THE LANGUAGE OF DECEIT:
ARE THERE RELIABLE VERBAL CLUES TO DECEPTION
IN THE INTERROGATION CONTEXT?

by

STEPHEN B. PORTER

B.Sc.(Hon.), Acadia University, 1992

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF MASTER OF ARTS
in

THE FACULTY OF GRADUATE STUDIES
(Department of Psychology)

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

September, 1994

© Stephen B. Porter, 1994
In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

(Signature)

Department of Psychology

The University of British Columbia
Vancouver, Canada

Date Sept. 1, 1994

DE-6 (2/88)
Abstract

In recent years, the need for enhanced methods of credibility assessment in criminal cases has become illuminated. Especially in cases of sexual assault, the words of the accused and complainant are often the sole evidence available to police. Consequently, researchers and practitioners have been searching for ways of differentiating truthful and deceptive accounts, focusing mainly on witnesses and victims. With its recent history, however, assessment based on verbal clues has been somewhat myopic and not well grounded in theory or integration. This thesis examined a general hypothesis, based upon conceptual information from a variety of perspectives, that reliable verbal indicators of deception exist in the interrogation situation. Sixty undergraduates were recruited for participation in research addressing "security effectiveness." Participants either committed a theft "to test the effectiveness of a new security guard" or carried out a similar but innocuous task. They were then asked to provide either: (1) a truthful alibi (2) a partially deceptive account (3) a completely false alibi or (4) a truthful confession regarding the theft to "an interviewer also hired for the purpose of investigating thefts." To increase motivation in the interview, honest and dishonest participants were offered a monetary incentive for convincing the interrogator of their veracity. The accounts
were then transcribed and examined for the presence of eighteen language variables. A multivariate analysis of variance (MANOVA) revealed a profile of three variables which significantly differentiated the truthful and deceptive accounts (amount of detail reported, coherence of the account, and admissions of lack of memory). For example, dishonest participants provided much less detail (although not fewer words) in relating an event than truthful participants describing a similar event. Implications for credibility assessment in forensic interrogations are discussed, emphasizing the need for establishing external validity with eclectic research strategies.
# TABLE OF CONTENTS

Abstract .......................................................... ii
Table of Contents ................................................. iv
List of Tables ..................................................... vi
List of Figures ...................................................... vii
Acknowledgment ..................................................... viii

Chapter 1.
INTRODUCTION ....................................................... 1

Chapter 2.
PERSPECTIVES ON MEMORY ACCURACY .......................... 9
2.1 The Undeutsch Hypothesis .................................. 9
2.2 Reality Monitoring .......................................... 17

Chapter 3.
PERSPECTIVES ON DECEPTION .................................... 19
3.1 Emotional - Motivational Approaches .................... 19
3.2 Cognitive Approaches ....................................... 23

Chapter 4.
THE PROPOSALS OF AVINOAM SAPIR .............................. 32
4.1 Ratio of Component Elements ............................... 32
4.2 Unnecessary Connectors ...................................... 33
4.3 Use of Pronouns .............................................. 34

Chapter 5.
SUMMARY OF THEORETICAL FRAMEWORK AND THE PRESENT STUDY .............................................. 36

Chapter 6.
METHOD .............................................................. 38
6.1 Participants .................................................... 38
6.2 Procedure ....................................................... 38
6.3 Instructional Manipulation ................................... 39
6.4 Interviewing .................................................... 41
6.5 Transcribing ..................................................... 43
6.6 Dependent Variables and Scoring Procedures ............ 43

Chapter 7.
RESULTS ............................................................. 50
7.1 Reliability Checks ............................................ 50
7.2 Preliminary Analyses ......................................... 51
7.3 Multivariate Analysis ......................................... 54
7.4 Relative Discriminating Ability of Criteria Showing Significant Effects ......................... 58
7.5 Followup Exploratory Analyses ............................. 60
Chapter 8.

Discussion ........................................... 66
8.1 Limitations of the Current Study .......... 79
8.2 Directions for Future Research .......... 84

References ............................................. 87

Appendix A: Debriefing Form ................. 96
Appendix B: Interrogation Protocol .......... 99
Appendix C: Standard Deviations for Table 1
Dependent Variables ......................... 103
List of Tables

Table 1: Mean Scores for the Four Conditions on All Dependent Variables .................. 53

Table 2: Discriminant Analysis of Criteria Showing Significant Effects for Classification of Truthful and Deceptive Accounts .................................................. 61
List of Figures

Figure 1: Number of Details Reported Across Conditions ................. 56

Figure 2: Coherence Ratings (0 - 3) Across Conditions .................. 57

Figure 3: Number of Admissions to Lack of Memory Across Conditions .................................. 59
Acknowledgment

The author would like to express appreciation to the interviewers and coders who contributed to this study. In particular, I would like to thank Deborah Tong, Barry Floyd, Karen Hausmann, Mark Suter, Monica Tymofievich, and Susan Bault who devoted much time and effort to this research in interviewing or coding data. I thank also my advisor, John Yuille, for his support and excellent advice.

