contexts, awareness of how malingers of trauma report events and respond on trauma scales relative to genuine victims is critical (Marshman, 2001). This research also investigated the degree to which falsified profiles of PTSD and other types of trauma symptomology can (or cannot) be maintained consistently over time, and what types of variations in falsified reporting are demonstrated over an extended time period and successive recall attempts.
Chapter 6. Introduction to the Present Study

Over the past several decades, the multi-faceted nature of traumatic memory has become a major focus of both psychological research and debate (e.g., Loftus & Davis, 2006; MacIntosh & Whiffen, 2005; McNally, 2003a; Peace & Porter, 2004). My review of the literature has demonstrated that – while consensus has been reached in some fundamental areas (e.g., memory is reconstructive) – controversies concerning the retention and detection of false allegations of trauma remain, and are under investigation in this dissertation.

Summary and Hypotheses

Several divergent perspectives (as reviewed) exist regarding the nature of traumatic memories for genuine experiences. It has been argued that trauma can impair (e.g., van der Kolk, 1996) or enhance (e.g., Porter & Peace, in press) memory relative to other types of events. Generally, research suggests that memories for personally-experienced traumatic events can be well-remembered and reliable over time (e.g., Berntsen, 2002; Peace & Porter, 2004), with both brief and lengthy follow-up periods (e.g., Merritt et al., 1994; Peterson & Parsons, 2005; Porter & Peace, in press; Wagenaar & Groeneweg, 1990; Yuille & Cutshall, 1986). However, empirical research evaluating the features of truthful versus fabricated reports of trauma is lacking. Much research using the reality monitoring framework has examined differences between real and imagined events based on self-reported (subjective) characteristics (e.g., Barnier et al., 2006; Johnson & Raye, 1981; Johnson et al., 1988; Masip et al., 2005; Mitchell & Johnson, 2000; Suengas & Johnson, 1988). Alternatively, past research in credibility
assessment (e.g., SVA/CBCA) has relied on objective measures (i.e., others’ ratings of unstructured production, detail, and consistency between statements) to detect deception (e.g., Merckelbach, 2004; Vrij, 2005). With few exceptions (e.g., Parker & Brown, 2000; Porter et al., 1999; Sporer, 1997), a detailed investigation of narrative features of both truthful and fabricated personally experienced traumatic events has not been conducted.

One goal of this dissertation was to combine these approaches to studying memory and credibility by evaluating the characteristics of truthful and fabricated reports of trauma using both subjective and objective measures of assessment (MAP; Porter et al., 1999). Based on this approach, it is possible that features of a verbal or written statement can be used to differentiate between truthful and fabricated trauma narratives. Comparing truthful and fabricated narratives using a within-subjects design allows for an assessment of deceptive behaviour as it may occur in everyday life. It is unlikely that individuals generate fabricated experiences without any basis in their own personally experienced realities. This may be particularly true with traumatic events – reliance on previous experiences would be used to enhance the credibility of a report of victimization (e.g., Loftus, 2003b). However, attempts to produce deceptive but convincing stories has resulted in a few reliable differences between truthful and fabricated reports due to the complexity of the task (DePaulo et al., 2003; Vrij, 2004). Truthful statements appear to be rich in detail, including both contextual and sensory detail, relative to fabricated narratives. On the other hand, fabricated statements contain more cognitive operations and require more mental effort as liars attempt to “keep their stories straight” (e.g., DePaulo et al., 2003; Granhag, Andersson, Strömwall, & Hartwig, 2004; Porter, Doucette, Woodworth, Earle, & MacNeil, 2006; Porter, Peace, & Emmett, in press; Vrij,
2000). In addition, fabricated narratives of trauma may have an exaggerated “over the top” quality, particularly in self-reports of memory features (Porter et al., 1999; also see Vrij, 2005 for review).

Based on the literature, it was predicted that narratives of truthful and fabricated traumas would be qualitatively different. While fabricated claims may have “the ring of truth”, I anticipated that deceptive reports would have discernable characteristics (both objective and subjectively assessed) that would be indicative of veracity – such as levels of detail, emotionality, and contextual information. Specifically, truthful narratives would be more detailed, emotional, contextually embedded, and plausible relative to fabricated narratives. On the other hand, fabricated narratives of trauma may be more coherent and have a rehearsed quality as a result of increased cognitive efforts to “keep the story straight”. Subjective indices of the narratives (self-rated memory and emotional characteristics) also were predicted to be exaggerated in the fabricated narratives in an attempt to appear credible.

Another goal of this dissertation was to examine how the features of retellings of truthful and fabricated narratives are retained or change over time. Witnesses to a crime often provide multiple statements about the event; they are re-interviewed on several occasions prior to testimony and may have to testify more than once. Only recently has research begun to address the effect of multiple interviews on the features of memory reports for truthful and fabricated experiences (e.g., Granhag & Strömwall, 2001b; Granhag & Strömwall, 2002; Granhag et al., 2006). Brewer et al. (1999) reported that testimonial inconsistencies often are considered synonymous with memory inaccuracy and potential fabrications, and less likely to be believed by potential jurors (see also Bell
& Loftus, 1989). Thus, the concept of “truth” is shaped by our perceptions and the indicators relied on in making assessments of credibility (e.g., Fisher & Cutler, 1995). Consistency is used as a cue to justify decisions that a person is either being deceptive or truthful (Gran Hag & Strömwall, 2001b). Some research has provided support for the trauma superiority argument in that traumatic experiences may be resistant to temporal deterioration relative to other types of emotional events (e.g., Porter & Peace, in press). This is argued to operate as a function of personal salience and rehearsal (e.g., Lampinen et al., 2003; Peterson & Whalen, 2001; Pezdek, 2003; Winningham et al., 2000). On the other hand, others have argued that traumatic memories do not operate under a special mechanism preserving their integrity over time (e.g., Talarico & Rubin, 2003).

This dissertation extended these perspectives to ascertain whether superior recall of genuinely experienced trauma would still be evidenced relative to fabricated allegations of trauma, which could be similarly well-rehearsed and emotionally provocative. It was predicted that the characteristics and consistency of truthful and fabricated claims would vary differentially over time. Based on the trauma superiority argument, it would be expected that both truthful and fabricated reports of trauma may share similar features during the initial phase of the study (especially in a within-subjects study). However, only truthful traumas would be sufficiently emotional and personally salient to facilitate improved (or more consistent) recall relative to fabricated traumas over time. It was hypothesized that genuine experiences would maintain higher levels of consistency (relative to deceptive claims) across the three time intervals as measured by direct questions about the traumatic experiences. However, it also was predicted that objective and subjective memory characteristics of false allegations may “level off” over
time and with retelling such that these demonstrate a degree of “memory” consolidation as well.

The last major goal of this dissertation was to examine response patterns associated with malingered traumatic stress. Symptoms of traumatic stress following trauma are highly subjective (e.g., Bury & Bagby, 2002; McGuire, 2002) and may be easy to successfully fabricate (e.g., Porter, Peace, & Emmett, in press; Rogers et al., 2005). The ability to reliably differentiate malingered (fabricated) from genuine (truthful) symptomatology in both criminal and civil contexts is of major importance. Research has demonstrated that involvement in litigation and compensation-seeking is associated with a higher proportion of false allegations of traumatic stress (e.g., Lees-Haley, 1997; Resnick, 1997; Rosen, 2004). However, the manner in which symptoms of truthful and fabricated traumatic stress differ in general and over time is unclear.

Drawing on the above approaches to memory and credibility assessment, self-reported symptoms pertaining to genuine and deceptive traumas were elicited. Based on my review of the research, it was predicted that fabricated symptomatology would show a different pattern of reporting relative to truthful traumatic stress. While studies have found few differences between malingerers and genuine victims (e.g., Frueh et al., 2005), other research has demonstrated some (if not small) variations between groups that warrant further investigation (e.g., Liljequist et al., 1998; McGuire, 2002; Porter, Peace, & Emmett, in press). Specifically, it was hypothesized that fabricated symptomatology would be associated with an “over the top” (exaggerated) quality. As such, levels of traumatic stress, PTSD, and other clinical symptomology (e.g., depression) would be inflated relative to genuine levels across all time periods. Deceptive claims would not
only be exaggerated during the initial phase, but this rate of responding is likely more consistent over time relative to genuine symptoms (which tend to recede with the passage of time).

**Study Design**

In order to accomplish the objectives of this study, a within-subjects design was used. The within-subjects approach used in this dissertation was modeled after several important studies (e.g., Johnson et al., 1988) and has been used in more recent empirical research evaluating the characteristics of different narratives (e.g., Caso et al., 2006; Kealy et al., 2006; Lampinen et al., 2003; Polage, 2004; Porter & Peace, in press; Porter, Peace, & Emmett, in press; Santtila et al., 1999; Sporer, 1997). While there are weaknesses to this approach (i.e., possible carry over effects), these were minimized by the time delay between reports (approximately one hour) and by counterbalancing the order of narrative recall. In addition, any possible carry over effects reflect the manner in which individuals are deceptive outside of the experimental context (i.e., using details or schemas from experienced events to generate fabricated statements). This type of design allows for a statistically powerful, detailed and realistic approach to the study of memory, consistency, and symptomatology in credibility assessment.

Participants were randomly assigned into one of two recall conditions based on the counterbalancing order in which they recalled truthful (TT) and fabricated (FT) traumatic experiences (within groups). The types of genuine traumatic experiences recalled were self-selected based on the participants' past history of victimization and their willingness to describe their past experience. Both objective and subjective memory features were
compared across truthful and fabricated reports using the Memory Assessment Procedure (MAP; see Porter et al., 1999). The MAP evaluated the following objective features of narratives: (1) *Number of details* (i.e., the number of distinctive pieces of information provided); (2) *Emotional components* (i.e., the number of emotional words used in reference to self and others); (3) *Time and place details* (i.e., the number of details provided indicating the time and location(s) of the experience); (4) *Coherence* (i.e., rating from 1 [not at all] to 7 [extremely] of the logical flow of the beginning, middle, and end of the narrative); (5) *Relevance* (i.e., rating from 1 [not at all] to 7 [extremely] of the extent to which information provided related to the event being described); (6) *Plausibility* (i.e., rating from 1 [not at all] to 7 [extremely] of the extent to which the report is realistic, logical, likely to occur in real life); and (7) *Number of words*. This procedure also assessed multiple subjective features of narratives using the Emotional Memory Survey (EMS; Porter et al., 1999). Participants rated their experiences from 1 (not at all/poor/none) to 7 (extremely/very/many times) on the following criteria: (1) anxiety/stress; (2) talking about event; (3) thinking about event; (4) vividness/clarity; (5) overall quality; (6) coherence; (7) sensory details; (8) confidence in accuracy of memory; (9) perceived credibility; (10) “turning off” emotions; (11) becoming emotional; (12) detached/distanced; and (13) emotional intensity. These measures allowed for a detailed examination of narrative features (and self-reported emotionality) as a function of trauma veracity and to test whether reliable indicators of credibility could be identified. The MAP was chosen for use in the present study to a) provide a holistic approach to memory assessment investigating both objective and subjective features, and b) to investigate the discriminative power of this technique to determine its utility in credibility assessment.
contexts. While the relative discriminability of Criteria-Based Content Analysis (CBCA) and Reality Monitoring (RM) approaches have been tested, no studies have evaluated the MAP in this respect. In addition, the former approaches rely on either objective or subjective assessment features of narratives whereas the MAP provides a combined approach to studying narrative differences (refer back to Chapter 3 for a detailed discussion of the MAP).

Each participant also was tested at three different times (within group) to assess the consistency of truthful and fabricated narratives. Participants were tested at Time 1, after approximately three months (Time 2), and six months (Time 3). At each time point, participants provided a narrative for the same truthful trauma narrative they described during the initial phase. Participants also generated a fabricated narrative of trauma at Time 1 and recalled the same fabricated narrative at Times 2 and 3. Consistency was assessed using the Consistency Questionnaire (CQ), designed for previous research on the long-term retention of trauma (see Peace & Porter, 2004; also see Porter & Peace, in press). A set of direct questions were used to elicit information about both traumas (i.e., what was the calendar date of the event?) and administered at each testing time. This enabled a comparison of variations in specific details about trauma between truthful and fabricated narratives over time. This type of information is particularly important in the legal system where victims are repeatedly asked to describe minor and major details of traumatic events (Bell & Loftus, 1989; Brewer et al., 1999).

Finally, participants were assessed on their levels of real and fabricated traumatic stress using several self-report measures, including the Impact of Event Scale - Revised (IES-R; Weiss & Marmar, 1997), the Post-traumatic Checklist (PCL; Weathers et al.,
1994), and the Trauma Symptom Inventory (TSI; Brière, 1995). All ratings associated with fabricated traumas were to be completed as if the event had been genuinely experienced and according to how participants thought a truthful victim would respond.

Conclusion

This dissertation is the first study, to my knowledge, to provide a detailed and comprehensive investigation of narrative features, consistency, and symptomology in an attempt to discern reliable cues to deception that can be used in credibility assessment. The limited research concerning false allegations of trauma experienced in adulthood was the major impetus for this dissertation. This research is important for the basic understanding of memory processes and retention, as well as its application to the criminal justice system.
Chapter 7. Method

Participants

Undergraduate participants were recruited through psychology courses at Dalhousie University according to the Department's Experimental Credit Point System, using the standard Participant Contact Form (see Appendix A for the Recruitment Advertisement Form). Advertisements were posted online and potential participants could sign up for a study “comparing truthful and fabricated events over time” in exchange for 2 credit points per experimental session (up to a maximum of 4 points per course) or $5.00 per session if credit point allotment was not available. To be eligible to participate, participants must have experienced some type of traumatic event (such as criminal victimization, assault, death of a loved one, animal attack, serious injury, etc.) after the age of 16. Participants were considered eligible if they were willing to provide a written description of this event. Participants were free to choose whatever traumatic event they had experienced and were comfortable discussing. All participants reviewed the detailed Informed Consent form (see Appendix B) and discussed this with the researcher to ensure their understanding of the sensitive nature of the study. Participants were free to leave at any point during the study if they became uncomfortable. This research was ethically approved by Dalhousie University’s Health Sciences Research Ethics Board (REB# 2004-876, see Appendix C).

During the initial experimental session (Time 1), a total of 291 participants took part. Of this sample, 233 (80.1%) were female and 58 (19.9%) were male, with a mean age of 19.64 years ($SD = 2.38$ yrs; range 16 – 33). The ethnicity of the sample was primarily Caucasian (84.5%); the remaining participants were Asian (4.1%), African
Canadian (2.1%), Aboriginal (.7%) or “Other” (8.6%) descent (e.g., Lebanese, Indian). Of this sample, 252 participants (rate of attrition = 13.4%) were successfully re-contacted and returned approximately three months later (Time 2: $M = 91.44$ days, $SD = 12.77$ days). For the final phase of the experiment (Time 3) approximately three months ($M = 72.13$ days, $SD = 16.31$ days) after Time 2, $N = 181$ participated (rate of attrition between Time 2 and Time 3 = 28.1%). The mean time interval between Times 1 and 3 was 163.84 days ($SD = 10.47$). Such a large sample was used to account for possible attrition due to the longitudinal design, to sample a wide variety of truthful and fabricated traumas, and to maintain statistical power. This study used a 2 (narrative type: truthful, fabricated) X 3 (testing time: initial, 3 months, 6 months) within subjects repeated measures multivariate design.

**Measures**

*Memory Survey Instructions.*

Written instructions were given to each participant. The instructions indicated which type of memory participants were to recall (truthful) or generate (fabricated) in which order (counterbalanced). In addition, the instructions gave information as to how to respond on questionnaires in that section of the experiment: “complete the written description and questionnaires as if the ‘fake trauma’ had happened to you and make your responses as convincing as possible”. Instructions varied for truthful traumatic experiences (see Appendix D) and fabricated traumatic experiences (see Appendix E).
Memory Report Forms.

The written narrative report forms utilized here have been used in previous research to elicit memories for traumatic events (e.g., Peace & Porter, 2004; Porter & Birt, 2001; Porter, Peace, & Emmett, in press). These forms provided participants with instructions to recall a specific event and provide as much detail as possible about the experience. For fabricated experiences, participants were asked to provide a narrative of a specific fabricated traumatic event that had never happened to them and provide as much detail as possible (see Appendices F and G).

Consistency Questionnaire (CQ).

This questionnaire was modified and expanded from a consistency measure used in a previous study (see Peace & Porter, 2004). The CQ was used to assess consistency in the details of memory reports over time (see Appendix H). Questions consisted of details from 23 factual categories (e.g., time of day, number of people at the scene, what the participant was wearing, etc.) that were applicable to both truthful and fabricated traumatic experiences. Participants provided specific responses to these questions, which were later coded for consistency over time.

The coding scheme compared responses from each time interval, and answers were rated as: (0) inconsistent, (1) partially inconsistent, (2) partially consistent, or (3) consistent. This coding scheme was used in previous research and found to be a reliable way to determine variations in responding over time (e.g., Peace & Porter, 2004; Porter & Peace, in press; see Appendix I for inclusion of this coding scheme). Due to the longitudinal design of this dissertation, comparisons were made between Time 1 (T1) to Time 2 (T2), Time 2 (T2) to Time 3 (T3), and across the extended interval of Time 1 (T1)
to Time 3 (T3). Responses were coded as inconsistent if they differed between T1, T2, and T3, and contained one or more contradictions. When the responses differed between phases, but there was no direct contradiction of information (e.g., omissions), they were coded as partially inconsistent. Partially consistent responses involved information (at least one component or fact) that was the same across T1, T2, and T3, as well as additional information that was not presented previously (and was not a contradiction). Finally, responses were coded as consistent when the words or semantic meaning of the responses were the same at each phase of the experiment. Total consistency scores were calculated for each memory type by summing the numerical ratings across the 23 event categories.

An intercoder reliability check was conducted on the consistency coding scheme by a trained coder (Rater 2) who followed the same coding guidelines as the primary coder (Rater 1). A subset \( n = 60 \) of narrative responses at Times 1 and 2 (20% of the sample)\(^\text{13}\) were re-coded based on the coding scheme (see Appendix I). Both raters were trained over multiple sessions by the primary investigator. In the first session, the coding scheme and sample item responses were discussed in detail. Each rater was given samples to code on their own, and responses were later compared between raters and the experimenter. Any questions and differences were reconciled through discussion and two more sample response sets were coded by each rater to ensure that both satisfactorily comprehended the coding scheme criteria. Both raters were kept blind to the hypotheses. The measurement of consistency (and corresponding coding) utilized in this dissertation.

\(^{13}\) Interrater reliability analyses were only conducted for the MAP and CQ for the initial coding set. Due to the high reliability of the coding schemes in previous research and in this dissertation, it is inferred that reliability of coding remained the same for the other time intervals. Further, the coders did not change during the duration of this study.
was similar to that used by other memory researchers using longitudinal designs (e.g., Lee & Brown, 2003; Neisser & Harsch, 1992; Peterson, Moores, & White, 2001; Porter & Peace, in press; Qin et al., 2003; Winningham et al., 2000).

*Coding the Qualitative and Self-Reported Features of Memory Reports.*

Coders rated eight qualitative features of the truthful and fabricated narratives based on the Memory Assessment Procedure (MAP; Peace & Porter, 2004; Peace & Porter, 2006; Porter & Birt, 2001; Porter & Peace, in press; Porter, Peace, & Emmett, in press; Porter et al., 1999). The MAP is not a psychological scale but rather a series of arbitrary criteria that provide objective information about the content of autobiographical narratives (see Chapter 3). The criteria address how a person reports a memory in narrative format (which can be objectively coded, such as amount of detail and coherence, references to emotional state, etc.) and how a person subjectively experiences a memory (items such as the memory's vividness sensory components, and stress level) which are rated by participants in a survey questionnaire called the Emotional Memory Survey (EMS).

The objective MAP criteria coded via tallying included: *Amount of detail* (pieces of distinctive information); *Emotional components* (number of references to one's own emotional state and to other's emotional state); *Time and Place Details* (number of references to time and location); and *Number of Words* (see Appendix J). *Coherence* (following a logical sequence from beginning to middle to end), *Relevance* (relevance/importance of the details mentioned to the event being described), and *Plausibility* (the extent to which the information provided is realistic, logical, and contains components that are likely to occur in real-life) were rated on a 7-point scale.
(not at all [1] to extremely [7]). Previous research established a high level of inter-rater reliability for all eight criteria (e.g., Peace & Porter, 2004: \( rs = 0.70-0.99 \)).

Eight research assistants (kept blind to the nature of the study) were trained on the scoring procedures for each of the MAP criteria. After reviewing the detailed scoring procedure with the experimenter, coders had the opportunity to ask any questions for the purposes of clarification. Next, coders were given several sample memory reports (that did not contain any identifying information as to the veracity of the report) to code to assure understanding of the operational definitions. They again met with the experimenter and other coders to review their scoring on MAP criteria as compared to memory reports coded by the author. Any differences in sample coding were reconciled through discussion and feedback. This process was repeated several times using sample memory reports until all coders were proficient in the MAP protocol. Following the training procedure, coders evaluated the MAP characteristics on all truthful and fabricated narratives (blind to the narrative type during coding) provided at Time 1, Time 2, and Time 3. Inter-rater reliability was assessed by having a third-party, blind to the hypotheses of this study and trained on the MAP criteria, re-code each narrative for 20% \( (n = 60) \) of the sample.

Twenty items were included in the EMS (see Appendix K) to address subjective features of the experiences. Items were self-rated using a 7-point scale (e.g., indicate your memory’s level of vividness from 1 (not at all vivid and clear) to 7 (completely vivid and clear). Items pertaining to memory characteristics were grouped together and included vividness/clarity, quality, coherence, sensory details, confidence, and credibility. Items for subjective ratings of emotion also were grouped to examine self-reported distress
(e.g., anxiety/stress, turning off emotions, becoming emotional, detached/distanced, emotional intensity).

In planning the analyses, all items were categorized into objective and subjective groups, which were evaluated separately. Items within the objective and subjective groups were conglomerated for the purpose of analyses to measure the utility of the MAP as a whole in credibility assessment. Previous multivariate analyses of the MAP in this respect have been conducted (e.g., Peace & Porter, 2004; Porter & Birt, 2001; Porter & Peace, in press; Porter, Peace, & Emmett, in press; Porter, Woodworth, et al., 2006).

Revised Impact of Event Scale (IES-R).