The author of this thesis is supported by a postgraduate fellowship from the Natural Sciences and Engineering Research Council of Canada (NSERC). The present research was funded directly by a grant from Division 41 of the American Psychological Association (APA).
Chapter 1

INTRODUCTION

And the Lord said unto Cain, "Where is Abel thy brother?" And he said "I know not: Am I my brother's keeper?" - Genesis 5:9

It did not take long for humans to employ language to attempt to circumvent penalty for their misdeeds. In this first documented case of such a phenomenon, Cain's verbal deception is quickly ascertained by the Deity who imposes a harsh punishment for his murderous transgression. Unfortunately, human investigators do not possess omniscience and must seek alternative methods of credibility assessment in criminal cases. The present thesis explores one method whose complexities and potential utility to this end have only recently been recognized—statement analysis.

Indeed, the question of whether truthful and deceptive accounts can be reliably differentiated based on certain language characteristics has been investigated systematically only in recent years (Landry & Brigham, 1992; Porter & Yuille, in press). Nonetheless, practical applications of statement analysis have been conducted by investigators since antiquity. Larsen (1926) contended that the earliest account of deception detection by a human occurs in the Bible when King Solomon is asked to decide which of two women claiming the same child is telling the
truth. He settled the dilemma by ordering that the child be severed into two pieces after which one woman renounced her claim and the other maintained silence. By examining these language behaviors, Solomon expeditiously determined the identity of the genuine mother (the retractor) and resolved the dispute. Hocking, Bauchner, Kaminski, & Miller (1979) described a papyrus scroll from 900 B.C. which articulates an Egyptian method of credibility assessment based in part on an analysis of the suspect's language behaviors:

The liar does not answer questions, or they are evasive answers; he speaks nonsense, rubs the big toe along the ground and shivers; his face is discoloured; he rubs the roots of his hair with his fingers.

More recently, Camps and Berger (1966) described investigators' utilization of verbal clues in detecting deception in the 1938 London murder case involving John Christie. Police located eight women in various stages of dismemberment and decomposition in the vicinity of Christie's residence. Most had been sexually assaulted at the approximate time of death. Upon questioning, Christie eventually admitted responsibility for the deaths of all the women. An analysis of his statements, however, exhibited a selective presentation of information and other techniques indicative of deception. The delusory elements of his statements became evident with systematic evaluation (mainly by noting discrepancies between his testimony and physical evidence). In retrospect, by comparing his statements to detailed scientific evidence available to the police, it was
possible to assess their accuracy and completeness, revealing Christie's deceptions. However, it is conceivable that a more sophisticated system of statement analysis could have prevented several deaths; Christie had been questioned earlier in the investigation and released, after which other women were killed and the wrong man was hanged for the crimes (Camps & Berger, 1966).

Documented cases of deception detection based on an examination of the accused's account are rare in relation to techniques founded on physiological attributes. In China, an ancient tradition was to require a suspect to chew rice and spit it out; if the rice remained dry, the suspect was deemed culpable in that a guilty conscience was thought to inhibit salivary secretions (Larsen, 1926). In medieval England, the detection of guilt was conducted by methods collectively known as "trial by ordeal." For example, if the accused were tossed into a river and failed to resurface, innocence was proclaimed. Alternatively, if the suspect lived he/she was deemed guilty and consequently drowned or burned at the stake. Other trials by ordeal included immersing the suspect's arm in boiling water for lengthy periods or prolonged bleeding to elicit a confession. These methods were founded on the premise that an innocent person was able to endure the ordeal longer than a guilty criminal (Lykken, 1981).

In modern times, increasingly sophisticated methods of credibility assessment based on physiology have emerged,
most notably the polygraph. Despite its renown, however, the polygraph has not gained unequivocal status as an effective detector of deception. For example, a major criticism of the polygraph is the elevated rate of false positive decisions it yields (e.g., Barland & Raskin, 1976; Horvath, 1977; Kleinmutz & Szucko, 1984), running counter to a legal system designed to protect the rights of the innocent (see Iacono & Patrick, 1987). As well, guilty people can "fool" the polygraph under certain circumstances (Honts, Raskin, & Kircher, 1994).

Other components of the detection arsenal include narcoanalysis, hypnosis, and voice analysis (Psychological Stress Evaluator). As in the case of the polygraph, however, serious shortcomings are associated with each technique that limit their efficacy as investigative tools (e.g., see Gunn & Gudjonsson, 1988). Some will inevitably contend that the experience or intuition of an investigator is superior to any technique designed for credibility assessment purposes. However, recent research has negated this assumption (e.g., DePaulo & Pfeifer, 1986; Kohnken, 1987), indicating that law enforcement officers generally perform at a similar level to untrained people. Ekman and O'Sullivan (1991) found that robbery investigators, psychiatrists, judges, professional polygraphers, college students, and a special interest group who had received training in detecting deceit performed at chance levels in determining the veracity of videotaped reports whereas
Secret Service agents did so slightly (but significantly) above chance level. Indeed, the effectiveness of intuition in any assessment context is questionable (e.g., see critiques of dangerousness assessments based on clinicians' intuition and experience by Cocozza & Steadman, 1978; Klein, 1976; Werner, Rose, & Yesavage, 1983).

The problems associated with traditional credibility assessment methods indicate an exigency for the development of novel, innovative detection techniques. This is especially salient if one considers the fact that often the sole evidence available to the police are the testimonials of the accuser and the accused. For example, in cases of adult and child sexual assault there is frequently little or no evidence to support or refute the allegation (e.g., Yuille, 1988). Cases of this type have been increasing at an alarming rate (see Finkelhor, 1986) causing tremendous difficulties for investigators and adjudicators in establishing fact.

Corresponding to the need for new detection techniques, a multitude of studies on the nonverbal correlates of deception have appeared in the literature (see Ekman, 1992; DePaulo, 1992) but comparatively very little research on verbal correlates. Yet, several researchers have argued that verbal behaviors are more reliable indices of deception than nonverbal behaviors (e.g., Bauchner, Kaplan, & Miller, 1980; deTurck & Miller, 1985; Hocking, Bauchner, Kaminski, & Miller, 1979; Zuckerman, DePaulo, & Rosenthal, 1981). In
the Zuckerman et al. (1981) study it was determined that the accuracy of detecting deception from verbal content (transcript) was higher than from any other channel (facial expressions, body language, etc.) spurring the authors to conclude that this evidence "is in direct contradiction to the belief that nonverbal channels are more likely to disclose deception than are verbal cues" (Zuckerman & Driver, 1985, p. 130). Also, as noted by Gudjonsson (1992), many interrogators appear to have "blind faith" in nonverbal clues to deception, even though they are often unreliable indicators of deception (Ekman, 1992). McCornack (1992) lamented that a "primitive view" of deceptive interaction exists partly because of the myopic focus on nonverbal cues associated with deception and argued that researchers must examine the content of the messages to derive meaningful principles of deception. In the past few years, a largely unresearched but promising series of methods for determining veracity based on an analysis of verbal clues in the accounts of a suspect or witness have emerged. Clearly, the genesis of a technology capable of establishing the truthhood of a person's statements would be welcomed by investigators of crime. Indeed, certain law enforcement agencies have expressed great interest in the development of such a procedure (Lesce, 1990).

In general, the statements of suspects elicited during the interrogation process are quite amenable to analysis. As described by Buckwalter (1983), a usual tactic of
interrogators is to solicit the suspect's free recall narrative account of the criminal act being investigated or his/her alibi. While the free recall account is being provided, few interruptions occur regardless of apparent fabrications in the story. This statement is audio- or video-taped and subsequently transcribed verbatim in written form. This unfragmented narrative account is ideal for structural and content analyses for the purpose of deception detection. It seems likely that in most interrogations a non-admitting guilty suspect will provide either a narrative relating a false alibi or one relating a distorted version of the circumstances or facts of the crime which is conducive to systematic analysis. On the other hand, statement analysis may have less applicability with denial statements (e.g., "I didn't kill anyone!"; "I know nothing about this crime"), statements of memory loss (e.g., "I was so drunk I just can't remember what happened"), or obviously cases in which the suspect refuses to make a statement (e.g., Moston & Stephenson, 1992, found that 16% of the criminal suspects in their sample utilized this option).

In summary, suspect interrogations represent a context in which lying indubitably occurs with high frequency. For certain types of crimes, the words of the complainant and accused are the only evidence available to police (e.g., a delayed report of sexual assault). Unfortunately, the validity of current methods of credibility assessment such as the polygraph is questionable, pointing to the need for
novel approaches. The present thesis investigated the potential utility of verbal clues in the detection of deception from the transcripts of interrogations with honest and deceptive suspects.

Although the idea for this thesis was galvanized and developed from applied issues relating to suspect interrogations, the present work is based upon several theoretical approaches. Up to this point, there has been little integration of the theoretical perspectives relevant to the development of statement analysis techniques. Three major theoretical approaches will be emphasized in the ensuing three chapters: (1) Memory-based approaches describing how accurate and inaccurate memory reports can be distinguished (2) Deception-based approaches describing how deceptive and truthful statements can be distinguished (3) the theories of Avinoam Sapir which contain aspects of the first two approaches but are sufficiently unique and broad in scope to warrant independent discussion.
Chapter 2

PERSPECTIVES ON MEMORY ACCURACY

2.1 The Undeutsch Hypothesis. One of the founding figures in the development of statement analysis was Undeutsch, a German psychologist. Undeutsch was instrumental in the genesis of a systematic procedure for determining the veracity of child witness reports referred to as Statement Reality Analysis (SRA) which later became known as Statement Validity Analysis (SVA). The theoretical basis for this technique, a proposition known as the "Undeutsch Hypothesis", simply asserts that accounts based on memory for an actually experienced event differ from fictitious accounts in structure, content, and quality. This argument has been reiterated by psycholinguists who have argued that when people shift from being truthful to deceptive their language behaviors are markedly altered (e.g., Dulaney, 1982). For example, it is argued that recall accounts for a true event contain a richness of detail not present in inaccurate or falsified accounts (Undeutsch, 1967). According to Arntzen (1983), a multitude of details in a statement is a good indication of credibility because it is difficult for witnesses to intentionally (or unintentionally) embellish a false testimony with many details because the details do not exist in memory. As well, witnesses may not wish to provide a multitude of details because of the increased difficulty of maintaining...
the story's consistency. As noted by Steller and Kohnken (1989), a memory report is considered detailed if it contains the exact description of the place, vivid person descriptions from different aspects, and a step-by-step description of the sequence of events.