The IES-R (Weiss & Marmar, 1997) was used to provide a subjective measurement of the level of traumatic stress associated with both truthful and fabricated traumatic (or stressful) experiences reported by the participants (see Appendix L). This 22-item self-report questionnaire involved rating the frequency of occurrence of various cognitive experiences in relation to a traumatic event (e.g., “I tried not to think about it”; “Any reminder brought back feelings about it”) on a scale marked 0 (not at all), 1 (rarely), 3 (sometimes), and 5 (often). The IES-R contains three subscales that measure the extent to which the traumatic event intrudes in daily life (intrusion), the degree of avoidance of memories and stimuli associated with the event (avoidance), and the level of increased arousal in day-to-day activities (hyperarousal). A total score was generated by combining the subscale scores. This provided an assessment of the overall level of impact of an event, with higher scores representing greater impact. This scale has been extensively used with trauma victims (Foa & Rothbaum, 1998).
Post-Traumatic Stress Disorder Checklist (PCL).

The PCL (Weathers et al., 1994; see Appendix M) is a 17-item self-report screening questionnaire for the presence of PTSD symptomology. Each item corresponds directly to the DSM-IV PTSD criteria of re-experiencing (e.g., “Have you had recurrent distressing dreams about the incident?”), avoidance and numbing (e.g., “Have you tried to avoid thoughts, feelings or conversations associated with the trauma?”), and hyperarousal (e.g., “Have you experienced an exaggerated startle response?”). Participants rated how often they had experienced the symptoms listed over the past month on a scale of 1 (not at all) to 5 (all the time). Dobie et al. (2002) found that the PCL performs well as a screening measure for PTSD (sensitivity 0.79, specificity 0.79) and is widely used in this capacity (e.g., Dobie et al., 2006). It has good psychometric properties, with moderate to high levels of interrater reliability, internal consistency, test-retest reliability, and convergent validity (Mueser et al., 2001).

Trauma Symptom Inventory (TSI).

Developed by Brière (1995), the TSI is a well-validated 100-item self-report measure of symptomology associated with trauma (refer to Appendix N). The TSI consists of two types of measurements: validity scales (atypical response, response level, inconsistent response) and clinical scales (anxious arousal, depression, anger/irritability, intrusive experiences, defensive avoidance, dissociation, sexual concerns, dysfunctional sexual behaviour, impaired self-reference, and tension reduction behaviour). Participants rated each item in relation to how frequently it had occurred in the six months following the event, from 0 (has not happened) to 3 (has happened often). The TSI has good psychometric properties in terms of both reliability and validity (Brière, 1995; Brière &
Elliott, 1997). Here, the TSI provided a comprehensive examination of symptom and validity profiles associated with truthful and fabricated reports of victimization.

**Demographic Questionnaire.**

Participants were asked to complete a brief questionnaire designed by the researcher (see Appendix O), requesting information pertaining to various demographic variables (age, gender, education level) as well as their past history of trauma and beliefs about memory phenomena (e.g., existence of repression/false memories).

**Debriefing Form.**

After completing all phases of the study, participants were given a debriefing form to read over and discuss with the experimenter (see Appendix P). This reviewed the nature and purpose of the study, provided information on traumatic memories, and asked questions to ensure understanding (as per the Psychology Department’s Subject Pool policy for the allotment of experimental credit points).

**Procedure**

**Time One.**

Participants were asked to read and sign an Informed Consent form prior to participation. After consent had been obtained, participants received a set of instructions relating to the reporting of their truthful and fabricated traumatic events. After review of the instructions, participants were requested to provide either a truthful or fabricated (counterbalanced) written account of a traumatic or highly stressful life experience (see Written Narrative Report forms), and complete a series of specific questions pertaining to event details on the Consistency Questionnaire (CQ). Subsequently, participants were
asked to complete several self-report measures with respect to the traumatic event they just provided: the IES-R, the EMS, the PCL, and the TSI. The inventories were completed after each written narrative. In the interim of the provision of truthful and fabricated narratives, participants were given a 10-15 minute break (optional). After written narratives and associated questionnaires were completed for each type of trauma, participants completed a brief demographic questionnaire. The MAP was used to assess the features of each type of narrative, and was completed by the experimenter (and coders) at a later time. The experimental procedure took approximately 1.5-2 hours.

Upon completion of the first phase of the experiment, participants were given the option of obtaining course credit points (2 points) or a cash sum ($5.00) for their involvement in this phase of the research. Participants also were given contact information for the Dalhousie University Counselling Services should any distress be experienced, as well as a follow up appointment card listing the date and time of the next phase of the experiment and the experimenter’s contact information. Participants were reminded that the additional phases of the study would test their memory for the traumatic events provided at Time 1.

*Times Two and Three.*

During the second and third phases of the experiment, participants were re-contacted and appointments were arranged with the researcher. Time 2 occurred approximately three months ($M = 91.44$ days, $SD = 12.77$ days) following the initial experimental session, and Time 3 three months following Time 2 ($M = 72.13$ days, $SD = 16.31$ days). During these sessions, participants completed the same measures and methodology as specified at Time 1, with the exception of the demographic
questionnaire. Upon completion of these phases of the experiment, participants were compensated with course credit points (2 points) or a cash sum ($5.00) for their involvement. The debriefing form was reviewed with participants at the conclusion of the study. Participants again were given contact information for the Dalhousie University Counselling Services should any distress be experienced. In addition, contact information for the research supervisor and primary investigator were provided in the event that the participant would like to discuss any aspect of the research. All experimental sessions were conducted by the author of this dissertation and took place within the Department of Psychology at Dalhousie University.
Chapter 8. Results

Preliminary Analyses

MAP Reliability Check.

The seven components of the MAP (amount of detail, emotional components relating to the self and others, time and place details, coherence, relevant details, and plausibility) were included in this analysis. Bivariate correlations were conducted; all MAP characteristics for both memory types were highly reliable and statistically significant (all ps < .05), with correlations between \( r = .77 \) and .99. Interrater correlations for objective MAP characteristics for truthful narratives were as follows: amount of detail, \( r(60) = .99 \); emotional components relating to self, \( r(60) = .95 \); emotional components relating to others, \( r(60) = .95 \); time and place details, \( r(60) = .99 \); coherence, \( r(60) = .78 \); relevant details, \( r(60) = .84 \); and plausibility, \( r(60) = .95 \). Interrater correlations for fabricated narratives were as follows: amount of detail, \( r(60) = .99 \); emotional components relating to self, \( r(60) = .98 \); emotional components relating to others, \( r(60) = .93 \); time and place details, \( r(60) = .98 \); coherence, \( r(60) = .83 \); relevant details, \( r(60) = .87 \); and plausibility, \( r(60) = .86 \). This range of correlations was acceptable and similar to those found in previous research using this measure (e.g., Peace & Porter, 2004; Peace & Porter, 2006; Porter & Birt, 2001; Porter & Peace, in press; Porter, Peace & Emmett, in press). No significant mean differences between raters were revealed on paired \( t \)-test analyses (ps > .05).

Consistency Reliability Check.

The reliability check indicated that the coded scores for both truthful and fabricated consistency questions were highly reliable. Specifically, results indicated that
the total consistency scores for Rater 1 (original scores) and Rater 2 (reliability check scores) were highly positively correlated for both truthful, $r(60) = .97, p < .001$, and fabricated, $r(60) = .96, p < .001$, narratives of trauma. The mean consistency score for truthful narratives between Times 1 and 2 did not differ according to rater: $41.28 \ (SD = 7.55)$ for the original scores and $39.52 \ (SD = 7.09)$ for the reliability check scores, $t(59) = 1.22, p > .05$. The same was true for the mean consistency score for fabricated narratives (for the T1-T2 interval): $35.78 \ (SD = 6.96)$ for the original scores and $33.69 \ (SD = 7.03)$ for the reliability check scores, $t(59) = 1.46, p > .05$.

*Order of Narrative Reporting.*

To test for any possible order effects of narrative presentation, counterbalancing manipulation was entered into all major multivariate analyses of variance (MANOVAs) reported below as a between subjects factor. No significant main effects or interactions were found (all $ps > .05$), and is excluded from the analyses reported.

*Themes of the Trauma Narratives.*

As expected, participants reported a variety of traumatic experiences which were classified according to the type of event. See Table 2 for the categories of truthful (TT) and fabricated (FT) traumas reported.

*Differences in Real and Fabricated Trauma Narratives*

Features of narratives were assessed using both objective (MAP) and subjective (EMS) scores. The means for objective (e.g., amount of detail, coherence, plausibility) and subjective (e.g., vividness, sensory detail, credibility) features are presented in Tables
Table 2.
Breakdown of categories of truthful (TT) and fabricated (FT) traumatic events reported.

<table>
<thead>
<tr>
<th>Traumatic Event Type</th>
<th>TT</th>
<th>FT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sexual/physical violence</td>
<td>11.8%</td>
<td>24.3%</td>
</tr>
<tr>
<td>Theft/robbery/mugging</td>
<td>4.9%</td>
<td>24.6%</td>
</tr>
<tr>
<td>Serious accident/injury</td>
<td>21.7%</td>
<td>24.6%</td>
</tr>
<tr>
<td>Death of loved one</td>
<td>22.3%</td>
<td>11.1%</td>
</tr>
<tr>
<td>Serious medical condition</td>
<td>9.4%</td>
<td>2.2%</td>
</tr>
<tr>
<td>Relationship related</td>
<td>15.0%</td>
<td>1.8%</td>
</tr>
<tr>
<td>Other (mixed)(^{14})</td>
<td>14.9%</td>
<td>11.4%</td>
</tr>
</tbody>
</table>

3 and 4 (respectively). Two repeated measures mixed MANOVAs were conducted to examine variations in narratives as a function of narrative type and time (independent variables) for both objective (MAP variables) and subjective (EMS variables) characteristics (dependent measures). In addition, self-rated “field of view” in memory (e.g., participant versus observer perspectives) was considered separately from the other analyses as it was a categorical variable. For truthful traumas, participants endorsed the participant perspective 52.6% at T1, 53% at T2, and 50.7% at T3. Similarly, the participant perspective was also endorsed for fabricated events at a rate of 50% for T1, 50.7% for T2, and 53.2% at T3. Chi-squared analyses revealed no significant differences in memory perspective endorsement between narratives or across time ($p < .05$).

\(^{14}\) The “other” category included events such as moving, failure/embarrassment, natural disaster, and getting lost.
Table 3.  
*Means (and standard deviations) for objective MAP features for truthful (TT) and fabricated (FT) narratives over Time 1 (T1), Time 2 (T2), and Time 3 (T3).*

<table>
<thead>
<tr>
<th>MAP Narrative Features</th>
<th>TT</th>
<th></th>
<th></th>
<th>FT</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T1</td>
<td>T2</td>
<td>T3</td>
<td>T1</td>
<td>T2</td>
<td>T3</td>
</tr>
<tr>
<td>Amount of detail</td>
<td>211.59</td>
<td>158.28</td>
<td>154.11</td>
<td>177.97</td>
<td>115.53</td>
<td>110.60</td>
</tr>
<tr>
<td></td>
<td>(129.18)</td>
<td>(91.17)</td>
<td>(93.18)</td>
<td>(112.29)</td>
<td>(73.24)</td>
<td>(70.59)</td>
</tr>
<tr>
<td>Emotional (self)</td>
<td>4.48</td>
<td>3.40</td>
<td>1.92</td>
<td>3.03</td>
<td>2.29</td>
<td>1.21</td>
</tr>
<tr>
<td></td>
<td>(4.87)</td>
<td>(4.78)</td>
<td>(2.46)</td>
<td>(4.56)</td>
<td>(3.23)</td>
<td>(1.93)</td>
</tr>
<tr>
<td>Emotional (other)</td>
<td>1.20</td>
<td>1.23</td>
<td>.73</td>
<td>.83</td>
<td>.69</td>
<td>.54</td>
</tr>
<tr>
<td></td>
<td>(2.34)</td>
<td>(2.32)</td>
<td>(1.20)</td>
<td>(1.58)</td>
<td>(1.24)</td>
<td>(1.02)</td>
</tr>
<tr>
<td>Time &amp; place details</td>
<td>18.72</td>
<td>16.56</td>
<td>14.77</td>
<td>14.88</td>
<td>11.72</td>
<td>10.02</td>
</tr>
<tr>
<td></td>
<td>(12.80)</td>
<td>(11.60)</td>
<td>(9.99)</td>
<td>(10.28)</td>
<td>(7.86)</td>
<td>(6.67)</td>
</tr>
<tr>
<td>Coherence</td>
<td>6.21</td>
<td>5.98</td>
<td>6.14</td>
<td>6.21</td>
<td>5.90</td>
<td>6.11</td>
</tr>
<tr>
<td></td>
<td>(.81)</td>
<td>(1.10)</td>
<td>(.90)</td>
<td>(.82)</td>
<td>(1.12)</td>
<td>(1.16)</td>
</tr>
<tr>
<td>Relevant details</td>
<td>6.02</td>
<td>5.79</td>
<td>5.93</td>
<td>5.91</td>
<td>5.53</td>
<td>5.63</td>
</tr>
<tr>
<td></td>
<td>(.98)</td>
<td>(1.06)</td>
<td>(1.04)</td>
<td>(.99)</td>
<td>(1.08)</td>
<td>(1.22)</td>
</tr>
<tr>
<td>Plausibility</td>
<td>6.07</td>
<td>6.00</td>
<td>6.04</td>
<td>5.46</td>
<td>5.23</td>
<td>5.20</td>
</tr>
<tr>
<td></td>
<td>(.80)</td>
<td>(.94)</td>
<td>(.78)</td>
<td>(1.23)</td>
<td>(1.28)</td>
<td>(1.30)</td>
</tr>
<tr>
<td>Word count</td>
<td>389.72</td>
<td>307.23</td>
<td>289.48</td>
<td>327.70</td>
<td>221.61</td>
<td>205.51</td>
</tr>
<tr>
<td></td>
<td>(217.02)</td>
<td>(179.57)</td>
<td>(172.96)</td>
<td>(191.38)</td>
<td>(138.54)</td>
<td>(133.28)</td>
</tr>
</tbody>
</table>
Table 4. 
Means (and standard deviations) for subjective EMS features for truthful (TT) and fabricated (FT) narratives over Time 1 (T1), Time 2 (T2), and Time 3 (T3).

<table>
<thead>
<tr>
<th>EMS Narrative Features</th>
<th>TT</th>
<th></th>
<th></th>
<th></th>
<th>TT</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T1</td>
<td>T2</td>
<td>T3</td>
<td></td>
<td>T1</td>
<td>T2</td>
<td>T3</td>
<td></td>
</tr>
<tr>
<td>Vividness/clarity</td>
<td>5.52</td>
<td>5.21</td>
<td>5.07</td>
<td></td>
<td>5.37</td>
<td>4.63</td>
<td>4.71</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.46)</td>
<td>(1.43)</td>
<td>(1.50)</td>
<td></td>
<td>(1.50)</td>
<td>(1.57)</td>
<td>(1.72)</td>
<td></td>
</tr>
<tr>
<td>Memory quality</td>
<td>5.48</td>
<td>5.27</td>
<td>5.17</td>
<td></td>
<td>5.36</td>
<td>4.68</td>
<td>4.62</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.60)</td>
<td>(1.47)</td>
<td>(1.48)</td>
<td></td>
<td>(1.68)</td>
<td>(1.67)</td>
<td>(1.69)</td>
<td></td>
</tr>
<tr>
<td>Coherence</td>
<td>5.33</td>
<td>5.16</td>
<td>5.10</td>
<td></td>
<td>5.02</td>
<td>4.53</td>
<td>4.46</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.42)</td>
<td>(1.40)</td>
<td>(1.44)</td>
<td></td>
<td>(1.56)</td>
<td>(1.45)</td>
<td>(1.54)</td>
<td></td>
</tr>
<tr>
<td>Sensory details</td>
<td>4.15</td>
<td>3.96</td>
<td>4.06</td>
<td></td>
<td>4.60</td>
<td>3.95</td>
<td>4.02</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.81)</td>
<td>(1.62)</td>
<td>(1.61)</td>
<td></td>
<td>(1.62)</td>
<td>(1.56)</td>
<td>(1.59)</td>
<td></td>
</tr>
<tr>
<td>Confidence in memory</td>
<td>6.07</td>
<td>5.67</td>
<td>5.30</td>
<td></td>
<td>5.63</td>
<td>4.44</td>
<td>4.53</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.28)</td>
<td>(1.35)</td>
<td>(1.50)</td>
<td></td>
<td>(1.40)</td>
<td>(1.67)</td>
<td>(1.56)</td>
<td></td>
</tr>
<tr>
<td>Credibility</td>
<td>5.93</td>
<td>5.41</td>
<td>5.35</td>
<td></td>
<td>5.41</td>
<td>4.45</td>
<td>4.42</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.32)</td>
<td>(1.43)</td>
<td>(1.45)</td>
<td></td>
<td>(1.40)</td>
<td>(1.65)</td>
<td>(1.55)</td>
<td></td>
</tr>
</tbody>
</table>

Objective Narrative Features.

The 2 (narrative type) X 3 (time) MANOVA on objective narrative (i.e., amount of detail, plausibility) properties revealed a main effect of narrative type, Wilks’ lambda = .49, $F(8,166) = 21.32, p < .001$ ($\eta^2 = .51$), and a main effect of time, Wilks’ lambda = .42, $F(16,158) = 13.46, p < .001$ ($\eta^2 = .58$). The narrative type X time interaction was not
significant, Wilks’ lambda = .87, F(16,159) = 1.53, p > .05 (η² = .13). Overall, truthful narratives were associated with significantly more detail, F(1,173) = 48.54, p < .001 (η² = .22), emotional components relating to the self, F(1,173) = 20.44, p < .001 (η² = .11), and others, F(1,173) = 12.07, p < .01 (η² = .07), time and place details, F(1,173) = 42.95, p < .001 (η² = .20), relevant details, F(1,173) = 11.92, p < .01 (η² = .06), higher levels of plausibility, F(1,173) = 64.37, p < .001 (η² = .27), and a greater number of words, F(1,173) = 53.85, p < .001 (η² = .24), relative to fabricated narratives (see Table 3). Truthful and fabricated narratives did not differ on their level of coherence (p > .05). All objective characteristics also decreased over time for both TTs and FTs (all ps < .05). Pair-wise comparisons indicated that plausibility and relevant details differed at both Times 1 and 2 (p < .05 and p < .01, respectively), and Times 1 and 3 (p < .01 and p < .05, respectively), but not between Times 2 and 3. Levels of coherence differed at both Times 1 and 2 (p < .001), and Times 2 and 3 (p < .01). Emotional components related to others did not differ between Times 1 and 2 (p > .01), but did for both Times 2 and 3 (p < .01) and Times 1 and 3 (p < .005)\textsuperscript{15}. The only significant narrative type X time interaction was for plausibility, F(2,346) = 4.39, p < .05 (η² = .03), suggesting that lower ratings over time were accounted for by the reduction of plausibility in fabricated narratives but not truthful narratives (see Figure 1).

\textsuperscript{15} The remaining MAP objective characteristics differed at all time intervals.
Figure 1. Mean ratings of objective plausibility (1-7) for truthful (TT) and fabricated (FT) trauma narratives at Times 1, 2, and 3 (T1, T2, T3) (including the standard deviation from the mean indicated as vertical lines).

Subjective Narrative Features.

Similar to the above analysis, a 2 (narrative type) X 3 (time) MANOVA on subjective memory features yielded an overall main effect of narrative type, Wilks’ lambda = .64, $F(6,154) = 14.35$, $p < .001$ ($\eta^2 = .36$), a main effect of time, Wilks’ lambda = .55, $F(12,148) = 9.94$, $p < .001$ ($\eta^2 = .45$), and a narrative type X time interaction, Wilks’ lambda = .83, $F(12,148) = 2.6$, $p < .01$ ($\eta^2 = .17$). Truthful narratives were rated as being more vivid, of higher quality, and more coherent, relative to fabricated narratives (all $ps < .001$; the effect sizes were $\eta^2 = .08$, .08, and .13, respectively). Overall levels of confidence in the accuracy of the narrative, $F(1,159) = 26.94$, $p < .001$ ($\eta^2 = .15$), and perceived credibility of the report, $F(1,159) = 51.91$, $p < .001$ ($\eta^2 = .25$), also were significantly higher in truthful than fabricated claims (see Table 4). Univariate analyses
of subjective memory characteristics yielded significant decreases over time (all $p$s < .001): vividness, $F(2,318) = 24.29$ ($\eta^2 = .13$), quality, $F(2,318) = 18.98$ ($\eta^2 = .11$), coherence, $F(2,318) = 15.54$ ($\eta^2 = .09$), sensory details, $F(2,318) = 11.41$ ($\eta^2 = .07$), confidence, $F(2,318) = 12.26$ ($\eta^2 = .07$), and credibility, $F(2,318) = 43.38$ ($\eta^2 = .21$), when collapsed over narrative type. Pair-wise comparisons indicated that all subjective features differed at both Times 1 and 2 ($p$s < .01), and Times 1 and 3 ($p$s < .01), but not between Times 2 and 3. However, there was a narrative type X time interaction for each. This effect indicated that self-reported vividness, $F(2,318) = 5.04$, $p < .01$ ($\eta^2 = .03$), quality, $F(2,318) = 5.26$, $p < .01$ ($\eta^2 = .03$), coherence, $F(2,318) = 4.33$, $p < .05$ ($\eta^2 = .03$), and confidence, $F(2,318) = 3.31$, $p < .05$ ($\eta^2 = .02$), decreased in a small (but significant) linear fashion for truthful traumas across the testing times relative to fabricated narratives (which decreased to a greater degree from T1 to T2 and remained relatively stable between T2 and T3). Sensory details changed little over time for truthful experiences, and were generally equivalent to fabricated traumas at T2 and T3 (with FT being elevated at T1 more so than TT), $F(2,318) = 6.05$, $p < .01$ ($\eta^2 = .04$). Self-ratings of how credible participants thought their reports would be to others decreased more between the T1-T2 intervals and were stable between T2-T3 for TTs, $F(2,318) = 3.65$, $p < .05$ ($\eta^2 = .02$), relative to FTs (refer to Figures 2a-2f).
Figure 2a. Mean ratings of self-reported vividness/clarity (1-7) for truthful (TT) and fabricated (FT) trauma narratives at Times 1, 2, and 3 (T1, T2, T3) (including the standard deviation from the mean indicated as vertical lines).

Figure 2b. Mean ratings of self-reported memory quality (1-7) for truthful (TT) and fabricated (FT) trauma narratives at Times 1, 2, and 3 (T1, T2, T3) (including the standard deviation from the mean indicated as vertical lines).
Figure 2c. Mean ratings of self-reported coherence (1-7) for truthful (TT) and fabricated (FT) trauma narratives at Times 1, 2, and 3 (T1, T2, T3) (including the standard deviation from the mean indicated as vertical lines).

Figure 2d. Mean ratings of self-reported sensory details (1-7) for truthful (TT) and fabricated (FT) trauma narratives at Times 1, 2, and 3 (T1, T2, T3) (including the standard deviation from the mean indicated as vertical lines).
Figure 2e. Mean ratings of self-reported confidence in memory accuracy (1-7) for truthful (TT) and fabricated (FT) trauma narratives at Times 1, 2, and 3 (T1, T2, T3) (including the standard deviation from the mean indicated as vertical lines).

Figure 2f. Mean ratings of self-reported credibility (1-7) for truthful (TT) and fabricated (FT) trauma narratives at Times 1, 2, and 3 (T1, T2, T3) (including the standard deviation from the mean indicated as vertical lines).
Objective and Subjective Features as Predictors of Narrative Type.

To determine whether objective and subjective features of narratives can reliably discriminate between truthful and fabricated narratives, as well as if predictors change over time, discriminant function analyses (DFA) were conducted. All narrative features (both objective and subjective)\textsuperscript{16} were entered simultaneously into the DFA model as independent measures at each of Times 1, 2, and 3, with narrative type as the grouping variable. The structure matrix provided by this analysis allowed for predictors in the model to be identified, with coefficients at the level of .30 or greater being considered meaningful predictors (see Tabachnick & Fidell, 2001, for discussion of multivariate predictors and standardized cut-off scores).

At Time 1, the DFA analysis yielded a significant result, Wilks' lambda = .82, $\chi^2_{(14)} = 108.82$, $p < .001$. Combined objective and subjective features correctly classified truthful and fabricated narratives at rates of 70.1% and 63.8%, respectively, for an overall classification accuracy of 67%. Based on an evaluation of the structure matrix, the strongest discriminators between narrative types were: plausibility (.65), self-rated credibility (.41), confidence (.36), time and place details (.35), emotional components relating to self (.32), and word count (.30)\textsuperscript{17}. At Time 2, the discriminant function again was significant, Wilks' lambda = .76, $\chi^2_{(14)} = 135.15$, $p < .001$. The overall level of discrimination at T2 was 72.4%, with 74.1% of truthful and 70.6% of fabricated

\textsuperscript{16} These features were the objective MAP characteristics (amount of detail, emotional components relating to self and others, time and place details, coherence, relevance, plausibility, word count) and subjective EMS variables (vividness, quality, self-rated coherence, sensory details, confidence in memory accuracy, and perceived credibility).

\textsuperscript{17} At T1, sensory details (-.29), amount of detail (.27), self-rated coherence (.21), emotional components relating to others (.20), relevance (.11), vividness (.11), quality (.08), and objectively-rated coherence (.003) did not discriminate meaningfully between narrative types.
narratives being correctly classified. This model was more sensitive to narrative discrimination and the structure matrix revealed a greater number of meaningful predictors: plausibility (.59), self-rated credibility (.55), word count (.48), amount of detail (.47), time and place details (.43), self-rated coherence (.41), vividness (.35), memory quality (.34), and confidence (.33)\textsuperscript{18}. Discriminant function analysis at Time 3 also yielded a significant result, Wilks' lambda = .71, $\chi^2_{(14)} = 115.87, p < .001$. Objective and subjective features continued to correctly classify narratives at an overall level of 72.3\%, with a high discrimination of truthful (74.7\%) and fabricated (69.9\%) traumas at T3. According to the structure matrix, meaningful predictors were: plausibility (.62), self-rated credibility (.49), time and place details (.45), word count (.44), amount of detail (.42), confidence (.41), and self-rated coherence (.34)\textsuperscript{19}.

Consistency of Real and Fabricated Traumas over Time

The consistency of truthful and fabricated narratives over time was evaluated using total and item scores from the Consistency Questionnaire (CQ). Mean scores are reported for comparisons between Times 1 and 2 (T1-T2), Times 2 and 3 (T2-T3) in Table 5. The mean level of consistency (derived from summing the scores from the 23 item categories) over the extended Time 1 to 3 (T1-T3) interval was 41.23 ($SD = 7.83$) for truthful narratives, relative to 35.02 ($SD = 6.81$) for fabricated narratives (see Table 6). Mean item scores ranged from 0 (inconsistent) to 3 (consistent).

\textsuperscript{18} At T2, emotional components relating to self (.26) and others (.25), relevance (.22), objectively-rated coherence (.07), and sensory details (.01) were not meaningful variables for discrimination.

\textsuperscript{19} Emotional components relating to self (.25) and others (.12), relevance (.23), vividness (.18), objective-rated coherence (.03) and sensory details (.01) did not meaningfully discriminate between narrative types at Time 3.
Table 5.  
*Mean item and total scores (and standard deviations) for truthful (TT) and fabricated (FT) traumas across Times 1 and 2 (T1-T2) and Times 2 and 3 (T2-T3) on the CQ.*

<table>
<thead>
<tr>
<th>CQ Item</th>
<th>T1-T2</th>
<th>T2-T3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TT (SD)</td>
<td>FT (SD)</td>
</tr>
<tr>
<td>Clothes</td>
<td>2.04 (1.05)</td>
<td>1.58 (1.02)</td>
</tr>
<tr>
<td>Weather</td>
<td>1.76 (1.05)</td>
<td>1.42 (1.08)</td>
</tr>
<tr>
<td>Location</td>
<td>1.99 (.99)</td>
<td>1.71 (1.08)</td>
</tr>
<tr>
<td>Time of day</td>
<td>1.70 (1.12)</td>
<td>1.53 (1.15)</td>
</tr>
<tr>
<td>Emotions before</td>
<td>1.39 (1.20)</td>
<td>1.39 (1.20)</td>
</tr>
<tr>
<td>Number people</td>
<td>1.52 (1.37)</td>
<td>1.65 (1.41)</td>
</tr>
<tr>
<td>Day of week</td>
<td>2.07 (1.12)</td>
<td>1.75 (1.30)</td>
</tr>
<tr>
<td>Duration</td>
<td>1.31 (1.19)</td>
<td>1.20 (1.15)</td>
</tr>
<tr>
<td>Personal items</td>
<td>2.06 (1.18)</td>
<td>2.04 (1.05)</td>
</tr>
<tr>
<td>Emotions during</td>
<td>1.49 (1.13)</td>
<td>1.60 (1.10)</td>
</tr>
<tr>
<td>Calendar date</td>
<td>1.89 (.97)</td>
<td>1.13 (1.02)</td>
</tr>
<tr>
<td>Smells</td>
<td>2.06 (1.22)</td>
<td>1.52 (1.31)</td>
</tr>
<tr>
<td>Visual details</td>
<td>1.28 (1.22)</td>
<td>1.04 (1.18)</td>
</tr>
<tr>
<td>Landmarks</td>
<td>1.24 (1.27)</td>
<td>1.24 (1.26)</td>
</tr>
<tr>
<td>Emotions after</td>
<td>1.32 (1.13)</td>
<td>1.25 (1.14)</td>
</tr>
<tr>
<td>Weapon</td>
<td>2.76 (.78)</td>
<td>2.54 (1.01)</td>
</tr>
<tr>
<td>Noises</td>
<td>1.53 (1.26)</td>
<td>1.20 (1.25)</td>
</tr>
<tr>
<td>Sensations</td>
<td>1.48 (1.19)</td>
<td>1.20 (1.19)</td>
</tr>
<tr>
<td>Others involved</td>
<td>2.28 (1.09)</td>
<td>2.36 (1.08)</td>
</tr>
</tbody>
</table>

154
<table>
<thead>
<tr>
<th>CQ Item</th>
<th>T1-T3 (Overall Level of Consistency)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TT</td>
</tr>
<tr>
<td>1. Clothes</td>
<td>2.12 (1.04)</td>
</tr>
<tr>
<td>2. Weather</td>
<td>1.70 (1.04)</td>
</tr>
<tr>
<td>3. Location</td>
<td>1.98 (.98)</td>
</tr>
<tr>
<td>4. Time of day</td>
<td>1.83 (1.14)</td>
</tr>
<tr>
<td>5. Emotions before</td>
<td>1.45 (1.23)</td>
</tr>
<tr>
<td>6. Number people</td>
<td>1.46 (1.23)</td>
</tr>
<tr>
<td>7. Day of week</td>
<td>2.06 (1.14)</td>
</tr>
<tr>
<td>8. Duration</td>
<td>1.21 (1.21)</td>
</tr>
<tr>
<td>9. Personal items</td>
<td>2.09 (1.12)</td>
</tr>
<tr>
<td>10. Emotions during</td>
<td>1.43 (1.11)</td>
</tr>
<tr>
<td>11. Calendar date</td>
<td>1.94 (1.00)</td>
</tr>
<tr>
<td>12. Smells</td>
<td>1.90 (1.23)</td>
</tr>
</tbody>
</table>

Table 6. 
Mean item and total scores (and standard deviations) for truthful (TT) and fabricated (FT) traumas over the extended Time 1 to 3 (T1-T3) interval on the CQ.
<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>13. Visual details</td>
<td>1.30 (1.21)</td>
<td>1.22 (1.19)</td>
</tr>
<tr>
<td>14. Landmarks</td>
<td>1.03 (1.21)</td>
<td>1.25 (1.30)</td>
</tr>
<tr>
<td>15. Emotions after</td>
<td>1.27 (1.12)</td>
<td>1.15 (1.08)</td>
</tr>
<tr>
<td>16. Weapon</td>
<td>2.49 (1.07)</td>
<td>2.44 (1.08)</td>
</tr>
<tr>
<td>17. Noises</td>
<td>1.45 (1.26)</td>
<td>1.17 (1.21)</td>
</tr>
<tr>
<td>18. Sensations</td>
<td>1.49 (1.19)</td>
<td>1.25 (1.20)</td>
</tr>
<tr>
<td>19. Others involved</td>
<td>2.34 (1.02)</td>
<td>2.26 (1.16)</td>
</tr>
<tr>
<td>20. Age at time</td>
<td>2.50 (1.07)</td>
<td>1.60 (1.45)</td>
</tr>
<tr>
<td>21. Actions before</td>
<td>1.54 (1.32)</td>
<td>1.59 (1.30)</td>
</tr>
<tr>
<td>22. Physical injury</td>
<td>2.40 (1.07)</td>
<td>2.06 (1.16)</td>
</tr>
<tr>
<td>23. Actions after</td>
<td>1.49 (1.17)</td>
<td>1.60 (1.19)</td>
</tr>
</tbody>
</table>

**Mean Consistency**

41.23 (7.83) 35.02 (6.81)

*Consistency Scores.*

To investigate the consistency of truthful and fabricated narratives of trauma, a 2
narrative type) X 3 (time) MANOVA with total consistency scores as the dependent
measure was conducted. There were main effects of narrative type, Wilks’ lambda = .67,
\[ F(1,96) = 46.36, p < .001 \ (\eta^2 = .33) \], and time, Wilks’ lambda = .41, \[ F(2,95) = 67.35, p < .001 \ (\eta^2 = .59) \], as well as a narrative type X time interaction, Wilks’ lambda = .91,
\[ F(2,95) = 4.52, p < .05 \ (\eta^2 = .09) \]. Overall, TTs \( M = 43.15, SE = .64 \) were more
consistent than FTs \( M = 38.17, SE = .60 \). Pair-wise comparisons indicated that scores
( across truthful and fabricated narratives) differed at each of the time intervals: T1-T2 \( p < .001 \), \( M = 39.35 \); T2-T3 \( p < .001 \), \( M = 44.56 \); T1-T3 \( p < .01 \), \( M = 38.08 \). For
truthful narratives, the scores at each time were: T1-T2 ($M = 42.08$, $SD = 7.50$); T2-T3 ($M = 46.43$, $SD = 6.72$); T1-T3 ($M = 41.23$, $SD = 7.83$). For fabricated narratives, consistency scores for each time were: T1-T2 ($M = 36.64$, $SD = 6.64$); T2-T3 ($M = 42.71$, $SD = 8.00$); T1-T3 ($M = 35.02$, $SD = 6.81$). The interaction indicated that these findings resulted from increases in consistency scores between T2-T3 relative to both T1-T2 and T1-T3 for both narrative types, with a slightly greater increase for FTs (see Figure 3a).

**Figure 3a.** Total consistency scores of truthful (TT) and fabricated (FT) trauma narratives over intervals between Times 1, 2, and 3 (T1-T2; T2-T3; T1-T3) (including the standard deviation from the mean indicated as vertical lines).

An additional MANOVA comparing consistency scores for both narrative types across T1-T2 and T1-T3 intervals indicated that both memory types decreased over time, Wilks’ lambda = .93, $F(1,96) = 7.01$, $p < .01$ ($\eta^2 = .07$), and that TTs were more
consistent overall relative to FTs, Wilks’ lambda = .67, $F(1,96) = 47.27, p < .001$ ($\eta^2 = .33$) (see Figure 3b). The memory type X time interaction was not significant.

*Figure 3b.* Total consistency scores of truthful (TT) and fabricated (FT) trauma narratives over the entire follow-up period between Times 1 and 3 (T1-T2; T1-T3) (including the standard deviation from the mean indicated as vertical lines).

To examine these differences further, mean scores for each of the 23 consistency categories were examined (see Tables 5 and 6) as a function of time and narrative type (using the initial and extended interval). The MANOVA yielded a main effect of narrative type, Wilks’ lambda = .33, $F(23,74) = 6.67, p < .001$ ($\eta^2 = .67$). There was no main effect of time, Wilks’ lambda = .73, $F(23,74) = 1.17, p > .05$ ($\eta^2 = .27$), or interaction effects, Wilks’ lambda = .78, $F(23,74) = .92, p > .05$ ($\eta^2 = .22$). Memory differences were revealed for the following consistency categories: clothes, $F(1,96) = 19.38, p < .001$ ($\eta^2 = .17$), time of day, $F(1,96) = 7.59, p < .01$ ($\eta^2 = .07$), day of the
week, \( F(1, 96) = 6.49, p < .05 (\eta^2 = .06) \), calendar date, \( F(1, 96) = 60.89, p < .001 (\eta^2 = .39) \), smells, \( F(1, 96) = 11.8, p < .01 (\eta^2 = .11) \), sensations, \( F(1, 96) = 9.78, p < .01 (\eta^2 = .09) \), age at time of event, \( F(1, 96) = 21.04, p < .001 (\eta^2 = .18) \), and physical injury, \( F(1, 96) = 16.34, p < .001 (\eta^2 = .15) \). The mean level of consistency was higher in TTs relative to FTs for each of these categories. Consistency scores for truthful narratives for T1-T2 were strongly correlated with both T2-T3, \( r(124) = .57, p < .001 \), and T1-T3 intervals, \( r(123) = .71, p < .001 \). Scores for T2-T3 and T1-T3 also were correlated, \( r(124) = .54, p < .001 \). Consistency scores for fabricated narratives also were highly positively correlated across intervals: T1-T2 v. T2-T3, \( r(117) = .45, p < .001 \); T2-T3 v. T1-T3, \( r(116) = .58, p < .001 \); T1-T2 v. T1-T3, \( r(116) = .61, p < .001 \). These data indicate that consistency was a relatively static characteristic of individual participants over time (within narrative types).

*Consistency as a Predictor of Narrative Type.*

To determine whether levels of consistency discriminated narratives in terms of their veracity, discriminant function analyses (DFAs) were used with the same cut-offs as identified above. This analysis, using total consistency scores for T1-T2, T2-T3, and T1-T3, was significant, Wilks' lambda = .85, \( \chi^2(3) = 37.06, p < .001 \). Truthful and fabricated narratives were discriminated based on total consistency scores at classification rates of 68.3% and 71.6%, respectively. The overall model correctly classified narrative types 69.9% of the time. Evaluation of the structural matrix indicated that consistency scores over the T1-T3 interval was the most meaningful predictor (.95), followed by T1-T2 scores (.86) and T2-T3 scores (.58). To determine which elements of consistency enabled the strongest degree of discrimination, consistency item scores (the 23 CQ questions)
were entered into the model for T1-T2 and T1-T3 (separately). The discriminant function for consistency items at T1-T2 was significant, Wilks’ lambda = .65, $\chi^2_{(23)} = 148.75, p < .001$. There were 5 consistency items that meaningfully predicted differences between truthful and fabricated narratives, with higher levels of consistency associated with truthful traumas relative to fabricated (see Table 5): calendar date (.65), physical injury (.45), age at time of trauma (.42), clothing worn at time of event (.31), and olfactory sensations (smells) associated with the event (.31). The remainder of the consistency items did not have high discriminative power (all items < .22). The overall model was able to correctly classify 75.7% of cases (76.5% truthful, 74.9% fabricated). Evaluation of consistency items after the extended time interval between T1-T3, the DFA model again was significant, Wilks’ lambda = .64, $\chi^2_{(23)} = 117.47, p < .001$. Analysis of the structural matrix indicated that the calendar date of the event (.58), age when it happened (.48), and the clothing worn at the time of the event (.37) remained meaningful discriminants\(^{20}\). Time of day (.30) and day of the week (.28) did not quite reach the cut-off. Each was more consistently recalled over time in truthful, relative to fabricated, narratives of trauma (see Table 6). In addition, these items continued to correctly classify truthful (78.7%) and fabricated (72.6%) narratives at an overall rate of 75.7%.

*Levels of Trauma Symptomology as a Function of Narrative Type*

Scores on each of the measures of trauma symptomology were evaluated to determine whether differences between truthful (TT) and fabricated (FT) symptom

\(^{20}\) The remaining 18 consistency items fell under the recommended .30 cut-off for meaningful discrimination (< .24).
reports were evident. Levels of self-reported emotional and cognitive characteristics also were evaluated.

Subjective Emotional and Cognitive Characteristics.

Participants self-reported various subjective emotional and cognitive features associated with truthful and fabricated traumas, including anxiety/stress, feelings of detachment, “turned off emotions”, emotional intensity, and frequency of thinking and talking about the event. The means (and standard deviations) for these characteristics (as measured by the EMS) are presented in Table 7. To compare these features of truthful and fabricated narratives of trauma, a 2 (narrative type) X 3 (time) MANOVA was conducted on subjective emotional and cognitive ratings as dependent measures. There was a main effect of time, Wilks’ lambda = .52, $F(14,144) = 9.52, p < .001 (\eta^2 = .48)$, and no main effect of narrative type, Wilks’ lambda = .93, $F(7,151) = 1.59, p > .05 (\eta^2 = .07)$, or an interaction, Wilks’ lambda = .92, $F(14,144) = .84, p > .05 (\eta^2 = .08)$. Overall, both narrative types decreased in self-reported anxiety/stress, $F(2,314) = 19.53, p < .001 (\eta^2 = .11)$, thoughts about the event, $F(2,314) = 33.63, p < .001 (\eta^2 = .18)$, turning off emotions, $F(2,314) = 20.61, p < .001 (\eta^2 = .12)$, becoming emotional, $F(2,314) = 21.11, p < .001 (\eta^2 = .12)$, and emotional intensity, $F(2,314) = 37.31, p < .001 (\eta^2 = .19)$ over time. Pair-wise comparisons indicated that anxiety ratings differed at both Times 1 and 2 ($p < .001$) and Times 1 and 3 ($p < .001$), but not between Time 2 and 3 ($p > .05$). Overall ratings of thinking about the event, turning off emotions, becoming emotional, and emotional intensity also differed at both Times 1 and 2 (all $ps < .001$), and Times 1 and 3 (all $ps < .001$), but not between Times 2 and 3 (all $ps > .05$).
Table 7.
Means (and standard deviations) for subjective EMS emotional features for truthful (TT) and fabricated (FT) traumas over Time 1 (T1), Time 2 (T2), and Time 3 (T3).

<table>
<thead>
<tr>
<th>EMS Emotional Features</th>
<th>TT</th>
<th>FT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T1</td>
<td>T2</td>
</tr>
<tr>
<td>Anxiety/stress</td>
<td>6.04</td>
<td>5.78</td>
</tr>
<tr>
<td></td>
<td>(1.16)</td>
<td>(1.32)</td>
</tr>
<tr>
<td>Talking about event</td>
<td>4.38</td>
<td>4.31</td>
</tr>
<tr>
<td></td>
<td>(1.92)</td>
<td>(1.82)</td>
</tr>
<tr>
<td>Thinking about event</td>
<td>5.62</td>
<td>5.22</td>
</tr>
<tr>
<td></td>
<td>(1.57)</td>
<td>(1.65)</td>
</tr>
<tr>
<td>Turned off emotions</td>
<td>4.13</td>
<td>3.76</td>
</tr>
<tr>
<td></td>
<td>(1.89)</td>
<td>(1.88)</td>
</tr>
<tr>
<td>Becoming emotional</td>
<td>4.48</td>
<td>4.23</td>
</tr>
<tr>
<td></td>
<td>(1.86)</td>
<td>(1.87)</td>
</tr>
<tr>
<td>Detached/distanced</td>
<td>3.55</td>
<td>3.22</td>
</tr>
<tr>
<td></td>
<td>(1.89)</td>
<td>(1.70)</td>
</tr>
<tr>
<td>Emotional intensity</td>
<td>5.21</td>
<td>4.77</td>
</tr>
<tr>
<td></td>
<td>(1.62)</td>
<td>(1.57)</td>
</tr>
</tbody>
</table>

IES-R.

The means for the total IES-R scores for truthful traumas were 46.55 ($SD = 16.75$), 40.71 ($SD = 18.17$), and 38.25 ($SD = 18.77$) at each of T1, T2, and T3, respectively. Fabricated traumas were associated with mean IES-R total scores at T1, T2,
and T3 of 54.13 (SD = 16.17), 47.06 (SD = 19.28), and 44.96 (SD = 20.91) respectively (See Table 8 for subscale scores over time, and the proportions of scores classified as mild, moderate, or severe). Fabricated trauma symptoms were classified as severe for 61.1% of participants at T1, 46.5% at T2, and 44.3% at T3. Severe levels of traumatic stress for truthful traumas at T1 (45.3%), T2 (30.3%), and T3 (28.5%) were significantly lower than those reported for fabricated traumas; T1: $\chi^2 = 16.26, p < .01$; T2: $\chi^2 = 52.33, p < .001$; T3: $\chi^2 = 48.77, p < .001$.