The majority of research addressing the validity of the Undeutsch Hypothesis in credibility assessment has been conducted with child witnesses. Statement Validity Analysis (SVA) for children's statements has been formalized and empirically evaluated. SVA has proven useful in assessing the credibility of child witnesses and victims. Although SVA has been developed formally in the past decade, it has been used for about forty years predominantly in Germany and Scandinavia. During the past few years, an international group of researchers have developed and elaborated the procedure (Undeutsch, Steller, Kohnken, Raskin, Esplin, & Yuille; see Raskin & Yuille, 1989; Yuille, 1988). The SVA procedure consists of two stages: (1) criteria-based content analysis (2) examination of other evidence in the case. Content Analysis involves examining the statement for evidence of features associated with actual experiences. Nineteen criteria are used to evaluate the credibility of the content of the statement (e.g., quantity of details, logical structure, accounts of subjective mental state).

Scientific evaluations have been conducted to measure the efficacy of the content analysis component of SVA in determining the veracity of children's statements. Yuille
(1988) had 49 Grades 1 and 3 students prepare one false and one truthful story. The children were given two days to prepare their narratives. SA correctly classified 74.4% of the false stories and 90.9% of the true accounts. Steller, Wellershaus and Wolf (1988; cited in Steller & Boychuk, 1992) conducted a study aimed at increasing ecological validity. Eighty-eight children from Grades 1 and 3 were provided a list of hypothetical experiences containing aspects of direct involvement, loss of control, and negative emotional tone (e.g., having dental work done or being attacked by an animal). They were asked to come up with a true and false story about one of these events. The true and false accounts were successfully distinguished by nine of the SVA criteria.

Unfortunately, there has been a dearth of studies addressing the effectiveness of the SVA procedure in differentiating genuine and fictitious recall accounts in the case of adults. It is, of course, possible that SVA is context-dependent and its proper niche is solely with children's accounts in cases of alleged abuse. Nonetheless, the two studies evaluating its utility in this regard have reported moderately high rates of accuracy in classifying adult statements (Landry & Brigham, 1992; Zaparniuk, Yuille, & Taylor, 1993). Landry and Brigham (1992) recently presented one of the first empirical studies extending the use of SVA to the accounts of adults. They studied adults using fourteen of the nineteen CBCA criteria. Participants
judged the truthfulness of the videotaped statements of twelve adults; six were true and six related an invented traumatic personal experience. For the trained subjects, ten of the fourteen criteria yielded significant differences in the predicted direction between evaluations of truthful and invented statements (reproduction of conversation, spontaneous corrections, and self-doubt were particularly effective) offering compelling evidence for SVA's potential as a tool for adult credibility assessment. On the other hand, two of the criteria were significantly more prevalent in the deceptive accounts (logical structure and comments on another's mental state). Subjects trained in CBCA outperformed the untrained subjects in detecting deception suggesting that training in adult SVA may be useful in enhancing one's ability to distinguish true and false accounts. In the Zaparniuk et al. (1993) study, participants were allocated to one of two groups; in one condition participants viewed one of two videotaped events and then recalled the event while in the second condition participants heard an audiotape of one of the events and then recalled the event as if they had witnessed it. SVA distinguished between the truthful and deceptive statements with a "reasonably" high degree of accuracy (mean hit rate = 76%).

Despite its apparent utility with witnesses, both children and adults, it is difficult to evaluate the efficacy of SVA with criminal suspects based on this
research. The deceivers and truthtellers in the foregoing studies were witnesses or victims, had not actually perpetrated an offense, and were provided little opportunity to prepare their stories prior to the interview. Additionally, the deceivers and truthtellers had little impetus for convincing the observer or interviewer of their veracity. Hence, leakage (signs of deception "leaking" out) in their language behaviors due to arousal may have been minimized. Thus, verbal characteristics of the accounts associated specifically with the experience of committing an offense, planning the lie, and being motivated to escape detection likely did not emerge. It is clear that the potential utility of SVA with criminal suspects has not been tested.