A 2 (narrative type) X 3 (testing time) repeated measures MANOVA was conducted with IES total and subscale scores as the dependent measures. The overall IES-R total scores were 42.13 (SE = 1.51) for truthful trauma symptoms and 47.84 (SE = 1.49) for fabricated symptoms. The MANOVA indicated main effects of narrative type, Wilks’ lambda = .65, $F(3,122) = 21.76, p < .001$ ($\eta^2 = .35$), and time, Wilks’ lambda = .59, $F(6,219) = 13.77, p < .001$ ($\eta^2 = .41$). There was no narrative type X time interaction, Wilks’ lambda = .95, $F(6,219) = .96, p > .05$ ($\eta^2 = .05$). Fabricated traumas were associated with higher levels of overall traumatic stress$^{21}$, $F(1,124) = 16.27, p < .001$ ($\eta^2 = .12$); avoidance, $M = 15.81, SE = .54, F(1,124) = 5.29, p < .05$ ($\eta^2 = .04$); and hyperarousal, $M = 14.72, SE = .54, F(1,124) = 42.65, p < .001$ ($\eta^2 = .26$); relative to truthful traumas, $M = 42.13 (SE = 1.51), 14.51 (SE = .57)$, and 11.29 (SE = .55), respectively. In addition, scores for IES-R total, $F(2,248) = 51.44, p < .001$ ($\eta^2 = .29$), intrusion, $F(2,248) = 48.76, p < .001$ ($\eta^2 = .28$), avoidance, $F(2,248) = 27.3, p < .001$ ($\eta^2 = .18$), and hyperarousal, $F(2,248) = 38.06, p < .001$ ($\eta^2 = .24$) decreased over time similarly for both types of trauma experiences (no interaction effect). Pair-wise

---

$^{21}$ The IES total score grand mean is reported above for both truthful and fabricated symptom reports.
Table 8.
*Means, standard deviations, and severity category for truthful (TT) and fabricated (FT) traumatic events over Times 1, 2, and 3 (T1, T2, T3) for IES-R scores.*

<table>
<thead>
<tr>
<th>IES-R Scores</th>
<th>TT</th>
<th>FT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T1</td>
<td>T2</td>
</tr>
<tr>
<td></td>
<td>n=274</td>
<td>n=246</td>
</tr>
<tr>
<td>Total Score</td>
<td>46.55</td>
<td>40.71</td>
</tr>
<tr>
<td>Avoidance Subscale</td>
<td>15.56</td>
<td>13.75</td>
</tr>
<tr>
<td>(6.48) (6.83) (7.46)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intrusion Subscale</td>
<td>18.16</td>
<td>15.87</td>
</tr>
<tr>
<td>(7.01) (6.70) (7.00)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hyperarousal Subscale</td>
<td>12.70</td>
<td>11.06</td>
</tr>
<tr>
<td>(6.64) (6.78) (6.79)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Severity Category</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mild</td>
<td>13.2%</td>
<td>22.2%</td>
</tr>
<tr>
<td>Moderate</td>
<td>41.5%</td>
<td>47.4%</td>
</tr>
<tr>
<td>Severe</td>
<td>45.3%</td>
<td>30.3%</td>
</tr>
</tbody>
</table>

Comparisons indicated that total and subscale scores differed from each other at each testing time (all ps < .01).

*PCL.*

The PCL also measured levels of intrusion, avoidance, and hyperarousal, and provided an overall assessment of PTSD. Total and subscale scores are presented in
Table 9. Truthful trauma was associated with mean total scores of 40.39 (SD = 13.45) at T1, 37.89 (SD = 13.87) at T2, and 36.63 (SD = 13.14) at T3. The mean level of fabricated symptomatology on the PCL was 50.31 (SD = 14.58), 44.83 (SD = 14.92), and 44.80 (SD = 15.80) for each T1, T2, and T3, respectively. Further, at T1, 37% of participants met PTSD criteria with respect to their truthful trauma, relative to 61.9% of fabricated claims, \(\chi^2 = 10.57, p < .01\). This pattern continued at T2, \(\chi^2 = 5.35, p < .05\), and T3, \(\chi^2 = 11.49, p < .01\); fabricated symptoms met PTSD criteria more often than symptoms associated with truthful trauma.

A 2 (narrative type) X 3 (time) repeated measures MANOVA was conducted with PCL total and subscale scores as the dependent measures. The MANOVA indicated main effects of narrative type, Wilks’ lambda = .71, \(F(3,126) = 17.38, p < .001\) (\(\eta^2 = .29\)), time, Wilks’ lambda = .63, \(F(6,123) = 11.91, p < .001\) (\(\eta^2 = .37\)), and a narrative type X time interaction, Wilks’ lambda = .88, \(F(6,123) = 2.7, p < .05\) (\(\eta^2 = .12\)). Overall, fabricated symptoms were associated with higher PCL total, \(F(1,128) = 37.9, p < .001\) (\(\eta^2 = .23\)), and subscale scores on intrusion, \(F(1,128) = 36.56, p < .001\) (\(\eta^2 = .22\)), avoidance, \(F(1,128) = 15.95, p < .001\) (\(\eta^2 = .11\)), and hyperarousal, \(F(1,128) = 50.8, p < .001\) (\(\eta^2 = .28\)), relative to truthful symptoms. Each of these scales also decreased, in general, over time (all ps < .001) when collapsed across memory type. Pair-wise comparisons indicated that PCL total scores as well as the intrusion, avoidance, and hyperarousal subscales differed at both Times 1 and 2 (all ps < .001), and Times 1 and 3 (ps < .01), but not between Times 2 and 3 (ps > .05). However, there was a narrative type X time interaction for PCL total scores, \(F(2,256) = 7.53, p < .01\) (\(\eta^2 = .06\)), and the avoidance, \(F(2,256) = 7.67, p < .01\) (\(\eta^2 = .06\)), and hyperarousal subscales, \(F(2,256) = 4.71, p < .05\) (\(\eta^2 = .04\)).
Table 9.  
Means, standard deviations, and percentages of possible PTSD for truthful (TT) and fabricated (FT) traumatic events over Times 1, 2, and 3 (T1, T2, T3) for scores on the PCL.

<table>
<thead>
<tr>
<th>PCL Scores</th>
<th>TT (Mean, SD)</th>
<th>FT (Mean, SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T1</td>
<td>T2</td>
</tr>
<tr>
<td>n=290</td>
<td>n=250</td>
<td>n=179</td>
</tr>
<tr>
<td>PCL Total Score</td>
<td>40.39</td>
<td>37.89</td>
</tr>
<tr>
<td>Intrusion Subscale</td>
<td>12.89</td>
<td>11.57</td>
</tr>
<tr>
<td></td>
<td>(4.81)</td>
<td>(4.62)</td>
</tr>
<tr>
<td>Avoidance Subscale</td>
<td>15.60</td>
<td>14.86</td>
</tr>
<tr>
<td></td>
<td>(5.02)</td>
<td>(5.03)</td>
</tr>
<tr>
<td>% Met PTSD Criteria</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YES</td>
<td>37.0%</td>
<td>31.2%</td>
</tr>
<tr>
<td>NO</td>
<td>63.0%</td>
<td>68.8%</td>
</tr>
</tbody>
</table>

This interaction indicates that lower ratings over time were accounted for by the deterioration of symptoms associated with truthful traumas to a greater extent (overall) than for fabricated symptoms (see Figures 4a-c)\(^{22}\).

\(^{22}\) Discriminant function analyses were not performed for IES-R or PCL scores due to their conceptual overlap, as well as the ability of the multivariate analyses to adequately address total and subscale scores as a function of narrative type.
Figure 4a. Mean PCL total scores for truthful (TT) and fabricated (FT) trauma symptoms at Times 1, 2, and 3 (T1, T2, T3) (including the standard deviation from the mean indicated as vertical lines).

Figure 4b. Mean PCL avoidance scores for truthful (TT) and fabricated (FT) trauma symptoms at Times 1, 2, and 3 (T1, T2, T3) (including the standard deviation from the mean indicated as vertical lines).
Figure 4c. Mean PCL hyperarousal scores for truthful (TT) and fabricated (FT) trauma symptoms at Times 1, 2, and 3 (T1, T2, T3) (including the standard deviation from the mean indicated as vertical lines).

TSI.

The means (and standard deviations) on the clinical and validity scales of the TSI for truthful and fabricated traumas are presented in Tables 10 and 11, respectively. Overall, the majority of TSI profiles were valid at each time period for both truthful (T1: 90.1%, T2: 94.4%, T3: 88.5%) and fabricated (T1: 84.8%, T2: 87.3%, T3: 84.4%) symptom reports.

A 2 (narrative type) X 3 (testing time) repeated measures MANOVA was conducted with the TSI clinical scales as the dependent measures\(^{23}\) (see Table 10). There

\(^{23}\) These scales were anxious arousal (AA), depression (D), anger/irritability (AI), intrusive experiences (IE), defensive avoidance (DA), dissociation (DIS), sexual concerns (SC), deviant sexual behaviour (DSB), impaired self-reference (ISR), and tension reduction behaviour (TRB).
were main effects of narrative type, Wilks’ lambda = .66, $F(10,149) = 7.52, p < .001$ ($\eta^2 = .34$), and time, Wilks’ lambda = .58, $F(20,139) = 5.02, p < .001$ ($\eta^2 = .42$), but no significant overall interaction between these two factors, Wilks’ lambda = .82, $F(20,139) = 1.48, p = .09$ ($\eta^2 = .18$). Overall, fabricated levels of symptomology (relative to truthful trauma symptoms) were higher for the following clinical scales: anxious arousal, $F(1,158) = 60.27, p < .001$ ($\eta^2 = .28$), anger/irritability, $F(1,158) = 9.45, p < .01$ ($\eta^2 = .06$), intrusive experiences, $F(1,158) = 46.75, p < .001$ ($\eta^2 = .23$), defensive avoidance, $F(1,158) = 20.23, p < .001$ ($\eta^2 = .11$), dissociation, $F(1,158) = 18.99, p < .001$ ($\eta^2 = .11$), sexual concerns, $F(1,158) = 6.45, p < .05$ ($\eta^2 = .04$), impaired self-reference, $F(1,158) = 6.87, p < .05$ ($\eta^2 = .04$), and tension reduction behaviour, $F(1,158) = 21.15, p < .001$ ($\eta^2 = .12$). In addition, all clinical subscale scores of the TSI decreased over time. Pair-wise comparisons indicated that each clinical scale differed at both Times 1 and 2 ($ps < .01$), and Times 1 and 3 ($ps < .01$), but not between Times 2 and 3 ($ps > .05$). The only exception to this was the dissociation scale, which decreased significantly at all three testing times (T1-T2: $p < .001$; T2-T3: $p < .05$; T1-T3: $p < .001$).

This analysis was repeated for the validity scales of the TSI (atypical responding, response level, and inconsistent response). The MANOVA yielded main effects of time, Wilks’ lambda = .82, $F(6,153) = 5.71, p < .001$ ($\eta^2 = .18$), and narrative type, Wilks’ lambda = .79, $F(3,156) = 14.03, p < .001$ ($\eta^2 = .21$), with no interaction ($p > .05$). Overall, atypical responding (ATR) differed significantly between truthful and fabricated symptom profiles, $F(1,158) = 36.90, p < .001$ ($\eta^2 = .19$). Specifically, fabricated profiles had higher ATR scores ($M = 57.26, SE = .98$) relative to truthful symptoms.
<table>
<thead>
<tr>
<th>TSI Clinical Scales</th>
<th>TT</th>
<th>FT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T1</td>
<td>T2</td>
</tr>
<tr>
<td></td>
<td>n=282</td>
<td>n=250</td>
</tr>
<tr>
<td>Anxious Arousal (AA)</td>
<td>55.91</td>
<td>52.56</td>
</tr>
<tr>
<td>(10.06)</td>
<td>(9.00)</td>
<td>(9.37)</td>
</tr>
<tr>
<td>Depression (D)</td>
<td>55.63</td>
<td>53.30</td>
</tr>
<tr>
<td>(10.38)</td>
<td>(9.38)</td>
<td>(10.02)</td>
</tr>
<tr>
<td>Anger/Irritability (AI)</td>
<td>56.01</td>
<td>54.39</td>
</tr>
<tr>
<td>(11.43)</td>
<td>(10.85)</td>
<td>(11.10)</td>
</tr>
<tr>
<td>Intrusive Experiences (IE)</td>
<td>58.22</td>
<td>55.54</td>
</tr>
<tr>
<td>(10.26)</td>
<td>(9.71)</td>
<td>(10.04)</td>
</tr>
<tr>
<td>Defensive Avoidance (DA)</td>
<td>56.68</td>
<td>54.28</td>
</tr>
<tr>
<td>(9.34)</td>
<td>(9.10)</td>
<td>(9.83)</td>
</tr>
<tr>
<td>Dissociation (DIS)</td>
<td>59.79</td>
<td>56.59</td>
</tr>
<tr>
<td>(11.64)</td>
<td>(9.85)</td>
<td>(12.57)</td>
</tr>
<tr>
<td>Sexual Concerns (SC)</td>
<td>52.48</td>
<td>50.12</td>
</tr>
<tr>
<td>(11.59)</td>
<td>(9.77)</td>
<td>(10.23)</td>
</tr>
<tr>
<td>Dysfunctional Sexual</td>
<td>56.78</td>
<td>54.64</td>
</tr>
<tr>
<td>Behaviour (DSB)</td>
<td>(15.55)</td>
<td>(12.81)</td>
</tr>
<tr>
<td>Impaired Self-Reference</td>
<td>57.10</td>
<td>55.17</td>
</tr>
<tr>
<td>(ISR)</td>
<td>(10.24)</td>
<td>(9.42)</td>
</tr>
</tbody>
</table>
Tension Reduction  59.78  57.77  56.24  64.06  61.74  60.19

(M = 52.63, SE = .78) across time. Response level (RL) and inconsistent responding (INC) were not significantly different between fabricated and truthful reports (ps > .05).

Scores on the ATR and INC validity scales differed at both Times 1 and 2 (p < .001 and p < .05, respectively), and Times 1 and 3 (p < .001 and p < .05, respectively). Scores for RL (indicating denial or minimization) differed at Times 1 and 3 only (p < .05).

Table 11.
Means, standard deviations, and percentages of endorsement for truthful (TT) and fabricated (FT) traumatic events over Times 1, 2, and 3 (T1, T2, T3) for TSI validity scales.

<table>
<thead>
<tr>
<th>TSI Validity Scales</th>
<th>TT</th>
<th></th>
<th></th>
<th>FT</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T1</td>
<td>T2</td>
<td>T3</td>
<td>T1</td>
<td>T2</td>
<td>T3</td>
</tr>
<tr>
<td></td>
<td>n=282</td>
<td>n=250</td>
<td>n=174</td>
<td>n=282</td>
<td>n=245</td>
<td>n=173</td>
</tr>
<tr>
<td>Atypical Response</td>
<td>55.27</td>
<td>52.56</td>
<td>51.59</td>
<td>61.54</td>
<td>56.90</td>
<td>56.10</td>
</tr>
<tr>
<td>(ATR)</td>
<td>(13.20)</td>
<td>(10.76)</td>
<td>(11.18)</td>
<td>(15.55)</td>
<td>(14.76)</td>
<td>(14.53)</td>
</tr>
<tr>
<td>Response Level</td>
<td>48.57</td>
<td>49.18</td>
<td>49.71</td>
<td>48.23</td>
<td>49.00</td>
<td>49.53</td>
</tr>
<tr>
<td>(RL)</td>
<td>(8.45)</td>
<td>(9.24)</td>
<td>(10.42)</td>
<td>(8.54)</td>
<td>(9.96)</td>
<td>(10.57)</td>
</tr>
<tr>
<td>Inconsistent Response</td>
<td>54.09</td>
<td>52.64</td>
<td>52.65</td>
<td>55.43</td>
<td>54.24</td>
<td>53.17</td>
</tr>
<tr>
<td>(INC)</td>
<td>(9.80)</td>
<td>(9.76)</td>
<td>(9.38)</td>
<td>(10.37)</td>
<td>(10.44)</td>
<td>(9.52)</td>
</tr>
<tr>
<td>Valid Profiles</td>
<td>90.1%</td>
<td>94.4%</td>
<td>88.5%</td>
<td>84.8%</td>
<td>87.3%</td>
<td>84.4%</td>
</tr>
<tr>
<td></td>
<td>(9.9%)</td>
<td>(5.6%)</td>
<td>(11.5%)</td>
<td>(15.2%)</td>
<td>(12.7%)</td>
<td>(15.6%)</td>
</tr>
</tbody>
</table>

171
| Score Range          | 17.4% | 26.0% | 32.8% | 12.1% | 21.2% | 23.7% | 50.2% | 53.6% | 47.7% | 40.5% | 46.9% | 47.4% | 31.2% | 17.6% | 17.8% | 46.4% | 30.7% | 26.6% | 1.1% | 2.8% | 1.7% | 1.1% | 1.2% | 2.3% | 12.1% | 7.2% | 6.9% | 26.6% | 16.7% | 15.0% |
|---------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| Endorse Unusual     |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| Symptoms            |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| Minimize            | 8.5%  | 10.0% | 11.5% |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| Behaviours          |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| Inconsistent        | 13.5% | 13.2% | 10.9% |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| Responding          |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |

To test the ability of the TSI scales (validity and clinical) to discriminate between memory types, as well as which scales lend the most predictive ability, discriminant function analyses (DFAs) were conducted on scores at each of Times 1, 2, and 3 (using the same cut-offs as previously; see Tabachnick & Fidell, 2001). At Time 1, the model was significant, Wilks’ lambda = .85, $\chi^2_{(13)} = 87.53, p < .001$. The strongest discriminator was intrusive experiences (.72); followed by anxious arousal (.62), atypical responding (.53), defensive avoidance (.52), tension reduction behaviour (.35), dissociation (.34), and sexual concerns (.31). Each of these scales had higher mean scores for fabricated symptoms relative to truthful symptoms (see Tables 10 and 11). The overall discriminability of the TSI at T1 was 66.5% (70.9% truthful; 62.1% fabricated). At Time
the TSI was able to discriminate between truthful and fabricated narratives at an overall rate of 62.8% (68.8% truthful; 56.7% fabricated), Wilks’ lambda = .88, $\chi^2_{(13)} = 61.03, p < .001$. An evaluation of the structure matrix indicated that anxious arousal (.74), intrusive experiences (.60), dissociation (.48), atypical responding (.46), defensive avoidance (.41), and tension reduction behaviour (.40), were able to meaningfully discriminate between narratives (all higher in FTs relative to TTs). At Time 3, the DFA analysis remained significant, Wilks’ lambda = .89, $\chi^2_{(13)} = 39.1, p < .001$. Again, elevated levels of anxious arousal (.70), intrusive experiences (.60), atypical responding (.50), tension reduction behaviour (.44), and defensive avoidance (.34), in fabricated symptom profiles served to meaningfully discriminate these from truthful symptoms. At T3, correct classification of truthful narratives occurred in 70.1% of cases and 56.1% of fabricated cases (overall classification 63.1%).
Chapter 9. Discussion

“The truth is rarely pure, and never simple”

(Oscar Wilde, *The Importance of Being Earnest*, 1895/1985)

As members of the Canadian criminal justice system have recognized, false allegations hold enormous harmful consequences for both accused parties and society in general. Despite the robust finding that legal professionals are generally not proficient at credibility assessment (e.g., Ekman & O’Sullivan, 1991; Porter, Woodworth, & Birt, 2000), little research has addressed how to better identify fabricated reports of traumatic victimization by complainants.

The reality monitoring (RM) framework evaluates differences in autobiographical memory based on self-reported (subjective) characteristics (e.g., Johnson & Raye, 1981; Johnson et al., 1988; Suengas & Johnson, 1988), whereas criterion-based content analysis (CBCA) relies on objective measures to detect deception (e.g., Merckelbach, 2004; Vrij, 2005). Research on both of these approaches primarily has involved truthful and deceptive “stories” of little consequence to the participant (i.e., lying about someone erasing a blackboard) and not personally salient. In this dissertation, the Memory Assessment Procedure (MAP) was used to provide a holistic and novel approach to studying the features of memory, and also utilized in the discrimination of truthful and fabricated narratives. With few exceptions (e.g., Barnier et al., 2005; Parker & Brown, 2000, Porter et al., 1999; Sporer, 1997), a systematic examination of the characteristics of truthful and fabricated narratives of traumatic victimization has not been conducted. Further, limited research has examined the features of truthful and fabricated traumas over time (e.g., Granhag et al., 2001; Suengas & Johnson, 1988). This dissertation
investigated the manner in which fabricated and truthful reports of trauma differ in terms of their narrative features and self-reported trauma symptomology over time.

*Are Truthful and Fabricated Narratives Qualitatively Different?*

One of the main hypotheses was that truthful and fabricated narratives of trauma may be qualitatively different. This notion was based on a detailed review of the literature suggesting that truthful statements are rich in detail and contextual information relative to fabricated narratives (e.g., Masip et al., 2005; Vrij, 2005). With respect to objective characteristics, one prediction was that truthful narratives would be more detailed, emotional, contextually established, and plausible relative to fabricated claims. These differences would result from the well-established finding that liars simplify their narratives in order to “keep the story straight” (e.g. DePaulo et al., 2003). Another prediction was that fabricated narratives may appear more coherent and rehearsed, reflecting a story-telling narrative style, relative to truthful narratives.

My findings support the first prediction: truthful narratives contained more detail overall, emotional components relating to self and others, contextual (time and place) information, relevant details, a greater number of words, and were rated as more plausible relative to fabricated statements. These results corroborate research using both CBCA and RM procedures for assessing the credibility of a statement (e.g., Granhag & Vrij, 2005; Masip et al., 2005; Vrij, 2005). Although the prediction was made that fabricated narratives would be rated objectively as more coherent than truthful narratives, my results do not directly support this prediction. Fabricated narratives were not more coherent than truthful traumas – comparable levels of coherence were found between narrative types.
Studies using the CBCA technique have, in general, found that unstructured production (i.e., less coherent, disorganized, unconstrained) to be indicative of truthfulness and less prevalent in deceptive statements (Granhag & Vrij, 2005). Truthful narratives have been associated with less coherence related to variations in memory recall (i.e., spontaneously remembering more details as they come to mind), relative to the overly structured and coherent quality of deceptive accounts due to rehearsal. In this study, the lack of a difference between narrative types on levels of coherence may have resulted from the nature of the within-subjects design. Subtle carry-over effects may have influenced the manner in which participants provided statements for each of the truthful and fabricated events, following a similar structural pattern for each. Alternatively, levels of coherence may not have differed due to reliance on real-life traumatic experiences when fabricating. This may involve activation of a schema or 'trauma script' (based on their personal experiences) which enables similar methods of reporting traumatic events regardless of veracity (see Anderson et al., 2000). The written narrative format used to elicit both narratives may have limited the detectable differences on this variable. A comparison of written versus oral narratives for truthful and fabricated traumatic experiences may reveal differences in the predicted direction. Most likely, however, is that the memory paradigm used in this dissertation did not allow for sufficient time for fabricated narratives to be rehearsed and become more coherent than truthful statements. It is possible that more time to prepare fabricated claims of victimization would be associated with greater levels of coherence relative to truthful memories (e.g., Alonso-Quecuty, 1992).

This study also evaluated subjective features of trauma narratives. As stated by DePaulo et al. (2003): “Liars usually make an effort to seem credible; truth tellers more
often take their credibility for granted” (p. 78). In line with this conclusion, one prediction was that subjective ratings would be inflated in fabricated claims of trauma in order to appear credible. This prediction was not supported by the results. Truthful narratives were rated as vivid, having a better overall memory quality, more coherent, and perceived as more credible relative to fabricated narratives. It is likely that this pattern resulted from participants’ personal knowledge of the veracity of the event which was revealed in their subjective ratings, whether intentional or not. This reflects metamemory processes (such as “feelings of knowing”) and the surrounding context of recall (i.e., “I know I was never sexually assaulted”) when participants are required to make assessments of their own memory features (e.g., Johnson et al., 1988; Kelly et al., 2002). In their review of deception detection literature, DePaulo et al. (2003) also argued that liars have more difficulty convincingly embracing their deceptive claims due to moral qualms, personal investment, and lack of knowledge and experience in the circumstances being lied about. Studies based on real and imagined events using RM techniques have frequently found that participants subjectively rate their own truthful experiences as more realistic, memorable, believable, and detailed (e.g., Arbuthnott et al., 2002; Kealy et al., 2006; Sporer, 1997). In addition, past research has demonstrated that subjective ratings of negative events yield differences as individuals are less likely to believe a distressing event they imagined or fabricated happened to them, and rate them accordingly (e.g., Kemp et al., 2003). It also is possible that the type of trauma reported (e.g., car accident, criminal victimization, death of a loved one) may influence the features of narratives and self-rated memory properties. The traumatic nature of the events in this study may have impeded participants’ abilities to respond as if the
fabricated trauma had happened. Thus, these overall findings support the theoretical basis of reality monitoring decisions when applied to subjective assessments.

Differences in the subjective features of narratives were especially prevalent between Times 1 and 2, and the overall Time 1 to 3 interval. Higher overall ratings associated with truthful narratives (as indicated above) gradually decreased over the six month interval. However, truthful and fabricated narratives did not differ between Times 2 and 3. While subjective ratings of memory features declined in a more linear fashion for truthful traumas, ratings for fabricated traumas dropped more (in general) between Times 1 and 2 and then levelled off between Times 2 and 3. The only exception to this pattern was for sensory details – fabricated (relative to truthful) narratives were rated as having more sensory components initially at Time 1, but ratings were equivalent to truthful narratives at Times 2 and 3. Contrary to these results, memory assessments using RM and/or CBCA techniques have found that sensory details (both objectively and subjectively rated) are more prevalent in truthful relative to fabricated narratives (see Masip et al., 2005). Fabrication inflation of sensory components may have resulted from perceptions common in popular crime-oriented television programs (e.g., Crime Scene Investigation; Law & Order) where investigators focus on sounds, smells, and sensations surrounding criminal victimization when questioning victims and witnesses. This may have been more pronounced when participants first generated their fabricated traumas, but the inflation of sensory details was not maintained over time due to simplification of their "stories". Based on these overall findings, it appears that subjective features of fabricated narratives become more similar to truthful narratives (and more difficult to distinguish) with repeated retellings (e.g., Granhag et al., 2001; Granhag & Strömwall,
2002; Suengas & Johnson, 1988). For example, Vrij (2004) argued that memory ratings may fluctuate to a greater extent during initial recall attempts, but these become more stable over time as the deceptive statement is solidified with successive retellings. The results of this study (pertaining to subjective ratings) support this view (discussed in more detail below). Further research involving subjective memory assessments is necessary to determine how these ratings may (or may not) differ over extended time intervals.

**Discrimination of Truthful and Fabricated Traumas**

As reviewed above, the overall features of truthful and fabricated narratives of trauma are different. In order to address the degree to which these features contribute to the ultimate discrimination of truthful and fabricated narratives, discriminant function analyses were conducted. No direct *a priori* predictions were made regarding the ability of the Memory Assessment Procedure (MAP) to discriminate narratives as no studies have examined its utility in that respect.

Overall, both objective and subjective components of the MAP correctly discriminated between truthful and fabricated narratives of trauma at a rate of 67% at Time 1, 72.4% at Time 2, and 72.3% at Time 3. The MAP features classified truthful traumas (70.1%, 74.1%, and 74.7%, respectively) at higher rates than fabricated traumas (63.8%, 70.6%, and 69.9%, respectively). This finding conforms to previous research on deception detection (e.g., Sporer, 1997; Vrij, Akehurst, et al., 2004b), and may result from a greater prevalence of specific criterion (e.g., amount of detail) in truthful relative to fabricated reports (e.g., Lucas & McKenzie, 1999). Taken as a whole, these findings are comparable to those demonstrated in previous studies; credibility assessments using
CBCA and RM have evidenced an overall rate of discrimination of approximately 72% (Granhag & Vrij, 2005).

Although discrimination of truthful and fabricated narratives is possible, few studies have examined changes in statements over time. Erdmann et al. (2004) reported that CBCA criteria lost their discriminative ability over the course of five interviews (with a 2 week interval between each interview) with children asked about a real or fictitious negative event (e.g., falling off a horse). That is, with repeated recall or questioning, their statements became more similar over time. Strömwall and Granhag (in press) found that discrimination ability of the RM technique decreased over a week interval, with correct classification rates of 85% at the initial interview and 79% at the second. They argued that there was a trend for truthful and fabricated events to become less distinct over time (also see Granhag et al., 2001). On the other hand, Alonso-Quecuty and Hernández-Fernaud (1997) found an increase in contextual and sensory information over successive statements (little time delay), particularly in truthful narratives. Although they did not use discriminative function analyses, their findings suggest that RM criteria were more discriminative after the third recall relative to the first recall. The findings of this study support previous research that has found increased discriminability over time with CBCA and RM models (e.g., Granhag et al., 2006; Strömwall et al., 2004; Vrij, 2004). Interestingly, the discriminative power of the MAP increased more between the Time 1 to Time 2 interval and then remained stable between Times 2 and 3. This likely results from a consolidation of details with repetition (e.g., Suengas & Johnson, 1988; Vrij, 2004), as discussed above.
Another important objective of this dissertation was to investigate features of narratives that are predictors of veracity. Ratings of plausibility (objective) and credibility (subjective) were the strongest indicators of narrative type across all three time periods in this study, with plausibility being the most predictive. The results of this study strongly support past research that has suggested liars tell less convincing stories (DePaulo et al., 2003). The extent to which a story “sounds plausible” is perhaps one of the most difficult criteria to quantify, and yet studies have repeatedly found it to be a diagnostic cue to deception (Granhag & Vrij, 2005). Although difficult to define operationally (but yet reliably coded in this study), plausibility ratings appeared to capture some quality of the narratives that other characteristics did not to the same extent. That is, the rating of plausibility allowed for a determination of whether a narrative had “the ring of truth”. This also follows from the argument made by Pezdek and Taylor (2000) that fabricated narratives may contain a more story-like quality adhering to a general schema about traumatic events. While schematic information also is present in truthful narratives, there would be more specific (episodic) details about the experience that make the story more plausible. Two examples of similar events (an outdoor attack) in the present sample provided (by different participants) as truthful and fabricated claims of victimization exemplify this point:

Example 1: “I had been out at the bar drinkin with a bunch of my friends and decided to walk home rather than cabbin it as I didn’t have the cash. It was pretty late, maybe 2:30 or so. I decided to cut through the cemetery off Spring Garden to save time getting home. Plus it was kinda misty and
cool out and the graves were all eerie. I guess I was in my own little world as I fell down close to this big tree. I thought I tripped at first but next thing I knew there was this guy standing over top of me. I didn’t hear him come up or anything and was terrified. I kind of jolted up then and tried to scramble away, my shoe got stuck on a tombstone on the ground, one of those flat grave marker things. That is when I heard it, when I was trying not to fall on my face again, Joe’s voice (my ex) telling me he wanted to talk to me, he’d been following me. I started laughing at that point until he grabbed my shoulders and shook me …”

Example 2: “Late one Tuesday night at around 2:00am, I was walking through the woods taking a short-cut home. It was a dark and dreary night, and the rain from earlier had muffled the sounds all around me. The dampness seeped into my bones, I felt so cold and isolated. Suddenly, off to my left a branch snapped in the woods, and footsteps came crashing down upon me as my assailant jumped from the darkness and grabbed me …”

It is evident that both narratives contain trauma schema information related to being attacked late at night. However, Example 1 (truthful) has more specific details about the experience whereas Example 2 (fabricated) is more story-like and appears less plausible. Previous studies also have found plausibility to be related to perceptions of credibility, and the accurate detection of deception (e.g., Hiltunen, 1996; Kraut, 1978; Landstrom et
al., 2005; Strömwall & Granhag, 2003b). This also would be associated with self-rated perceptions of credibility and confidence levels in the accuracy of a report. Collectively, these relationships provide a likely explanation for why subjectively rated credibility and confidence also were predictors of veracity at all three testing times.

Objective ratings of time and place details (contextual information) and number of words also served as reliable predictors of narrative type over time. Research with both CBCA and RM techniques has suggested that these features are more common in truthful relative to fabricated narratives and are generally retained in memory (Granhag & Vrij, 2005). Similarly, the amount of detail has been consistently evidenced as an indicator of deception (Vrij, 2005). Here, although the amount of detail was not a strong predictor of veracity at Time 1, it became more useful as a cue to credibility over time (Times 2 and 3). This likely resulted from increased investment and ease of providing detail in fabricated narratives during the initial phase (generation), but increased difficulty in maintaining this level of detail (from memory) at Times 2 and 3. Related to this, self-rated coherence also emerged as a predictor of veracity at Times 2 and 3. It is likely that as participants had more difficulty remembering the details they previously provided in their fabricated narratives, they felt their stories were less coherent and “put together” relative to their truthful traumas. In other words, participants experienced “fabrication deflation” relative to their truthful events (e.g., Polage, 2004).

The analysis also revealed that emotional components relating to the self only was a predictor of veracity at Time 1 (i.e., above the .30 cut-off recommended by Tabachnick & Fidell, 2001). Although truthful narratives contained more of these details relative to fabricated narratives at Times 1, 2, and 3, it is possible that this feature decreased in its
discriminative value as other stronger predictors (such as amount of detail and self-rated coherence) emerged at Times 2 and 3. That is, larger differences between narratives for the amount of detail and levels of coherence were present resulting in a stronger discrimination value. Further, the mean level of emotional details relating to oneself at Time 3 was relatively small limiting the extent to which true differences could be detected. Finally, subjective ratings of memory vividness and quality were moderate predictors of narratives type at Time 2 only. Research by Sporer (1997) reported a similar finding where truthful narratives were rated higher in vividness/ clarity, but this was only evident after a delay. Based on analysis of the mean vividness and quality ratings, these features were predictors at Time 2 only due to the largest mean difference at this time (i.e., fabricated claims decreased more between Times 1 and 2 relative to truthful narratives). Overall, objective features assessed by the MAP (e.g., plausibility, time and place details, amount of detail, and number of words) were strong and reliable predictors of veracity.

**Consistency of Truthful and Fabricated Narratives**

Not only are the objective and subjective features that differentiate narratives important, but also the degree of consistency associated with narratives over time. Due to the legal significance of testimonial consistency in perceptions of credibility (e.g., Bell & Loftus, 1989; Brewer et al., 1999), narrative changes over time are a critical, but understudied, component of veracity determinations (e.g., Granhag & Strömwall, 2001b). While it is commonly believed that liars will have difficulty “telling a good story” in a reliable manner, little empirical evidence has been conducted to substantiate this view.
This dissertation evaluated narrative features and responses to direct questions for both trauma types over three successive testing times. Due to differences in the source of the "memory", it was predicted that truthful narratives of trauma would maintain higher levels of consistency relative to fabricated narratives. In addition, the effects of rehearsal or retelling were predicted to influence fabricated narratives to a greater extent, such that initial decreases may be evident but level off as details become consolidated in memory. The results of this dissertation support the former prediction, but not the latter.

Truthful narratives of trauma had higher total consistency scores (measured by the CQ) than fabricated narratives at each of the time intervals measured. In addition, both truthful and fabricated narratives decreased in consistency over the extended six month (164 day) interval. Interestingly, levels of consistency for truthful and fabricated traumas were higher between Times 2 and 3 relative to initial (Time 1-2) and overall (Time 1-3) testing periods. This likely resulted from the different length of the testing periods due to difficulties in scheduling participant appointments (i.e., no shows). The interval from Time 1 to 2 was approximately 91 days, whereas Time 2 to 3 a shorter interval of 72 days. As such, participants may have demonstrated higher levels of consistency as they were able to better recall responses to questions from the previous testing period. After a lengthier delay, this effect may have deteriorated. Participants also may have chosen a few key features to try to remember consistently over time and rehearsed these more. To determine if this increase in consistency was an artifact of the testing period, the consistency of responses at Times 1 and 3 were compared. These results demonstrated that consistency decreased (not increased) over the extended time period, as would be predicted, for both narrative types. Therefore, the elevated level of consistency over the
T2-T3 interval appeared to be an artifact of the shorter time period and this elevation was not maintained over time.

Overall, truthful narratives were recalled more consistently than fabricated narratives, and both decreased over time. The lack of any interaction when evaluating consistency over the six month interval indicated that fabricated narratives did not stabilize over time due to rehearsal and memory consolidation, but continued to decline in the same manner as truthful narratives. Although total mean consistency scores for Times 1-2 and Times 1-3 were different between narratives, truthful and fabricated traumas changed only a small degree over time (approximately 1-1.5 points on the consistency scale). While professionals and laypersons alike often perceive the “truth” as being steadfast and deception as more variable (e.g., King, 2006), the results of this dissertation do not confirm this idea with respect to direct questions about an event. Studies have reported that fabricated narratives of trauma are less consistent relative to truthful narratives (e.g., Mazzoni & Memon, 2003; Polage, 2004), but few (if any) have found similar rates of change in narratives.

These results can be explained by considering memory for genuine versus imagined (or deceptive) events. As discussed previously, there are differences between recalling a genuine memory versus retelling of a story (e.g., Dudukovic, Marsh, & Tversky, 2004; Granhag et al., 2006; Lampinen et al., 2003; Suengas & Johnson, 1988). Truthful traumas are based in reality, with a detailed memory that has been encoded, stored, and recalled. As evidenced in this study, memories of truthful trauma were highly detailed with contextual and perceptual information. On the other hand, fabricated narratives have no basis in reality, thus perceptual and contextual information must be
generated and is unlikely to leave a lasting “impression” in memory. Recollection of genuine events is derived from the memory of the original experience, whereas recollection of fabrication involves remembering a prior episode of being deceptive. More detailed and consistent reports are associated more so with truthful claims due to difficulties in embellishing a fabricated report with details generated solely from the imagination and maintaining this over time (Arntzen, 1983).

In addition, the superior recall of trauma relative to other types of emotional experiences has been repeatedly demonstrated in the literature (e.g., Alexander et al., 2005; Porter & Peace, in press). While both narratives elicited in the present study were traumatic in nature, a genuinely experienced traumatic event is more personally salient and would undoubtedly be well-retained in memory relative to one that was imagined (e.g., Pezdek, 2003; Porter et al., 1999). However, truthful traumas would not be fixed and indelible but exhibit minor changes in consistency over time due to the natural variations in memory – such as remembering another detail about the event or temporarily forgetting a detail previously reported (e.g., Granhag & Strömwall, 2002; Loftus, 2003b; Parker & Brown, 2000; Turtle & Yuille, 1994). On the other hand, false allegations of trauma may be distinctive in their invention, but are not directly experienced and are more difficult to maintain consistently over time. The design of this dissertation afforded participants the opportunity to rehearse and retell their fabricated narratives on multiple occasions which likely preserved the gist of their deceptive claims (such as the main details)(e.g., Altmann, 2003; Colwell, Hiscock, & Memon, 2002; Winningham et al., 2000). Items on the Consistency Questionnaire (CQ) primarily surveyed main details about the event (e.g., time of day, location, emotional state) that
may have been retained as “gist” information (see Pezdek, 2003). Analysis of specific consistency items yielded no effect of time (i.e., ratings did not change significantly overall when collapsed across narrative type) and no interaction effects. Overall, truthful narratives were more consistent in details pertaining to clothing, time of day, day of the week, calendar date, smells, sensations, age at the time of the event, and physical injuries, relative to fabricated claims. These categories reflect more factual details that genuine victims may naturally recall, and alleged victims may use to “anchor” their story (to the best of their ability) over time. Few changes in responses over the six month interval indicate that consistency was a relatively static characteristic within each narrative type. Further research evaluated the consistency of the narratives themselves and the specific details mentioned at Times 1, 2, and 3 (rather than responses to direct questions on the CQ) may yield a different pattern of results and show more variation in narratives over time.

Consistency scores also discriminated between narrative types. Analyses using the total consistency scores at each of the time intervals revealed that truthful and fabricated narratives could be correctly classified 69.9% of the time. In particular, total scores for each narrative type between Times 1 to 3 was the strongest predictor of veracity, closely followed by total consistency scores obtained by comparing Time 1 and 2 responses. Total scores on the CQ between Times 2 and 3 were less predictive than the former. It is apparent that the overall level of consistency measured over an extended time period (i.e., six months) would provide the most useful information for credibility determinations. This is important in criminal justice settings given the lengthy delays between initial statements of victimization to police and subsequent statements during court proceedings.
months or years later (e.g., Eurele & Turtle, 1999). My findings suggest that inter-
interview consistency is less important (possibly due to natural variations in recall and
wording with repeated questioning) relative to the overall level of consistency between
the first report of an incident and the final recall. In addition, specific consistency items
pertaining to factual details about the event evidenced higher classification scores than
evaluating total consistency scores (summation of 23 consistency items). During the first
time interval (Times 1 to 2), the calendar date of the event, physical injuries, age at time
of trauma, clothing, and smells associated with the event were meaningful predictors of
narrative type. Collectively, these items were able to correctly discriminate between
narratives at a level of 75.7%. A comparison of consistency items across the Time 1 to 3
interval indicated that the calendar date of the event, age at the time of the event, and
clothing worn when the event happened remained predictive following the six month
time period (overall classification rate of 78.7%). In each of these cases, consistency item
scores were higher for truthful relative to fabricated traumas. These findings support
studies using CBCA and RM techniques that have found contextual embedding to be
highly important for discrimination. Details relating to the context of the event (e.g., date,
age, clothing) are unlikely to be forgotten in truthful memories (and particularly of
traumatic events; also see Porter & Peace, in press). On the other hand, greater variation
in reporting these specific details is evident in fabricated claims of traumatic
victimization. The enhanced discriminability of specific consistency items suggests that
these may serve as deceptive markers or flags for investigators to pay particular attention
to when interviewing genuine and alleged victims of trauma. In addition, these markers
may vary according to the type of alleged trauma. Future research evaluating consistency across trauma types is necessary to elucidate this relationship.

**Symptoms of Truthful and Fabricated Trauma**

Traumatic victimization also is associated with subsequent psychological sequelae, such as stress, hyperarousal, intrusive experiences, and attempting to avoid reminders of the experience. Individuals fabricating traumatic events are likely to feign this type of symptomology to appear genuine. As a result, levels and patterns of symptom reporting are important to consider when assessing the credibility of victimization claims, particularly in legal contexts. For example, research has demonstrated that litigation and compensation-seeking scenarios are associated with a higher proportion of false allegations of traumatic stress (e.g., Lees-Haley, 1997; Resnick, 1997; Rosen, 2004). Symptoms of trauma are common, highly subjective (e.g., Bury & Bagby, 2002; McGuire, 2002), and may be easy to successfully fabricate. Some studies have found few differences between malingerers and genuine victims (e.g., Frueh et al., 2005), whereas other research has demonstrated variations between groups that warrant further investigation (e.g., Liljequist et al., 1998; McGuire, 2002; Porter, Peace, & Emmett, in press). It remains unclear as to how symptoms reported in a genuine versus malingered fashion differ in general, and over time. As such, it was predicted that fabricated symptomology would be associated with an “over the top” quality. Levels of traumatic stress, PTSD, and other clinical symptomology (e.g., depression) would be inflated in fabricated relative to truthful reports across all time periods.
This prediction was supported by the results of this study. Overall, fabricated traumas were rated as having higher total and subscale mean scores on the IES-R, PCL, and TSI, relative to truthful traumas. This confirms previous research demonstrating "fabrication inflation" (e.g., McGuire, 2002; Porter et al., 1999; Rosen, 2004). However, evaluation of the mean scores indicated that while inflated, these were not exaggerated to such an extent as to appear grossly abnormal (e.g., endorsing 5 for each symptom on the PCL). For example, on the TSI, mean scores for fabricated traumas were not outside the "normal" range of responding that signifies clinical significance. This may indicate that participants’ demonstrated some level of sophistication in their deceptive symptom reports. Alternatively, participants may have become fatigued when completing the symptom measures (i.e., 100 questions on the TSI) and might have responded in a more standard fashion (i.e., endorsing a pattern of 2-3-2-3-2 etc.).

That being said, neither the IES-R nor the PCL have validity scales that examine patterns of responding to determine the validity of a symptom report whereas the TSI does. Results indicated that of the three TSI validity scales (atypical responding, response level, and inconsistent responding), fabricated symptom reports only were elevated on levels of atypical responding (ATR) relative to truthful reports. The ATR scale measures endorsement of statistically unusual symptomology and elevations on this scale can occur from patterned and indiscriminate responding (i.e., 2-3-2-3-2) which supports the explanation above. Overall, the TSI validity scales were not good indicators of the veracity of a report. Of the fabricated reports, the majority were deemed valid profiles.