Inconsistent evidence regarding the validity of the Undeutsch Hypothesis derives from empirical studies comparing the amount of detail in true and fictitious memory accounts. The issue of whether truthful and deceptive accounts generally differ in length and degree of detail is unresolved. Some authors (e.g., deTurck & Miller, 1985; Knapp, Hart, & Dennis, 1974; Kraut, 1976; Mehrabian, 1971) have contended that people providing false statements are more laconic than people delivering truthful reports (i.e., their accounts are not as detailed) which appears compatible with the Undeutsch Hypothesis. Macdonald and Michaud (1987) argued that "the suspect who gives only brief answers is almost certainly lying through concealment of information"
Correspondingly, Knapp et al. (1974) conducted an extensive examination of the language of deceit and found that deceivers used significantly fewer words, fewer different words, fewer past tense verbs (was, were, etc.), more other references (they, them, etc.), fewer group references (we, our, etc.), and fewer self-references (I, me, my, etc.). In contrast, Harrison, Halek, Raney, and Fritz (1978) reported that deceivers in their study were significantly more garrulous in their descriptions of an event than people providing truthful accounts, particularly when recounted in the presence of an interviewer. Similarly, Arcaro (1990) and Macdonald and Michaud (1987) asserted that an excessively detailed alibi is indicative of deception because an innocent person would not recall insignificant activities coinciding with the time of the offense. In this view, excessive detail is a clue to a rehearsed response. Zuckerman & Driver (1985) conducted a meta-analysis examining twelve visual behaviors (e.g., pupil dilation, smiling, postural shifts), six linguistic behaviors (e.g., response length, latency, speech errors), and five verbal behaviors (e.g., self-references, irrelevant information, negative statements) that had been cited at least twice in the literature as being indicative of deception. Response length was significantly negatively correlated with deception. They found that the behaviors most strongly associated with deception (p.<.001) included pupil dilation and adaptors (visual), speech hesitations
(paralinguistic), and immediacy (-) and negative statements (verbal). Other significant relationships between language behaviors and deception were speech errors, pitch, irrelevant information, and leveling (frequency of overgeneralized statements indicated by terms such as "every", "all", "nobody", "none", etc.).

One of the premier studies to examine clues to deception from multiple behavioral indices addressed this issue. deTurck and Miller (1985) compared the videotaped accounts of deceptive undergraduates who had been induced by a confederate to cheat on a task to win a prize for performance to those who had been engaged in a task-related discussion with the confederate. Dependent measures included five nonverbal (number of blinks, adaptors, feet/leg gestures, eye contacts) and four verbal (number of speech errors, number of pauses, length of response latency, and message duration) cues to deception. The results indicated that six cues (adaptors, hand gestures, pauses, speech errors, response latency, and message duration[-]) were significantly associated with deceptive accounts. The authors argued that these findings support the assertion that verbal cues are more important than nonverbal cues in detecting deception. Stiff and Miller (1986), employing a similar methodology to the deTurck et al. (1985) study, found that deceptive statements were predicted by "verbal content" \( r = -.40 \), measured by how plausible, definite, clear, and concise the response was rated). statements of
personal responsibility (r=.21), self-references (r=.21), and number of words (r=-.21). In summary, the majority of studies addressing the account length/veracity relationship have found a positive correlation lending support to Undeutsch's original claim.

One of the first published formal applications of statement analysis to an actual crime considered the Undeutsch Hypothesis (i.e., the amount and coherence of details) in evaluating credibility (Yuille & Cutshall, 1989). These authors examined the 1972 case of Jeffrey MacDonald whose wife and two children had been brutally murdered. MacDonald was subsequently found guilty of the crimes and imprisoned. MacDonald maintained that a group of drug-crazed hippies had entered his home and committed the offenses. A woman named Helena Stoeckley corroborated his account, claiming for years that she was one of the intruders of whom MacDonald spoke. Yuille and Cutshall examined Stoeckley's statements and concluded that "the logical consistency of her statements, particularly the relationship of her account to the general pattern of the murders, the spontaneous nature of her description, and the wealth of details in her statements provide evidence which supports the credibility of Helena Stoeckley's account" (p. 188).
2.2 Reality Monitoring. A seminal paper by Johnson and Raye (1981) described elaborately how memories for past perceptions of actual events can be distinguished from past acts of imagination, using a model termed "Reality Monitoring." This model posits that perceptual memories are associated with a higher level of external sensorial information (e.g., "It was a large cedar home") whereas self-generated (imagined) memories are the result of creative cognitive processes and are thus associated with more internally-generated cognitive details and subjective idiosyncratic information ("I remember seeing the home and thinking how lovely cedar homes are").

Schooler, Gerhard, & Loftus (1986) have demonstrated that memory descriptions of genuine and imagined (suggested) events differ in ways congruous with the theory of reality monitoring. Some participants were induced through suggestion to recall an object that had not actually been witnessed (employing Loftus' postevent misinformation paradigm). Content analysis revealed that genuine memories (in written form) contained more references to attributes of the memory stimulus, less reference to cognitive processes occurring during encoding, fewer "hedges" (e.g., "It seems to me" or "I believe"), and fewer words. In a second study, Schooler, Clark, and Loftus (1988) employed more critical memory items and asked participants for verbal rather than written accounts. Again, descriptions of genuine memories contained more sensory details, fewer references to
cognitive processes, fewer verbal hedges, and fewer self-references (i.e., pronoun "I").