---
24 Only two clinical subscales on the TSI were not significantly different between narratives (depression and deviant sexual behaviour) despite higher mean scores for fabricated relative to truthful reports.
(i.e., 84.8% at Time 1, 87.3% at Time 2, and 84.4% at Time 3). In addition, truthful symptom profiles were considered valid 90.1% (Time 1), 94.4% (Time 2), and 88.5% (Time 3) of the time – meaning that some profiles were incorrectly classified as invalid. The validity scales appeared better at successfully classifying a truthful report as valid (a true-positive), primarily due to the discriminative power of the ATR validity scale. The strongest clinical predictors were anxious arousal, intrusive experiences, tension reduction behaviour, and defensive avoidance. The subtleties of content for these particular scales may have been more difficult for individuals to reliably fabricate relative to scales assessing depression or sexual concerns. Overall accuracy of the TSI in discriminating between truthful and fabricated reports was 66.5% at Time 1, 62.8% at Time 2, and 63.1% at Time 3. While significantly different than chance levels, sole reliance on the TSI to determine veracity would not be advised.

Subjective reports of emotional distress (from the Emotional Memory Survey) also were assessed in this study (e.g., anxiety/stress, thoughts about the event, turning off emotions, becoming emotional, emotional intensity). While levels of subjective emotionality decreased over time for both truthful and fabricated traumas, no difference between narrative types was revealed (although mean fabricated responses were, in general, higher than truthful ratings). As discussed previously, metamemory processes may have interfered with participants’ ability to effectively fabricate emotional responses to non-existent traumas. In particular, these findings may be partially dependent on the questions used to measure subjective emotionality. Symptoms reported on standardized measures of traumatic stress (i.e., IES-R, PCL, TSI) demonstrated the inflated pattern that was predicted. It is possible that ratings of generalized emotionality on the EMS did not
demonstrate specificity or sensitivity relative to lists of symptoms that participants endorsed on standardized measures. Future research comparing subjective ratings on the EMS with additional assessments of emotionality is necessary to determine the utility of these items on this measure.

It was also predicted that fabricated symptom reports would be exaggerated at a higher level and more consistently over time, relative to truthful symptoms (which tend to recede with the passage of time). This prediction was not supported by the results related to the IES-R and TSI scales. Mean total and subscale scores for these measures decreased over time for both fabricated and truthful traumas. However, scores on the PCL (a direct measure of PTSD) did support this prediction. While all scores decreased over time on this measure, these changed at different rates for truthful and fabricated traumas. Specifically, PCL total, avoidance, and hyperarousal scores deteriorated to a greater extent (overall) for truthful relative to fabricated symptom reports. This finding is consistent with research that suggests individuals fabricate clinical levels of PTSD to a greater extent and maintain higher levels of symptomology relative to acute (genuine) PTSD symptoms that recede over time (McCann, 1998; Rogers, 1997; Schiraldi, 2000). It is unclear why this finding was not evidenced for the IES-R scale, which also measures levels of traumatic stress and PTSD symptomology. The mapping of PTSD criteria from the DSM-IV onto the PCL questions (17 items) directly assesses this disorder, whereas the IES-R (23 items) may contain some extraneous factors that limit the results. Further evaluation and comparison of these (and other) measures of PTSD symptomology in the context of malingering is critical for understanding changes in fabricated symptoms (e.g., Rogers et al., 2005). Certain scales may be more or less sensitive to subtle variations over
time in symptom endorsement. In addition, levels of symptomology may vary according to the type of trauma (e.g., criminal victimization, accident), and be subsequently related to differing levels of symptomology for fabricated claims. Standardized and consistent approaches to deception detection are essential in clinical and forensic contexts (e.g., Koch et al., 2006; McCann, 1998; Rogers, 1997). While results in this research indicate that fabricated symptomology is inflated, reliable clinical indicators of deceptive symptoms were not found on the measures used in this study.

Limitations and Strengths of this Research

Researchers conducting empirical evaluations of complex topics such as traumatic memory, recall consistency, and credibility assessment struggle in finding a balance between experimental control and ecological validity. Some researchers argue that experimental studies conducted in the laboratory are not sufficient for testing the utility of techniques used for credibility assessment (e.g., Kleinmuntz & Szucko, 1982; Lykken, 1998; Steller, 1989; Undeutsch, 1984). However, field research does not provide the experimental control over influential variables to draw any definitive conclusions. In order to further elucidate the relationships between these variables, it is necessary to apply structured methods to the recall of real-life personally salient traumatic experiences, as well as the generation of fabricated narratives. However, there are limitations associated with this type of research.

In any type of applied experimental paradigm involving traumatic memory, it is increasingly difficult to obtain substantiation of traumatic experiences. This is known as ground truth – establishing with absolute certainty when individuals are lying and when
they are telling the truth. Vrij (2004) mentions this as a problem in most experimental studies of deception and its detection. There was no way to establish (due to ethical and practical restraints within the sample) whether participants’ truthful traumas were indeed genuine and that their fabricated traumas had not happened in reality. While it is possible that a small number of the “truthful” traumas were falsified, I do not believe this was a substantial problem. First, the recruitment of participants was based on having experienced a traumatic event since the age of 16 and being willing to provide details about this event to the experimenter. All participants were reminded prior to commencing the study that they must have experienced a genuine traumatic event and were asked if they had an event in mind that satisfied this criterion. Most participants reported a recent traumatic event they had experienced without any hesitation. Second, participants did not appear to have difficulty when completing the truthful trauma portions of the experiment – they did not spend a great deal of time thinking or deliberating about answers relative to the time they took preparing their fabricated narratives. Third, the numerous differences between narrative features, consistency, and symptomology demonstrated in this study would have been limited if participants were providing two fabricated reports of trauma (e.g., Vrij, 2002). Not only were the findings of this study highly significant, they also were comparable to previous research on credibility assessment techniques (see Masip et al., 2005, and Vrij, 2005, for reviews).

The methodology of the study also is associated with several other limitations. The first is the use of an undergraduate sample. However, past studies on credibility assessment techniques using applied (e.g., victims/perpetrators) or undergraduate samples

---

25 Several participants had experienced multiple traumatic events since the age of 16 and asked which they could use for the study (whatever they were most willing to discuss).
have found comparable results between these groups (Vrij, 2005). Future research with a community sample and verified cases of victimization (e.g., clientele from Victim Services) is needed to further address these issues. As a result, the generalizability of the current findings may only apply to credibility assessment in traumatized university populations. The second potential limitation is the within-subjects design and the relatively brief time between each narrative type (approximately one hour). It is possible that subtle carry-over phenomena occurred when providing two memory reports and led participants to use a more standardized or uniform way of reporting that may have constrained true differences (Merckelbach, 2004). However, the results of this study do not support the notion that carry-over effects limited, in general, the differences between narrative types. In addition, liars often incorporate strategies into their lies that make claims appear more plausible and convincing, such as incorporating details from previous (genuine) experiences (Vrij, 2004). This study was modeled after several seminal studies (e.g., Johnson et al., 1988), including the design chosen for this research (e.g., Kealy et al., 2006; Lampinen et al., 2003; Merckelbach, 2004; Polage, 2004; Porter & Peace, in press; Porter, Peace, & Emmett, in press; Rubin et al., 2003; Santila et al., 1999; Sporer, 1997). Additional analyses examining the features of only the narratives provided first would further elucidate this finding. In future research using a within-subjects approach, it is important to separate the two accounts by days or weeks to avoid any possible carryover effects (although counterbalancing would reduce this problem somewhat). Studies using between-subjects designs (participants provided a truthful or deceptive story) have found comparable differences on CBCA and RM criteria as those where participants provided both truthful and deceptive stories (Vrij, 2005).
Another potential limitation of the present research was the lack of preparation time that participants had to “plan” their fabricated narratives. Alonso-Quecuty (1992) reported that when participants had 10 minutes to prepare a fabricated narrative, their stories were more richly detailed and had more RM criteria relative to truthful statements. Vrij (2004) noted that it may be more difficult for the liar to maintain a consistent story when required to concoct a claim “on the spot”, as this involves generating plausible responses and remembering what was previously said to avoid contradictions. However, the robust effect of plausibility and other objective and subjective narrative features in this dissertation speak to the ability of measures to discriminate between reports even if a lie is told spontaneously or if it was constructed with preparation time. In addition, the relatively small (although significant) decrease in consistency of fabricated narratives over a six month period suggests that liars were able to generate memorable responses during the initial experimental session. Future research should investigate the influence of preparation in the features of fabricated narratives, especially those of a highly emotional nature and serious consequence. It could be predicted that the motivation to successfully deceive in high stakes situations would lead to greater preparation time and fewer potential variations in statement consistency as a result (e.g., Lees-Haley, 1984).

Related to preparation is the level of motivation of the individual. On a basic level, participants may not have been sufficiently motivated to complete the experimental sessions as instructed and may have demonstrated fatigue with the process (e.g., partial questionnaire completion, copying responses from real to fabricated questionnaires to “save time”, etc.). As such, more effortful attention to the task may have revealed a different pattern of results. However, the consistent and strong effects revealed in this data
appear to indicate otherwise. Other studies using a similar design have reported that motivation did not appear to be undermined by the task of providing two narratives as participants still provided high levels of detail that were comparable to previous research (e.g., Merckelbach, 2004). On a more complex level, participants may not have been sufficiently motivated to be convincingly deceptive. Participants were instructed to make their fabricated narratives as realistic as possible and that their goal was to successfully be able to fool a judge into thinking their trauma was real. They also were told that their reports would be assessed by students and legal professionals (e.g., judicial members, lawyers, police officers) in order to determine the veracity of the traumatic memory report. However, the motivation to lie may not have been realistic as there were no “stakes” for generating highly plausible fabricated claims. It is difficult to establish realistic motivators in experimental research that prompt participants to tell convincing, but deceptive, accounts of trauma. Detailed evaluations of fabricated claims of victimization, and corresponding motivations, should be conducted in field studies to fully ascertain the role of motivation in deceptive claims of trauma. Were the “stakes” higher (e.g., being charged for obstruction of justice), greater variations in memory and symptom accounts between real and fabricated accounts may be evidenced (Köhnken, 2002).

That being said, previous studies have addressed the role of motivation in deceptive behaviour. Cues to deception are more easily discerned in motivated relative to unmotivated liars (e.g., DePaulo, Blank, Swaim, & Hairfield, 1992). Several experimental studies have found that high-stakes lies involved more emotion, content complexity, and impression management, and were easier to detect than low-stakes lies (e.g., DePaulo,

---

26 This instruction was approved during REB review. In addition, statements will be shown to criminal justice professionals in future studies.
LeMay, & Epstein, 1991; Lane & DePaulo, 1999; Porter, McCabe, Woodworth, & Peace, in press; Porter & Yuille, 1996; Vrij, Harden, et al., 2001). Higher levels of motivation to lie successfully are associated with greater difficulty at lying, a phenomenon known as the “motivational impairment” effect (e.g., DePaulo et al., 1992; DePaulo & Kirkendol, 1989; DePaulo et al., 2003). In the present study, it would be expected that increased levels of motivation to successfully fabricate a report of trauma (e.g., revenge, compensation) would be associated with more indicators of deception that would enable truthful and fabricated claims to be differentiated. Future research investigating false allegations for motivated purposes such as compensation is necessary to determine the influence of motivation on narrative features.

Last, the findings in this study may have been limited (to some extent) by the selection of the measures (apart from those mentioned previously in the discussion such as the CQ). These results suggest that some objectively rated narrative features are strong predictors of narrative veracity. The inclusion of a larger number of objective criteria in a memory assessment procedure may assist in the discrimination of truthful and fabricated traumas. One purpose of this study was to determine the utility of the MAP in discriminating between narrative types. While the MAP demonstrated comparable levels of discrimination with both CBCA and RM techniques, it also may be associated with some of the same criticisms of the latter methods. All of the criteria assessed by the MAP are, in general, more prevalent in truthful relative to fabricated claims. The inclusion of assessment criteria that are specifically present in deceptive claims (e.g., cognitive operations) may be useful in future research. Additional studies need to evaluate the ability of the MAP to detect deception relative to other approaches and establish
formalized scoring and credibility cut-offs. Only the CBCA/SVA technique has, to date, a scoring procedure and ranges specified that correspond to varying levels of credibility (Bradford, 1994). In addition, use of a broader variety of symptom measures (with validity scales) may have revealed a different pattern of findings. Finally, personality and individual difference measures also may influence credibility assessments. For example, Merckelbach (2004) has conducted extensive research on the role of fantasy proneness and dissociative tendencies on memory reporting as well as deception abilities.

This study also had several strengths worth noting. To my knowledge, no previous studies have provided a concurrent examination of traumatic memory, consistency, and symptomology within the context of credibility assessments. In addition, the discriminative ability of the MAP and CQ has not previously been assessed in this manner. This permits the direct comparison of these measures to well-established techniques for detecting deception through statement analysis such as CBCA and RM. While discrimination rates of the MAP features (collectively) were significantly above chance level, Vrij (2005) cautioned that these rates are not high enough to be used as primary evidence in criminal courts. That being said, the MAP demonstrated good discrimination between narratives, and higher rates of deception detection relative to untrained human judges (see Vrij, 2004). Research with both professionals and laypersons has demonstrated that they lack valid knowledge related to empirically-based cues to deception and often perform at around chance levels (or lower) without specialized training or specific feedback (e.g., Bond & DePaulo, in press; Ekman & O’Sullivan, 1991; Kassin & Fong, 1999; Kaufmann, Drevland, Wessel, Overskeid, & Magnussen, 2003; Porter, McCabe, et al., in press; Porter et al., 2000). Several studies that specifically
have used samples of truthful and false or fabricated memory reports in their deception
detection paradigms have found similar results. For example, Campbell and Porter (2002)
found that judges were better at correctly determining the credibility of false
(mistaken/implanted) memories (60%) relative to true memory reports of negative
childhood experiences (53%). Porter, Campbell, Stapleton, and Birt (2002) had
undergraduate participants judge the credibility of four truthful and four deceptive
accounts of emotionally laden childhood events. They found that the overall level of
accuracy for the sample was 57.6%. While content cues present in the narratives were
relied upon in judgments, these were negatively correlated with successful deception
detection. A recent honours thesis evaluated credibility assessment in an undergraduate
population using a sample of truthful and fabricated claims of sexual trauma (Peace et al.,
2006). Results indicated that when judges were confronted with claims of highly
traumatic events of a sexual nature (e.g., sexual assault), they performed very poorly at
determining the veracity of the report (overall accuracy 45.3%). Truthful accounts of
sexual victimization were correctly identified 49.6% of the time, relative to 41% of
fabricated narratives. Taken together, the rates of deception detection in samples of
"judges" are significantly lower than those found in this study.

Further, the design and sample size allowed for a statistically powerful, detailed
and realistic approach to credibility assessment (for review, see Granhag & Strömwall,
2004; and Masip et al., 2005). An additional strength of this research was the extended
six month period used to measure variations in narratives. Empirical studies that have
addressed both consistency and credibility have used relatively short (i.e., days or a
week) durations (e.g., Alonso-Quecuty, 1992; Granhag & Strömwall, 2002; Lampinen et
al., 2003; Suengas & Johnson, 1988), with few exceptions (e.g., Erdmann et al., 2004). In addition, this study also approached truthful and fabricated narratives of trauma prospectively. Narratives of recent traumatic experiences were obtained and compared to the generation of a fabricated claim. Participants’ memories were studied “forward” in time on three separate occasions, rather than recalling events (both truthful and deceptive) that happened in the distant past. Prospective analyses are being used more frequently in the clinical literature on trauma (e.g., Arbuthnott & Arbuthnott, 1999; Erickson et al., 2001; Harvey & Bryant, 1999), and are necessary for a controlled examination of consistency and credibility. Due to the legal reliance on testimonial consistency, the approach used in this dissertation provided empirical data and support for applied cognitive and forensic issues.

The need for reliable techniques in credibility assessment is important given the obvious difficulty of these judgments. One such difficulty is the tendency to “err on the side of caution”, particularly with claims of traumatic victimization. It is difficult to believe that an individual would falsely allege they were victimized (unless a motivation is clear). More often than not, individuals tend to believe “tales of trauma” when they may, in fact, be false. The effect of this “truth bias” has been well-documented (e.g., DePaulo, Charlton, Cooper, Lindsay, & Muhlenbruck, 1997; Merkelbach, 2004; Ruby & Brigham, 1997) and is particularly relevant in forensic settings. Future research should continue to address various factors associated with deception detection in truthful and fabricated narratives of victimization (e.g., individual differences, handedness, etc.), the reliability of content-based techniques (such as CBCA, RM, and MAP), and comparisons
of the ability of these techniques to discriminate truthful from fabricated traumas relative to judge samples.

In conclusion, this dissertation has demonstrated that the features of truthful and fabricated narratives of traumatic victimization differ as a function of narrative type and time. In general, truthful traumas contained more detail, contextual information, emotional details, and were considered more plausible relative to fabricated traumas. The details of truthful narratives also were more consistent than fabricated narratives over an extended six-month interval, although levels of consistency decreased over time. Fabricated symptoms were inflated and also may serve as indicators of possible deception, particularly in assessments of post-traumatic stress disorder. The findings of this dissertation indicate that the Memory Assessment Procedure has utility as a tool for deception detection in truthful and fabricated narratives, demonstrating similar levels of discrimination as established techniques such as CBCA and RM. Further research on MAP criteria and their application to statement analysis is warranted. The MAP is a comprehensive approach to memory assessment, thus the use of additional criteria drawn from previous approaches may aid in enhancing its discriminative ability. Ultimately, "there is no single piece of information that invariably marks a memory as an accurate reflection of the past" (Johnson & Raye, 1998, p. 137). As recommended by Porter, Campbell, et al. (2003), credibility assessments should always consider the following: (1) the context in which the report was made (i.e., compensation); (2) the content of the description (i.e., plausibility, details); (3) individual characteristics and trauma profiles (i.e., personality features, exaggerated symptoms); and (4) whether corroboration exists.
Use of the MAP in conjunction with other measures of memory features and symptomatology conforms to these recommendations and can assist in determining the veracity of a report. While further research on truthful and fabricated narratives of trauma is certainly needed, the current findings provide useful information for pinpointing possible false allegations and differentiating fact from fiction.
References


205


Dunnett, S. (1862). An essay concerning the philosophy of memory and the laws which govern it: Also, of the nature and power of thought. Brantford: Samuel Dunnett.


Appendix A: Recruitment Advertisement

Research Participant Contact Form
Academic Year 2003/2004

Code Number: 2004-876 REB Date Approved: _______________

Staff Member Responsible: Dr. Steve Porter Telephone No.: 494-6934 (Dr. Porter)
494-8093 (K. Peace)

Contact Person’s email: kpeace@dal.ca (preferred method of contact)

Experimenter(s): Kristine Peace, Ph.D. Student

Duration of Experiment: 1.5-2 hours during each phase (3 phases)

Description of the Research Project

Study Title: Characteristics of True and Fabricated Traumatic Events Over Time

Purpose of the research project: The objective of this research is to determine how memory characteristics and traumatic stress are related to real and falsified reports of traumatic or stressful life experiences over time.

Who can participate: Anyone can participate, provided they have experienced some sort of negative, traumatic, or stressful event (e.g., death of a loved one, criminal victimization, accident, serious injury, theft, etc.) since the age of 16, are comfortable and willing to provide a written description of this event, are registered in a course that allows credit points, and have not previously participated in the study. We are interested in all types of stressful or traumatic emotional events that you may have experienced.

What you will be asked to do: You will be asked to provide a complete written description of a true traumatic or stressful event that has happened to you since the age of 16, as well as generate a falsified written account of a traumatic event that never happened to you. You will be requested to fill out a form asking your age, sex, and provide some indication of your past history of negative experiences (e.g., number of experiences, overall level of stress). You will also be asked to fill out several questionnaires relating to memory, emotional stress, and self-perceptions after providing each written narrative. You may choose not to respond to some items if you are uncomfortable doing so. There are three phases of the experiment (separated by three months each) and you will be asked to complete the same questionnaires at each phase.

Possible risks and discomforts: Participants will be asked to recall a traumatic or stressful autobiographical memory, which may cause psychological distress in some individuals. Participants may withdraw from the study at any time if they no longer wish to participate. However, participants get to choose the negative experiences that they are comfortable talking about, so the risk of distress is minimized. Further psychological referral resources will be available if
necessary. Previous studies have been conducted at Dalhousie and other universities, and there were no psychological difficulties reported by the participants when recalling and writing about traumatic experiences.

Compensation: Participants will be awarded either two credit points (e.g., one credit point per hour of participation, with an anticipated study duration of one-and-a-half to two hours) or a small cash sum of $5.00 at the end of each phase of the experiment. Thus, if you are not enrolled in a credit point course when completing the follow-up phases of the study, you may choose the cash option. You can earn up to a total of 6 credit points (but a maximum of 4 per course, where available), or $15.00 cash.

It is important that you write down the code number of the experiment, name of the staff member responsible, telephone number and email, and keep a record of this information.

If you wish to participate in this experiment, please contact the experimenter.
Appendix B: Informed Consent Form

INFORMED CONSENT FORM

TITLE: A COMPARISON OF TRUTHFUL AND FABRICATED EVENTS OVER TIME

LOCAL PRINCIPAL INVESTIGATOR: Kristine A. Peace, B.A., M.Sc., Psychology Department, Dalhousie University, Halifax, Nova Scotia, B3H 4J1, ph. 902-494-8093, fax 902-494-6585, kpeace@dal.ca.

DEGREE PROGRAM: Doctoral Candidate in Experimental (Forensic) Psychology

SUPERVISOR: Stephen Porter, Ph.D., Psychology Department, Dalhousie University, Halifax, Nova Scotia, B3H 4J1, ph. 902-494-6934, fax 902-494-6585, sporter@dal.ca.

CONTACT PERSON: If you have any questions or concerns about this study, or require further information, please do not hesitate to contact Kristine Peace (494-8093, kpeace@dal.ca) or Dr. Stephen Porter (494-6934, sporter@dal.ca).