Nonetheless, as pointed out by Leippe, Manion, & Romanczyk (1992), the memory stimuli employed by Schooler et al. (1986; 1988) were noncomplex (e.g., a stop sign). Leippe et al. (1992) suggested that even more frequent and richer distinguishing cues may have emerged in recall accounts concerning a more complex event. Addressing this possibility, Leippe et al. (1992) staged an interactive, elaborate event in which participants were given a bogus skin sensitivity test including touching and conversation. Participants were subsequently interviewed to gauge their memories for the event. Leippe et al. (1992) then divided the accounts into high and low accuracy and compared them on a number of measures. It was found that the accurate accounts were delivered more confidently, contained less instances of verbal hedging, less admissions of memory failures (e.g., "I don't know"), and were lengthier than the inaccurate accounts. As well, people judging the accuracy of the accounts generally overused confidence and underused the other cues in determining the accuracy of the accounts. The authors concluded from these results that "with further research and insight, it may be possible to train people to look for them (cues) and so become better judges of memory reports" (p.195).
3.1 Emotional-Motivational Approaches.

Inability to lie is far from being love of truth.
Be on your guard. — Nietzsche’s Zarathustra

Presumably, lying is a difficult undertaking for most people especially when the consequences for a failed lie are of considerable magnitude. As noted by Simpson (1992), a liar acts self-consciously against a convention containing a moral dimension. Accordingly, emotional and motivational factors have been implicated as being critical in differentiating truth and falsehood in the interrogation context (see Ben-Shakhar & Furedy, 1992 for a comprehensive review of the various theoretical perspectives considering these factors in physiological deception detection).

According to the theory of the "punishment mechanism", fear of failure to deceive leads to a specific physiological response associated with the fear (Ben-Shakhar & Furedy, 1992). Davis (1961) argued that the greater the consequences of being detected, the greater the physiological response during lying and, therefore, the greater the likelihood of detection. In an interrogation situation, a suspect with guilty knowledge will be expected to exhibit high arousal due to this fear of detection, leading to discernible changes in verbal behaviors. Nonetheless, as pointed out by Inbau, Reid, and Buckley (1986), signs of nervousness during an interrogation in both
guilty and innocent suspects can result from many factors independent of their level of veracity (e.g., concern that police and family will discover other transgressions; concern for what will happen to them while in custody and during the interrogation). However, these authors argue that guilty suspects will experience higher levels of anxiety than innocent suspects because of their knowledge of the crime.

Other hypotheses relating deceptive behavior to emotional arousal include the "conditioned response theory" which postulates that lying produces autonomic reactions because the related questions are associated with negative experiences (i.e., because past lying behaviors were frequently linked with aversive stimuli) and the "conflict theory" (see Kohnken, 1985) positing that observable psychological reactions result from the conflicting needs to lie (situational demand) and tell the truth (societal or familial demand).

It is evident that the motivation for being deceptive is an important consideration in deception detection by statement analysis. Ekman and Friesen (1969) asserted that it is most difficult to lie successfully when the consequences of deception detection are great. More recently, DePaulo and colleagues (e.g., DePaulo & Kirkendol, 1989) have referred to this phenomenon as the "motivational impairment effect." Physiological research has demonstrated an association between increased motivation and higher rates
of detection (e.g., Gustafson & Orne, 1963; Elaad & Ben-Shakhar, 1989). If this theory holds for observable behaviors (including language behaviors), liars and truth-telling persons would be expected to behave differently to a greater degree when there are serious perceived consequences from detection. Obviously, the impetus is high for a guilty suspect to successfully deceive his/her inquisitor. When the ramifications are this significant, deception clues, or "leakage", will likely be emitted through speech, voice, body, and face (Ekman, 1992) despite the liar's attempts to conceal emotional portrayals. The present study is founded in part on the theory that emotional and motivational factors influence people to "make verbal mistakes that can provide both leakage and deception clues" (Ekman, 1992, p.85).

One aspect of Zuckerman et al.'s (1985) meta-analytic study concerned motivation for lying. Motivation was considered substantial if participants had been promised some monetary reward for doing well on the deception task or if the task had been described as a measurement of a skill. Highly motivated deception was associated with response length (−), speech rate (−), speech errors, speech hesitations, pitch, negative statements, self-references (−), and leveling more so than unmotivated deception. The criteria related to language behaviors proved quite useful in differentiating truthful and deceptive statements,
particularly when motivation was high. The authors concluded the following:

It appears that although the content of the speech may be quite controllable, its semantic structure may be as leaky as the body and/or tone of voice. Consequently, the search for cues to deception must focus on both verbal and nonverbal behaviors (Zuckerman et al., 1985, p. 144).

Experimental studies have found that the experience of lying may result in an elevated frequency of speech errors (e.g., Kohnken, 1985), possibly related to the motivational impairment effect. Comber and Canter (1983) conducted a study to determine whether malicious (hoax) and non-malicious (truthful) fire alarm calls could be distinguished on the basis of their psycholinguistic characteristics. Eight malicious and eight non-malicious calls were compared in terms of the number of speech errors of various kinds occurring in the call (number of times the operator and caller spoke simultaneously, number of filled pauses ["umm", "ahh"] in the caller's speech, number of questions asked by the caller, and the number of false starts in the operator's and caller's speech). Although no single variable significantly differentiated malicious and non-malicious calls (probably due to low statistical power resulting from the sample size), the mean differences were in the predicted direction for all variables leading the authors to conclude that "it was possible to differentiate the two types of calls based on the overall configuration of speech behavior within the call" (p. 460).
3.2 Cognitive Approaches.

For words, like Nature, half reveal
And half conceal the soul within. -
Alfred Lord Tennyson

The interrogation may be viewed as a context in which deceptive information is employed as an adaptive response to a dangerous situation. As noted by McCornack (1992), researchers have recently argued that deceptiveness reflects a functional adaptation to the demands of a complex communication situation. People are often confronted with interactions in which the competing goals of conveying information and minimizing the damage which might result from the information are operating. Thus, people present deceptive messages in ways fostering control and manipulation of dangerous information. Building on this premise, McCornack (1992) formulated "Information Manipulation Theory" (IMT) postulating that deceptive messages function deceptively because they covertly violate principles guiding conversational exchanges. IMT suggests that although deceptive messages are infinite in number, the information in them varies in systematic, identifiable ways. Given that conversational interactants possess expectations regarding the quantity, quality, manner, and relevance of the information presented, it is likely that deceivers violate these expectations in attempts to deceive the listener. In the suspect interrogation, IMT would predict that the information presented by a guilty suspect would differ from that of an innocent suspect on these four
measures (quantity, quality, manner, and relevance) because of the "danger" which sensed by the deceiver. Deceptive interactants are likely to fail to fulfill the informational requirements of the interaction, provide information that is irrelevant to the central conversational topic, and provide information in a less orderly, coherent fashion with fewer attempts to avoid ambiguity and confusion.

Others have also contended that information processing and attentional strategies are important considerations in a deception detection task. Waid & Orne (1981) have suggested that physiological detection may depend on how the suspect processes the information presented to him/her in the interview. Suspects may avoid detection by ignoring the stimulus or responding mechanically to the stimulus without "deeply" processing it (Craik & Lockhart, 1972). It may be possible for suspects to circumvent the problem of deception leakage by encoding information in the interview on a superficial level. One interviewing strategy which may be effective in counteracting the processing of interview stimuli on a superficial level is the use of interrogative probes (e.g., "Come on. It is obvious that you are lying to me. What really happened?"). Such negative feedback may make it very difficult for a suspect to encode and respond to information superficially (see Gudjonsson & Clark, 1986). It is possible that verbal cues to deceit may be differentially useful for credibility assessment purposes in
the context of a confrontive, negative interrogation situation (cf. Stiff & Miller, 1986).

The superficial-encoding strategy may be made more effective by planning its implementation prior to the interview. As noted by DePaulo, Lanier, and Davis (1983), lies in the laboratory are usually characterized by low levels of cognitive planning (including the story to be presented and strategies for avoiding detection), yielding findings that may not be generalizable to forensic contexts. Most experimental research on deceptive communications has examined the nature of the spontaneous, unprepared lie. However, a guilty person usually has the opportunity to prepare for an interrogation or courtroom testimony by inventing and rehearsing his/her story (Carpenter, 1990). He/she has anticipated lying and, as the critical questions are asked, provides lies that are rehearsed and that he/she is mentally prepared to tell. Additionally, as noted by Buckwalter (1983), fabrication and deception are predominant in the daily lives of many criminals for survival on the street and to avoid incarceration so they may be quite practiced at and proficient at fabricating information. In devising generalizable experimental research addressing statement analysis it is apparent that this preparation factor must be considered, although few studies have done so.

O'Hair, Cody, and McLaughlin (1981) investigated deception of factual information for both prepared and
spontaneous lies. Truthful and deceptive accounts were compared on paralinguistic measures and body leakage clues. Subjects who had anticipated lying showed shorter response latencies and less postural shifting than truth tellers. Secondly, the prepared liars showed higher head-nodding rates, lower laugh/smile rates, and shorter message durations. This is consistent with Zuckerman et al. (1981) who argued that efforts to conceal deceptions by controlled restraint of behavior can be counterproductive because the controlled behavior now appears overcontrolled and unnatural. Additionally, not all behavioral areas can be controlled equally resulting in leakage from different communication channels.

Alonso-Quecuty (1990) showed subjects a film depicting a physical assault on a woman by a man in the context of an abortion dispute. All subjects provided both an accurate account of the event (in which the aggressor was convicted) and an account in which information was falsified (the man was acquitted). Half of the subjects provided their accounts immediately following the film while the others were given ten minutes to prepare their accounts. The falsified accounts contained less words in the immediate statements but more words in the delayed statements leading the author to conclude that "the length of the statements may be a good detector of deception when the individual has disposed of the necessary time to elaborate them." Since this is generally the situation in criminal cases, the
length of the statement may be a useful consideration in forensic contexts. As well, Alonso-Quecuty's (1990) finding may help resolve the inconsistencies described earlier regarding the research addressing the account length/veracity relationship. Notably, in the Alonso-Quecuty (1990) study the delayed accounts were approximately twice as long as the immediate accounts for both the truthful and deceptive statements, again illustrating the effect of preparation on testimony.