We invite you to take part in a research study being conducted by Kristine Peace who is a graduate student at Dalhousie University as part of her doctoral dissertation. Your participation in this study is voluntary and you may withdraw from this study at any time without any negative repercussions. The study is described below. This description tells you about the risks, inconvenience, or discomfort which you might experience. Participating in this study will give you the option of obtaining bonus credit points or a small cash compensation for each session, and will help the researchers learn things that will benefit others. You should discuss any questions you have about this study with Kristine Peace (principal investigator) or Stephen Porter (supervisor).

PURPOSE OF THE STUDY:
The objective of this research is to determine how memory characteristics and traumatic stress are related to real and falsified reports of traumatic or stressful life experiences over an extended period of time. The researchers are trying to better understand how memory is influenced by traumatic or negative experiences, and how memories for emotional events (both truthful and fabricated) are affected by the passage of time.

STUDY DESIGN:
This study uses a longitudinal design, with participants being involved in three phases of the experiment. Phase 1 is the initial appointment, Phase 2 occurs three months later, and Phase 3 occurs six months later. All participants from Phase 1 will be re-contacted for the follow-up phases to book appointments that are convenient for them. Participants will be given additional information and fully debriefed once the study is completed.
WHO CAN PARTICIPATE IN THE STUDY:
Anyone can participate, provided they have experienced some sort of negative, traumatic, or stressful event (e.g., death of a loved one, criminal victimization, accident, serious injury, theft, mugging, etc.) since the age of 16, are comfortable and willing to provide a written description of this event, and have not previously participated in Phase 1 of the study. We are interested in all types of stressful or traumatic emotional events that you may have experienced. If you are enrolled in a course that currently allows credit points, you will have the option of obtaining 2 credit points for each phase of this experiment, or you may opt for the cash sum of $5.00.

WHO WILL BE CONDUCTING THE RESEARCH:
The primary investigator, Kristine Peace, is conducting this study and will be involved with running participants in each phase of the experiment. You will also have contact with a research volunteer who will administer all the questionnaire packets during your experimental sessions.

WHAT YOU WILL BE ASKED TO DO:
You will be asked to provide a complete written description of a true traumatic or stressful event that has happened to you since the age of 16, as well as generate a falsified written account of a traumatic event that never happened to you. You will be requested to fill out a form asking your age, sex, and provide some indication of your past history of negative experiences (e.g., number of experiences, overall level of stress). You will also be asked to fill out several questionnaires relating to memory, emotional stress, and self-perceptions after providing each written narrative. You also may choose to omit any item on the questionnaires that you are not comfortable answering, without loss of compensation or course credit. Some participants will receive a trauma fact sheet prior to generating their fabricated trauma report, giving information on how victims of criminal or non-criminal trauma tend to respond on questionnaires. You will fill out the questionnaires individually or in a group setting. Completion of the questionnaires will take approximately 1.5 – 2 hours. For the second and third phases of this experiment (after three months and after six months, respectively), you will be asked to describe both your true and fabricated traumatic event in the same format as Phase 1, and complete the same questionnaires with respect to each event. The time period of this study (all three phases) will allow the entire experiment to be completed within one academic year, in order to be convenient for students.

Note: You will also have the option of giving permission for your data (e.g., transcribed memory reports and symptom profiles) to be used in future research studies. This may involve general re-analysis of the data and examination of the effect of different variables using the original information collected. In addition, studies involving the detection of true and false trauma, using samples of students and legal professionals, may be conducted. It is important that research be done on victim credibility, in order to determine how legal professionals decide (and how accurate they are) that a victim is telling the truth or lying. All names and any information that is highly personal or identifying (e.g., specific locations, names) will be removed from the transcripts and pseudonyms will be used.

POSSIBLE RISKS AND DISCOMFORTS & BENEFITS:
There are no physical risks or discomforts associated with this study. Participants will be asked to recall a traumatic or stressful autobiographical memory, which may cause mild psychological distress in some individuals. Participants may withdraw from the study at any time if they no longer wish to participate. However, participants get to choose the negative experiences that they are comfortable talking about, so the risk of distress is minimized. Further psychological referral resources can be obtained, if necessary, through the Dalhousie University Counseling Services (Room 409, Student Union Building, 6136 University Avenue, 494-2081). This service offers crisis counseling, and appointments can be made by phone or drop-in in the rare event of a crisis situation. These services are available to all Dalhousie students at their request (no special
arrangement has been made for this study). Any information you disclose to a counselor is not a part of this study, and is strictly confidential between you and a counselor. This service is open from 9:00am-8:30pm (T-W) and 9:00am-5:00pm (M-Th-F) during October to March, and open from 9:00am-5pm (M-T-W-Th-F) during April to September. Data collection from all participants will occur during the hours of operation of the Counseling services, so that these services will be available to you during and following the study if needed. Previous studies have been conducted at Dalhousie and other universities that have involved the recall of traumatic events (e.g., Peace & Porter, 2004). Finally, you will be offered credit points or financial compensation for each phase of the experiment, which will provide a small benefit for participants either academically or financially. There is a remote possibility that reports of genuine criminal victimization that are currently under investigation may be subject to subpoena by legal professionals. In this case, the researcher would be forced to disclose your records, but only the information relevant to the investigation (e.g., real memory report & real symptoms). Although this situation is possible, it is highly unlikely that this would occur.

COMPENSATION/EXPENSE REIMBURSEMENT:
All participants in this study will have two options available for compensation for their time and involvement in this research. Participants can choose whether they would prefer a credit point allotment (2 points towards any course within the Psychology Department offering credit points, up to a maximum allotment of 4 points per course where allowed, as per departmental guidelines) or a cash sum of $5.00. These options are available at each phase of the experiment. For example, a participant may be taking a 1000 or 2000-level psychology course in the Fall term that allows them to obtain 4 points total. In this case, participants in this experiment can choose the credit points at Time 1 and Time 2 (which totals 4 points) and apply them towards their psychology course (if these phases are completed by the end of the Fall course), and get a cash sum ($5.00) at Time 3. If they are in a different psychology course at Time 3 (e.g., in the Winter term) that allows credit points, they could get credit points for the final phase of the experiment as well, and apply the 2 points earned at Time 3 to this other course. This decision is made by participants at each phase of the experiment, and depends upon the availability of the credit point option for each student. Cash will be awarded when credit points are not available.

CONFIDENTIALITY & ANONYMITY:

Anonymity: All information collected from this study will be maintained in an anonymous manner. All participants will be assigned a Participant Identification number, and informed consent forms will be kept on file separately from the memory survey packets so that responses cannot be associated with individual participants. Once each memory report has been transcribed, any significant names or dates that may lead to identification of the participant will be removed from the file. Only the primary investigator (Kristine Peace) will have access to your contact information so that the follow-up phases of this experiment can be scheduled, and this will be kept in a secure location. Individual participants will not be identified in any reports or publications derived from this research project.

Confidentiality: Once the data is collected, all information (e.g., memory reports, symptom profiles) will be coded and scored, and stored as numerical data in a statistical database. Memory reports will also be transcribed to avoid the possibility of identification of handwriting, and these will be password protected files with restricted access. Original copies of the information completed by the participant will be stored in files for the duration of the study, due to the longitudinal component of the experiment, and a Master copy of contact information will be securely maintained by the primary investigator only. If you are currently in psychological counseling pertaining to the genuine event you choose to discuss in this research, and would like copies of your data to give to your psychologist, you can request these and consent to their release. The researcher will only be authorized to release these records (without identifying

248
information) to you, in which you can give them to your counselor if you wish. All files will be
categorized by the participant's ID number in order to maintain anonymity. Files will be kept in a
locked filing cabinet in the Porter Forensic Lab (Rooms 3409 and 3412). The only individuals
who will be able to access these files are the principal investigator, supervisor, and research
assistants who have been screened and have signed confidentiality forms assuring that they will
not disclose the content of file information (even though no personally identifying information
will be associated with the data). The only limit to confidentiality arises in the case of subpoena
(discussed in the Risks section of this consent form). If you consent to having your data used in
future research studies involving (1) re-analysis of the original data set; and/or (2) detection of
true and false trauma in student and legal samples, no identifying information will be used in the
transcripts and locations/names/events may be modified to further assure your anonymity and
confidentiality. Written narratives that contain highly specific details that may render the
individual identifiable to legal personnel (if a criminal victimization) will not be used in future
research.

Data Retention: Upon completion of this study, including the follow up phases, all data
that was completed in handwritten format will have been transcribed and coded. All material will
be reviewed to assure no personally identifying information is contained in the files. Original
handwritten memory reports will be destroyed at this point, being shredded and discarded as per
University Policy for confidential material. The transcribed and coded data will be securely
maintained on file for a duration of five years, in locked filing cabinets, as required by the
University Policy on Research Integrity.

QUESTIONS:
If you have any questions or concerns regarding this study that may affect your decision to
participate, or anything mentioned in this informed consent form, please feel free to discuss them
with the primary investigator or the research assistant prior to your participation in this
experiment. You can also ask questions at any time throughout the duration of the experiment.
You may also contact the persons listed above. Please note that the interviewer will inform you
of any new information not contained in this consent form that may influence your decision to
participate in the study. You are also entitled to receive a copy of this consent form for your
records if you would like one.

PROBLEMS OR CONCERNS:
In the event that you have any difficulties with, or wish to voice concern about, any aspect of
your participation in this study, you may contact Patricia Lindley, Director of Dalhousie
University's Office of Human Research Ethics Administration for assistance: (902) 494-1462,
patricia.lindley@dal.ca.

CONSENT SIGNATURES:
Please read each of the statements below, and indicate your consent (if applicable) by signing on
the appropriate lines after each statement:

(1) "I have read the explanation about this study. I have been given the opportunity to discuss it
and my questions have been answered to my satisfaction. I hereby consent to take part in this
study. However I realize that my participation is voluntary and that I am free to withdraw from
the study at any time".

PARTICIPANT: __________________________ DATE: ____________________

249
(2) "I am aware that the design of this study involves my participation in future phases of this study after three months (Time 2) and six months (Time 3). I hereby consent to being re-contacted for possible participation in follow-up phases of this study"

PARTICIPANT: ______________________________ DATE: ______________

(3) "I consent to having the data collected from my participation in this experiment used for further analysis of the information collected (e.g., a more detailed examination of the written narratives over time in another experiment). I am aware that my data will be coded in such a way that I cannot be identified from the information, and consent to its use in this respect"

PARTICIPANT: ______________________________ DATE: ______________

(4) "I consent to having the data collected from my participation in this experiment used in a future research study involving the detection of true and false traumas by students and legal personnel. I am aware that my data, with all possible identifying information such as places, dates, and names altered (and my own name not associated with the data), will be potentially shown to students, police officers, judges, and lawyers. I am aware that data will be coded and presented in such a way that I cannot be identified from the information, and consent to its use in this respect"

PARTICIPANT: ______________________________ DATE: ______________

RESEARCHER: ______________________________ DATE: ______________
Appendix D: Truthful Trauma Instructions (TT)

*** PLEASE READ BEFORE PROCEEDING ***

MEMORY SURVEY INSTRUCTIONS

REAL TRAUMA

In this section of the experiment, please write a description of a traumatic or stressful event that has happened to you since the age of 16. You can choose whatever event you are willing to discuss. Please remember that you may withdraw at any time.

We are interested in emotional memories for all types of events that are considered traumatic. We are especially interested in memories for traumatic experiences that involve criminal victimization (e.g., assault, sexual violence, theft). If you have been criminally victimized and you are comfortable and willing to discuss this event, please provide a description of this event and answer the questionnaires in this section of the experiment reflecting your feelings about the criminal victimization.

If you have not experienced an event involving criminal victimization, we are also interested in serious traumatic events (e.g., car accident, serious injury). If you have experienced an event of this type, and are comfortable and willing to discuss this event, please provide a description of this event and answer the questionnaires in this section.

If you have not experienced a serious traumatic event, we are also interested in a negative or stressful event in your life since the age of 16. If you are comfortable and willing to discuss this event, please provide a description of this event and answer the questionnaires in this section.

Please complete the written memory description and the questionnaires in this section according to the experience you have chosen. Try to answer all the questions completely and accurately.

TT
Appendix E: Fabricated Trauma Instructions (FT)

*** PLEASE READ BEFORE PROCEEDING ***

MEMORY SURVEY INSTRUCTIONS

FAKE TRAUMA

In this section of the experiment, please write a description of a traumatic event that has never happened to you. This event could be a report of some form of criminal (e.g., assault, robbery) or non-criminal (e.g., accident, injury, death of a loved one) form of victimization. Please choose a category of traumatic event that is different from your real traumatic experience. Please complete both your written description and associated questionnaires as if this "fake trauma" had happened to you, and make your responses as convincing as possible. Your goal in this section of the experiment is to produce a believable (but fabricated) traumatic memory report, that legal professionals would judge to be credible.
Appendix F: Written Narrative Report Form (TT)

ID # ________

Written Narrative Memory Report

TT

Please take a few moments and think back to the most traumatic or stressful life event that you have experienced since the age of 16 years old. You may choose whatever event you are willing to discuss. Please remember that you may withdraw at any time. However, we are interested in emotional memories for all types of events that are considered traumatic, including criminal victimization (e.g., assault, sexual violence, theft), serious traumas (e.g., car accident, serious injury), and other stressful/negative life events (e.g., end of relationship, loss of job). Choose a specific event, as opposed to a series of events or a drawn out traumatic period. Next, take your time and report everything that you can remember about the incident. Make sure to leave nothing out. Start at the beginning of the incident and give a complete description of the entire incident. You may use the back of this page as well. If you require more space, please continue your description on a sheet of paper and attach it to this booklet after this page. Please be sure to include your age at the time of the incident.
Appendix G: Written Narrative Report Form (FT)

ID # __________  
Written Narrative Memory Report  
FT

Please take a few moments and make up a memory report for a traumatic event that has NOT happened to you in real life. Create a fictitious report of what you “remember” about the event, and write out your memory report as if this event has occurred (since the age of 16). Choose a specific event, as opposed to a series of events or drawn out traumatic period. Next, take your time and report everything that you can remember about the incident. Make sure to leave nothing out. Start at the beginning of the incident and give a complete description of the entire incident. You may use the back of this page as well. If you require more space, please continue your description on a sheet of paper and attach it to this booklet after this page. Please be sure to include your age at the time of the incident.
Appendix H: Consistency Questionnaire (CQ)

ID ______

CQ

This questionnaire is designed to examine details that you recall about your traumatic or negative experience. Please complete these questions about the event you just wrote about to the best of your ability. If you do not remember a particular item, please indicate this. Try to be specific in your responses, and focus in on the specific incident discussed.

1. What were you wearing when the event occurred?

2. What was the weather like outside when the event occurred?

3. What was the location(s) of the event?

4. What time of day did the event occur (start)?

5. How were you feeling right before the event?

6. How many people (other than yourself) were present when the event occurred?

7. What day of the week did the event happen on?

8. What was the duration of the event (e.g., how long did it last)?

9. Did you have any personal belongings (e.g., wallet, purse, backpack, etc.) with you, and if so, what were they?

10. How were you feeling during the event?

11. What was the calendar date of the event?
12. Do you remember any particular smells during the event, and if so, what were they?

13. Are there any visual details during the event that stick out in your mind (if yes, describe)?

14. Were there any landmarks in the environment that you associate with the event (e.g., tree in a park, location in house), and if so, what were they?

15. How were you feeling immediately after the event?

16. Was there a weapon used during the event? If so, what was it?

17. Are there any noises or things that were said that stick out to you during the event? If so, please describe.

18. Are there any physical sensations that occurred during the event (e.g., sick to stomach, touch sensations, pain)?

19. What was your relationship with the other person(s), if any, involved in the event?

20. How old were you when the event occurred?

21. What were you doing immediately before the event happened?

22. Were you physically injured during the event, and if so, please describe?

23. What did you do immediately after the event?
Appendix I: CQ Coding Scheme

Consistency Coding Scheme

Following each memory report, participants were asked a set of 23 direct questions referring to general event categories (e.g., weather, time of day of event, emotional state before and after event, etc.). These questions can be found in Appendix K (preceding pages). Consistency questions will be administered at all experimental phases to allow for an examination of consistency for both real and fabricated traumatic memories across time. Responses to these questions will then be coded for their overall level of consistency between T1, T2, and T3. They will be first coded by the primary investigator, and then by a trained graduate student in psychology, in order to determine intercoder reliability. The coding scheme allows for one of four ratings: inconsistent, partially inconsistent, partially consistent, and consistent. These are defined as follows:

Inconsistent. Responses that would be coded inconsistent involve those that differ at T1 and T2, and contain one or more contradictions of facts. See the examples of these below:

Example 1 – What was your emotional state right before the event?
T1: Terrified.
T2: Kind of detached.

Example 2 – Do you remember what the duration of the event was?
T1: 20 minutes.
T2: Maybe 2 minutes.

Partially Inconsistent. Responses in this category generally refer to those that are inconsistent due to an omission at T1 or T2. For example, a participant provides a response at T1 to what time of day the event occurred, and at T2 says that they do not remember. This is inconsistent in nature as the responses differ, however there is no contradiction of facts presented. Another case where a response would be considered partially inconsistent, is if participants provided a response at either of the experimental periods which had one component correct, however then added additional information that contradicted what was said at the opposing time period. See the examples of these below:

Example 1 – Do you remember what day of the week this event occurred on?
T1: No, I don’t remember.
T2: Wednesday.

Example 2 – Do you remember what the weather was like when the event occurred?
T1: It was warm out, clear night.
T2: Warm, but overcast, felt like it might rain.

Partially Consistent. Responses coded as partially consistent involve ones that have at least one component the same but have additional information added that was not provided previously (but is not a contradiction of facts). Another case in which a response might be coded as partially consistent is when participants make statements about time or date that are “similar” (e.g., within a half an hour, or a day apart) but not identical to the previous response. See the examples of these below:
Example 1 – Did you have any personal belongings with you when the event occurred?
T1: Book and house keys.
T2: Book.

Example 2 – Do you remember what time of day the event occurred?
T1: 7:00am.
T2: Early morning.

Example 3 – Do you remember what the calendar date of the event was?
T1: July 2, 2001
T2: July 3, 2001
(Note: Two components in this response were correct – the month of July and the year 2001, and the actual date is “similar” as only a day apart but not identical).

Consistent. Responses in this category refer to those that are exactly the same, or the semantic meaning of the response means the same at both times (e.g., especially for questions involving emotional labeling), or where two or more details of the responses are the same and there are no contradictions in additional information. See examples of these below:

Example 1 – Do you remember what day of the week the event occurred on?
T1: Sunday night.
T2: Sunday night.

Example 2 – What was your emotional state just after the event?
T1: Angry, upset, didn’t believe it.
T2: Felt unreal, mad, upset.
(Note: Angry and mad refer to the same emotion, thus are coded as consistent. Similarly, statements such as “didn’t believe it” and “unreal” refer to the same sentiment).
Appendix J: Memory Assessment Procedure (MAP)

ID #: Memory Assessment Procedure (MAP) Criteria

Memory Type (please circle): Truthful Traumatic Fabricated Traumatic
Trauma Type (please circle): Criminal Victimization Non-Criminal Victimization

Amount of Detail:
This criterion refers to how much detail is contained within a memory report. Each distinctive piece of information offered by the participant is scored as one point (e.g., “I assaulted a man in the woods behind our house” contains four details).

Emotional Components: (Self) (Other)
All emotional words in each of the memory descriptions are counted. Examples of emotional words include upset, scared, fear, afraid, angry, sad, guilty, shame, excited, aroused, nervous, anxious, happy, proud etc. Also included are any words that follow the words “feel” or “felt” (e.g., “I felt wonderful” or “I remember feeling terrible”). Any words that convey emotion should be included. The emotional component is divided between the number of emotional references related to the self (e.g., I felt scared; I was very excited) and those relating to others (e.g., she looked scared; I could tell he was happy).

Time and Place Details:
All details mentioned in the memory report relating to the time of the event or surrounding experiences (e.g., 8am, morning, evening) and the places in which the experience(s) occurred (e.g., living room, home, grocery store, street) should be counted. These give the reader examples and clarification of when and where the event was taking place, and are primarily descriptive of the scene.

Coherence:
This measure refers to how a memory report hangs together and follows a logical sequence with a beginning, middle, and end reported in that order. That is, it reflects how logical and sensible the report is to the listener or coder. The coder rates the memory report’s coherence on a 1-7 point scale with anchors of 1 (not at all coherent) to 7 (extremely coherent).

Relevancy:
Relevancy was defined as the extent to which the information provided by participants is relevant to the event being discussed. For example, when a participant is describing an assault they experienced, and proceeds to detail the assault and experiences within the content of the assault, relevant information has been provided. When (s)he responds by explaining in detail why they were doing a particular activity and then only briefly mentions the assault, much irrelevant information is provided. Coders rate the relevance of information provided on a scale of 1-7 (not at all relevant details to extremely relevant details provided).

Plausibility:
Plausibility is defined as the extent to which the information provided is realistic, logical, and contains components that are likely to occur in real-life. The memory report would be rated as higher in plausibility if it appears to be a valid, likely, or an acceptable event. For example, when a participant is describing getting an electric shock and they state that they wandered unknowingly into a high voltage area, this information is low in plausibility as these areas are fenced off. Coders rate the plausibility of information provided on a scale of 1-7 (not at all plausible to extremely plausible details).

Number of Words:
To obtain a measure of the length of memory accounts, the total number of words in each free narrative part of the interview is counted.

261
Appendix K: Emotional Memory Survey (EMS)

ID# ______

Emotional Memory Survey (TT)

*Please answer the questions below about the real traumatic or stressful life event you just described by circling the answer that best applies to your experience.*

1. **If a visual image comes with the memory for this event, can you see yourself in the memory (i.e., can you see yourself acting and/or your face in the memory)?** ...
   
   a. I can *always* see myself in the memory
   b. I can *never* see myself in the memory
   c. The memory changes so that I can see myself in the memory image only *some of the time*.

2. **Please indicate the how much anxiety or stress, if any, you experienced at the time of this incident**...

   1. very little  
   2  3  4  5  6  7  
   4. moderate amount
   5. extreme amount

3. **Please rate your degree of intoxication from drugs or alcohol when the event occurred (i.e., how high or drunk were you)?** ...

   1. not at all intoxicated
   2  3  4  5  6  7  
   4. somewhat intoxicated
   5. extremely intoxicated

   *(i.e., no drugs or alcohol used)*

4. **Please indicate how often you talked about the event with other people since it happened (on average)...**

   1. never
   2  3  4  5  6  7  
   4. some times (few times/year)
   5. many times (7+ times/year)

5. **Please indicate how often you have thought about the event since it happened**...

   1. never
   2  3  4  5  6  7  
   4. some times (few times/year)
   5. many times (7+ times/year)

6. **Please indicate how vivid and clear your memory is for this event...**

   1. not at all vivid/clear
   2  3  4  5  6  7  
   4. moderately vivid/clear
   5. completely vivid/clear
7. Compared to your other memories, what is the quality of the memory for this event like...

1 2 3 4 5 6 7
very poor average very good

8. Compared to your other memories, how coherent is your memory for this event?

1 2 3 4 5 6 7
not at all moderately very
cohesive cohesive cohesive

9. Please indicate how many sensory details (e.g., visual images, smells, etc.) are associated with your memory for this event ...

1 2 3 4 5 6 7
very few moderate many sensory
amount details

10. Please respond to each of the following questions by circling yes’ or ‘no’ to indicate whether the statement has been a part of your memory for the event...

a. Can you see things in the memory? Yes No
b. Can you hear things in the memory? Yes No
c. Can you smell things in the memory? Yes No
d. Can you physically feel things in the memory? Yes No
e. Can you taste things in the memory? Yes No

11. Please indicate how often you have “turned off” your emotions when thinking about or discussing this event ...

1 2 3 4 5 6 7
never some times many times
(few times/year) (7+ times/year)

12. Please indicate how often you become emotional (experiencing positive or negative feelings) when thinking about or discussing this event ...

1 2 3 4 5 6 7
never some times many times
(few times/year) (7+ times/year)

263
13. Please indicate how often you feel detached or distanced from your feelings when you think or talk about this event ...

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>never</td>
<td>some times</td>
<td>many times</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(few times/year)</td>
<td>(7+ times/year)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

14. How do you feel when you talk about this experience with others? ...

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>happy</td>
<td>indifferent</td>
<td>upset</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

15. Please rate how emotionally intense your memory is for this event ...

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>not very emotional</td>
<td>moderately emotional</td>
<td>highly emotional</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

16. How confident are you that your memory for this event is accurate?

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>not at all confident</td>
<td>moderately confident</td>
<td>very confident</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

17. How credible do you think others would find your memory for this event?

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>not at all credible</td>
<td>moderately credible</td>
<td>very credible</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

18. Was there ever a period of time where you could not remember the event?

a. Yes
b. No (i.e., I always remembered the event)

*If Yes, what was longest period of time that you *COULD NOT* remember it? Please indicate the length of period of time and how old you were during this time period...

✓ Longest time period: ________ (e.g., minutes, days, weeks, years & how many?)
✓ Age: ________
19. To your knowledge, has your memory for this event changed over time?
   a. Yes.__________

      *If yes, please explain how and why you think it changed:*

      ____________________________________________________________
      ____________________________________________________________
      ____________________________________________________________

   b. No.

20. Has your memory for this event ever spontaneously “popped” into your mind?
   a. Yes. ______

      *If Yes, please explain what you were doing at the time:*

      ____________________________________________________________
      ____________________________________________________________
      ____________________________________________________________

   b. No.
Appendix L: Impact of Event Scale – Revised (IES-R)

**ID #: __________**

The Impact of Event Scale - Revised

Below is a list of difficulties people sometimes have after stressful life events. Please read each item, and then indicate how distressing each difficulty has been for you **DURING THE SEVEN DAYS FOLLOWING THE EVENT** with respect to ________________, how much were you distressed or bothered by these difficulties?

<table>
<thead>
<tr>
<th>1. Any reminder brought back feelings about it</th>
<th>Not at all</th>
<th>A little bit</th>
<th>Moderately</th>
<th>Quite a bit</th>
<th>Extremely</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. I had trouble staying asleep</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>3. Other things kept making me think about it</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4. I felt irritable and angry</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5. I avoided letting myself get upset when I thought about it or was reminded of it</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>6. I thought about it when I didn’t mean to</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>7. I felt as if it hadn’t happened or wasn’t real</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>8. I stayed away from reminders about it</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>9. Pictures about it popped into my mind</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>10. I was jumpy and easily startled</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>11. I tried not to think about it</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>12. I was aware that I still had a lot of feelings about it, but I didn’t deal with them</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>13. My feelings about it were kind of numb</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>14. I found myself acting or feeling as though I was back at that time</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>15. I had trouble falling asleep</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>16. I had waves of strong feelings about it</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>17. I tried to remove it from my memory</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>18. I had trouble concentrating</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>19. Reminders of it caused me to have physical reactions, such as sweating, trouble breathing, nausea,</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

266
<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>20. I had dreams about it</strong></td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td><strong>21. I felt watchful or on-guard</strong></td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td><strong>22. I tried not to talk about it</strong></td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

(Weiss & Marmar, 1997)
Appendix M: Post-traumatic Checklist (PCL)

ID#: ____________

PCL

The following is a list of experiences that people may have after a traumatic or stressful life event. Please read over the following list, and indicate how often or frequently these experiences happened to in the month following the event. Please rate each item based on the following scale:

1 = not at all; 2 = rarely; 3 = sometimes; 4 = often; 5 = all the time

<table>
<thead>
<tr>
<th>ITEM</th>
<th>RATING</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Have you had recurrent or intrusive distressing recollections of the event, including images, thoughts, or perceptions?</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>2. Have you had recurrent distressing dreams about the event?</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>3. Have you experienced acting or feeling as if the traumatic event were recurring (including a sense of reliving the experience, illusions, hallucinations, and dissociative flashback episodes, including those that occur on awakening or when intoxicated)?</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>4. Have you had intense psychological distress at exposure to internal or external cues that symbolize or resemble an aspect of the traumatic event?</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>5. Have you had physiological reactivity on exposure to internal or external cues that symbolize or resemble an aspect of the traumatic event?</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>6. Have you tried to avoid thoughts, feelings, or conversations associated with the trauma?</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>7. Have you tried to avoid activities, places, or people that arouse recollections of the trauma?</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>8. Have you experienced an inability to recall an important aspect of the trauma?</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>9. Have you experienced a markedly diminished interest or participation in significant activities?</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>10. Have you experienced feelings of detachment or estrangement from others?</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>11. Have you experienced a restricted range of affect (e.g., unable to have loving feelings)?</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>12. Have you experienced a sense of foreshortened future (e.g., do not expect to have career, marriage, children, or normal life span)?</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>13. Have you had difficulty falling or staying asleep?</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>14. Have you been irritable or experienced outbursts of anger?</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>15. Have you had difficulty concentrating?</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>16. Have you been hypervigilant (e.g., oversensitive to loud noises, easily startled or scared)?</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>17. Have you experienced an exaggerated startle response?</td>
<td>1 2 3 4 5</td>
</tr>
</tbody>
</table>

(Weathers, Litz, Huska & Keane, 1994)
Appendix N: Trauma Symptom Inventory (TSI)

Please read all of these instructions carefully before beginning. Mark all of your answers on the accompanying answer sheet and write only where indicated. DO NOT write in this item booklet.

On the answer sheet, please write your name, the date, your age, your sex, and your race in the spaces provided.

This questionnaire contains 100 items describing experiences that may or may not have happened to you. Rate these items based on how you felt during the SIX MONTHS following the falsified traumatic or stressful event you have just described. Please rate these items based on how you would have felt if the event had happened to you.

Circle 0 if your answer is NEVER; it has not happened at all in the last 6 months. 0 1 2 3.
Circle 1 or 2 if it has happened in the last 6 months, but has not happened often. 0 1 2 3
Circle 3 if your answer is OFTEN; it has happened often in the last 6 months. 0 1 2 3

If you make a mistake or change your mind, DO NOT ERASE! Make an “X” through the incorrect response and then draw a circle around the correct response.

Please answer each item as honestly as you can. Be sure to answer every item. You can take as much time as you need to finish the TSI.
In the last 6 months, how often have you experienced:

1. Nightmares or bad dreams
2. Trying to forget about a bad time in your life
3. Irritability
4. Stopping yourself from thinking about the past
5. Getting angry about something that wasn’t very important
6. Feeling empty inside
7. Sadness
8. Flashbacks (sudden memories or images of upsetting things)
9. Not being satisfied with your sex life
10. Feeling like you were outside of your body
11. Lower back pain
12. Sudden disturbing memories when you were not expecting them
13. Wanting to cry
14. Not feeling happy
15. Becoming angry for little or no reason
16. Feeling like you don’t know who you really are
17. Feeling depressed
18. Having sex with someone you hardly knew
19. Thoughts or fantasies about hurting someone
20. Your mind going blank
21. Fainting
22. Periods of trembling or shaking
23. Pushing painful memories out of your mind
24. Not understanding why you did something
25. Threatening or attempting suicide
26. Feeling like you were watching yourself from far away
27. Feeling tense or “on edge”
28. Getting into trouble because of sex
29. Not feeling like your real self
30. Wishing you were dead
31. Worrying about things
32. Not being sure of what you want in life
33. Bad thoughts or feelings during sex
34. Being easily annoyed by other people
35. Starting arguments or picking fights to get your anger out
In the last 6 months, how often have you experienced:

36. Having sex or being sexual to keep from feeling lonely or sad
37. Getting angry when you didn’t want to
38. Not being able to feel your emotions
39. Confusion about your sexual feelings
40. Using drugs other than marijuana
41. Feeling jumpy
42. Absent-mindedness
43. Feeling paralyzed for minutes at a time
44. Needing other people to tell you what to do
45. Yelling or telling people off when you felt you shouldn’t have
46. Flirting or “coming on” to someone to get attention
47. Sexual thoughts or feelings when you thought you shouldn’t have them
48. Intentionally hurting yourself (for example, by scratching, cutting, or burning) even though you weren’t trying to commit suicide
49. Aches and pains
50. Sexual fantasies about being dominated or overpowered
51. High anxiety
52. Problems in your sexual relations with another person
53. Wishing you had more money
54. Nervousness
55. Getting confused about what you thought or believed
56. Feeling tired
57. Feeling mad or angry inside
58. Getting into trouble because of your drinking
59. Staying away from certain people or places because they reminded you of something
60. One side of your body going numb
61. Wishing you could stop thinking about sex
62. Suddenly remembering something upsetting from your past
63. Wanting to hit someone or something
64. Feeling hopeless
65. Hearing someone talk to you who wasn’t really there
66. Suddenly being reminded of something bad
67. Trying to block out certain memories
68. Sexual problems
69. Using sex to feel powerful or important
70. Violent dreams
In the last 6 months, how often have you experienced:

71. Acting “sexy” even though you didn’t really want sex
72. Just for a moment, seeing or hearing something upsetting that happened earlier in your life
73. Using sex to get love or attention
74. Frightening or upsetting thoughts popping into your mind
75. Getting your own feelings mixed up with someone else’s
76. Wanting to have sex with someone who you knew was bad for you
77. Feeling ashamed about your sexual feelings or behavior
78. Trying to keep from being alone
79. Losing your sense of taste
80. Your feelings or thoughts changing when you were with other people
81. Having sex that had to be kept a secret from other people
82. Worrying that someone is trying to steal your ideas
83. Not letting yourself feel bad about the past
84. Feeling like things weren’t real
85. Feeling like you were in a dream
86. Not eating or sleeping for 2 or more days
87. Trying not to have any feelings about something that once hurt you
88. Daydreaming
89. Trying not to think or talk about things in your life that were painful
90. Feeling like life wasn’t worth living
91. Being startled or frightened by sudden noises
92. Seeing people from the spirit world
93. Trouble controlling your temper
94. Being easily influenced by others
95. Wishing you didn’t have any sexual feelings
96. Wanting to set fire to a public building
97. Feeling afraid you might die or be injured
98. Feeling so depressed that you avoided people
99. Thinking that someone was reading your mind
100. Feeling worthless
Appendix O: Demographic Questionnaire

ID#: __________  

Demographic Questionnaire

Gender (circle one): MALE FEMALE

Age: ________________

Ethnic Background (please check one of the boxes below):

☐ Caucasian
☐ African Canadian
☐ Aboriginal
☐ Asian
☐ Other (specify):

Education Level – Year of Study (please check one of the boxes below):

☐ First year
☐ Second year
☐ Third year
☐ Fourth year
☐ Five or more years

Psychology Background – Number of Courses (please check one of the boxes below that indicates the number of psychology courses you have taken):

☐ One to three
☐ Four to nine
☐ Ten or more

Do you believe that people can repress memories for traumatic and painful life events and later recover them?

☐ Yes
☐ No

Have you ever recovered a memory of a past event that you could not remember previously?

☐ Yes
☐ No
If yes, please provide a brief description of the event and how you came to remember it:

_____________________________________________________________________________________
                                                                                       
                                                                                       
                                                                                       
                                                                                       
Do you believe that people can falsely recall events that never really occurred?

☐ Yes
☐ No

Have you ever recalled an event that you know did not happen to you? (e.g., have you ever recalled a false memory?)

☐ Yes
☐ No

If yes, please provide a brief description of the event and how old you were at the time:

_____________________________________________________________________________________
                                                                                       
                                                                                       
                                                                                       
                                                                                       
In your lifetime, how may traumatic or stressful events have you experienced?

☐ One to three
☐ Four to nine
☐ Ten or more
Out of the following categories, please indicate what types of traumatic events you have experienced in your lifetime (check all that apply, more than one category can apply to each trauma):

- Physical trauma (e.g., assault, severe injury)
- Sexual trauma (e.g., sexual assault, harassment)
- Psychological trauma (e.g., being degraded, name-calling)
- Financial trauma (e.g., stolen funds, financial hardship)

On average, how much traumatic stress have you experienced during your lifetime?

- Mild amounts of traumatic stress
- Moderate amounts of traumatic stress
- High/Severe amounts of traumatic stress

Have you ever been the victim of a crime?

- Yes
- No

If yes, please explain the type and number of times you have been victimized as well as what age you were at the time:

Type: ___________________________ Age: ________
Type: ___________________________ Age: ________
Type: ___________________________ Age: ________
Type: ___________________________ Age: ________
Type: ___________________________ Age: ________

Note: If you have experienced more than five criminal victimizations in your lifetime, please indicate the remaining types and ages on the back of this sheet, or ask the researcher for an additional piece of paper.

Have you ever had any problems with your memory?

- Yes
- No

If yes, please briefly explain:

__________________________________________________________________________
__________________________________________________________________________
__________________________________________________________________________
__________________________________________________________________________
Do you consider yourself to be a good “liar” or good at fooling other people?

☐ Yes
☐ No

Please explain what characteristics do you think make you a “good” or a “poor” liar:

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________
Appendix P: Debriefing Form

Debriefing Form

The study of traumatic and negative life experiences has thrived in the last few decades, with many researchers examining the qualitative and phenomenological characteristics of traumatic memories (e.g., Bernsten, Willert & Rubin, 2003). Empirical evidence suggests that memories for traumatic experiences differ from other types of emotional events in ways such as vantage point and number of details (Porter & Birt, 2001). However, there remains relatively little known about the characteristics of fabricated reports for trauma and victimization. Previous research has shown that fabricated reports of childhood experiences are qualitatively different from true experiences; being more detailed, vivid, and associated with higher ratings of confidence (Porter, Yuille & Lehman, 1999).

In addition, recent research examining low, moderate, and severe levels of traumatic stress demonstrated that levels of psychological impact (including levels of post-traumatic stress disorder symptoms) are associated with different memory characteristics. Specifically, severe traumatic memories were more detailed, vivid, and contained more sensory details than memories for moderate and low traumatic experiences (Peace & Porter, 2004). These findings support the idea that trauma does not impair memory processes, and may serve to enhance memory for these types of events.

In this study, information about symptoms that people experience after traumatic events may have been given to you at the beginning of the study. This was done in order to examine the effect that “knowledge” or training may have had on your ability to successfully fabricate trauma symptoms. Participants were given either factual training, or inflated training (which exaggerated the degree to which people endorse trauma symptoms). Attached is a copy of the factual training sheet that you may keep so that you will have correct and accurate information about responses to trauma.

To date, no previous studies have directly examined true and fabricated reports of victimization in adulthood. The purpose of this study is to determine if participants displaying different levels of traumatic stress (e.g., presence or absence of post-traumatic stress disorder symptoms) differ in their true and fabricated traumatic memory reports. Another important issue to consider is whether the experience of past traumatic events is associated with participants being able to formulate more credible, but fabricated accounts of a trauma, and whether these real and fabricated accounts can be differentiated based on the patterns of traumatic stress reported. For example, are individuals who have experienced severe or multiple traumatic events better at “faking trauma”? What types of symptoms do people report about traumatic events that never really happened? These results are directly related to forensic contexts where legal professionals are continually faced with the task of determining the credibility of reports of victimization.

In accordance with the Psychology Department Guidelines, you must correctly answer the following two questions to be awarded course credit:

1. How do fabricated reports of childhood experiences differ (qualitatively) from true experiences?

2. Identify two characteristics of severe traumatic memories that are different from less stressful traumatic experiences.

277
For further information about traumatic memories, please refer to the following sources:


Appendix Q: Criterion Descriptions

Assessment components for criterion-based content analysis (CBCA), reality monitoring (RM), and memory assessment procedure (MAP) techniques.

<table>
<thead>
<tr>
<th>Procedure/Criterion Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CBCA</strong></td>
</tr>
<tr>
<td>1. <em>Logical structure</em> – statements that are coherent and consistent; details form a logical account of a sequence of events.</td>
</tr>
<tr>
<td>2. <em>Unstructured production</em> – information is not provided in chronological time sequence; no underlying pattern/structure.</td>
</tr>
<tr>
<td>3. <em>Quantity of detail</em> – a significant amount (or abundance) of detail about the event.</td>
</tr>
<tr>
<td>4. <em>Contextual embedding</em> – statements that place event in its spatial and temporal context; references to time &amp; space.</td>
</tr>
<tr>
<td>5. <em>Description of interactions</em> – description of interrelated actions and reactions</td>
</tr>
<tr>
<td>6. <em>Reproduction of conversation</em> – speech reproduced in its original form</td>
</tr>
<tr>
<td>7. <em>Unexpected complications</em> – elements incorporated into statement that are unexpected or unforeseen</td>
</tr>
<tr>
<td>8. <em>Unusual details</em> – inclusion of detail that is not unrealistic but uncommon or unlikely to occur</td>
</tr>
<tr>
<td>9. <em>Peripheral details</em> – descriptions that are not central to the event</td>
</tr>
<tr>
<td>10. <em>Accurately reported details misunderstood</em> – actions and details that don’t make sense to witness but do to interviewer</td>
</tr>
<tr>
<td>11. <em>Related external associations</em> – reports of details that are related but not a part of event or immediate focus</td>
</tr>
<tr>
<td>12. <em>Subjective mental state</em> – description of feelings or thoughts at time of the event</td>
</tr>
<tr>
<td>13. <em>Attribution of perpetrator’s mental state</em> – description of perpetrator’s thoughts, feelings, or motives during event</td>
</tr>
<tr>
<td>14. <em>Spontaneous corrections</em> – corrections to statement made without prompting from interviewer</td>
</tr>
<tr>
<td>15. <em>Admitting lack of memory</em> – admitting that some parts of statement may not be accurate or cannot remember</td>
</tr>
<tr>
<td>16. <em>Raising doubts about own testimony</em> – objections against veracity of statement, acknowledging apparent contradictions</td>
</tr>
<tr>
<td>17. <em>Self-deprecation</em> – inclusion of unfavourable, self-incriminating details</td>
</tr>
<tr>
<td>18. <em>Pardoning perpetrator</em> – explanations or rationalizations for perpetrator’s behaviour</td>
</tr>
<tr>
<td>19. <em>Details characteristic of offence</em> – description of events typical for the type of experience under investigation</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>RM (MCQ/JMCQ)</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. <em>Clarity/vividness</em> – ratings of how dim/clear, colourful, and vague/vivid the memory is</td>
</tr>
<tr>
<td>2. <em>Sensory information</em> – details about visual, sound, smell, touch, and taste information during the event</td>
</tr>
<tr>
<td>3. <em>Spatial information</em> – details about the environment, specific locations, and spatial arrangements of persons/objects</td>
</tr>
<tr>
<td>4. <em>Time information</em> – memory for time, year, season, day, hour, duration of the event</td>
</tr>
<tr>
<td>5. <em>Reconstructability</em> – whether event is confusing/comprehensible, and details can be understood in context of event</td>
</tr>
<tr>
<td>6. <em>Emotion and feelings</em> – feelings and intensity of emotion at the time of the event and during recall</td>
</tr>
<tr>
<td>7. <em>Realism/plausibility</em> – contains a simple/complex or bizarre/realistic storyline</td>
</tr>
<tr>
<td>8. <em>Cognitive operations</em> – records of organization, elaboration, retrieval, “thoughts” about the memory</td>
</tr>
<tr>
<td>9. <em>Amount of detail</em> – ratings of how detailed memory is from sketchy to very detailed</td>
</tr>
<tr>
<td>10. <em>Familiarity</em> – is the setting of the event familiar or not</td>
</tr>
<tr>
<td>11. <em>Observer/participant perspective</em> – recalled from a spectator standpoint or a participant standpoint</td>
</tr>
<tr>
<td>12. <em>Consequentiality</em> – the implications of the event and what it reveals or says about the person</td>
</tr>
</tbody>
</table>

279
### MAP Objective

1. **Amount of detail & number of words** – quantity of unique information contained in statements and number of words
2. **Emotional components (self & other)** – number of emotional details about oneself and others involved
3. **Time & place details** – quantity of details relating to the context of the event including the time/duration and location
4. **Coherence** – rating of extent memory report hangs together and follows a logical sequence
5. **Relevant details** – rating of extent to which information provided is relevant to the event in question
6. **Plausibility** – rating of extent to which information is realistic, logical, and likely to occur in real-life

### MAP Subjective

7. **Memory perspective** – can participants see themselves in the memory (observer) or not (participant)
8. **Anxiety/stress** – level of anxiety or stress experienced at time of the event
9. **Talking about the event** – rating of frequency of talking about the event with others
10. **Thinking about the event** – rating of frequency of thoughts about the event
11. **Memory vividness** – how vivid and clear is the memory for this event
12. **Memory quality** – compared to other memories, the overall quality of memory for the event
13. **Memory coherence** – does the memory seem coherent and logical
14. **Emotional intensity** – level of emotional intensity (how strong of feelings) when recalling this event
15. **Confidence in memory accuracy** – level of confidence in the accuracy of the memory
16. **Perceived credibility** – how credible would others find the memory for the event
17. **Lack of memory** – was there ever a period of time where you could not remember the event
18. **Memory changes over time** – has the memory changed over time