Kohnken (1985) has investigated the utility of the information processing approach in detecting the deception of eyewitnesses. He contended that providing a deceptive eyewitness testimonial represents a more difficult cognitive task than giving an honest account and that with cognitive load symptom substitutions should appear. Based on previous studies he concluded that cognitive load can result in an increase in speech hesitations (cf. Macdonald & Michaud, 1987, p. 36), a reduction of the informational value of applied words, and an increase in repetitions. To test his information processing model, Kohnken devised an experiment in which subjects viewed a film depicting a theft. Prior to interviews, participants were divided into a truthful condition (instructed to tell the truth about the film), a deception condition (instructed to implicate the wrong person as the thief and offered an LP record if the lie was successful), or a hearsay condition (subjects did not view the film but heard a verbal report about its contents and
were asked to convince the interviewer that they had viewed it; if successful they would receive an LP record). Thirteen dependent variables were examined including speech rate, filled pauses, self-reflections, self-corrections, type token ratio (TTR; see below), and mean word length. Results indicated that the groups could be reliably distinguished by most criteria and a multiple discriminant equation was derived to efficiently differentiate the groups (the TTR exhibited the highest discriminating power). This prompted Kohnken (1985) to conclude that his hypothesis had been accurate; telling lies is a more complex task than telling the truth and this increased task difficulty is reflected in speech behavior. Thus, an information processing approach appears to be appropriate in investigating speech and deception in witnesses. The present work extends this approach to the deceptive language of (mock) criminal suspects which has not yet been examined.

Another cognitive approach to statement analysis considers the complexity of the verbal responses in the interrogation context. Osgood (1960) argued that "behavior under increased drive, including encoding behavior, should become more stereotyped - the alternatives selected at all choice points should tend to be the most familiar, the most practiced, and the most expected." According to this view, a person providing testimony under greater motivation to appear credible should display lesser lexical diversity. Deviations in lexical diversity have proven useful in
DECEPTIVE LANGUAGE

identifying deceptions in the accounts of suspects. Lexical diversity is usually measured by calculating the number of distinct words ("types") over the total number of words ("tokens") in a statement, referred to as the "type-token ratio" (TTR) in Forensic Psychology. For example, in Lincoln's "government of the people, by the people, for the people" the six types and ten tokens yield a TTR of .60 whereas the statement "the boy left the store after his purchase of milk" yields a .90 TTR (nine types, ten tokens). In other words, a high TTR suggests that a person communicating employs a broad vocabulary, whereas a relatively low one reflects a communicator's preference for stereotypic language. The TTR has long been employed by linguists as a descriptor of language behavior but has only recently been applied to establishing veracity for law enforcement purposes. Hollien (1990) posited that the higher one's motivation, the lower one's TTR. Certain psycholinguists have argued that deceivers show higher TTRs than those of people uttering truthful statements.

Carpenter (1990) argued that the TTR is an excellent index of an individual's internal psychological condition. He found that when adults shift from being truthful to lying there are some clearly discernible changes in their TTRs. Carpenter provided several case studies in which the suspect's TTR had increased significantly from his/her mean TTR at points in the account later proven deceptive. As Carpenter notes, knowledge of the crime allows guilty
suspects to know which of their statements are false or self-incriminating and must be phrased with greater caution during interrogation or courtroom testimony. Although not detectable to listeners, the additional time needed to phrase such statements results in a higher frequency of comparatively uncommon words occurring in particular components of a statement. Similarly, when hidden knowledge of events constrains a suspect at some points during testimony, the additional split-seconds needed to formulate those sentences provides time for more different words, relatively uncommon to the individual, to be spoken.

Carpenter noted that in criminal investigations suspects attempt to portray credibility and thus many of their statements are truthful. It may be necessary to systematically examine changes within an individual's discourse to obtain meaningful information regarding veracity. Truthful statements should yield TTRs close to the person's mean TTR whereas significantly higher TTRs are indicative of deception.

Thus far, there is insufficient evidence to conclude that the TTR has utility in detecting deception in criminal cases. Empirical studies have supported the assertion that statements uttered by people experiencing apprehension or caution about possible adverse reactions from listeners have higher TTRs than those with no threat (e.g., Dulaney, 1982). Consequently, it is not known whether a high TTR in an interrogation is necessarily indicative of deception; it
could result from anxiety unrelated to guilt (see Inbau, Reid, & Buckley, 1986). It is possible, as Carpenter has asserted, that examining TTR variations within a suspect's statement could indicate deceptive aspects of the statement. Dulaney (1982) placed participants in a situation designed to elicit truthful and untruthful statements. By employing computer programs, he found that when falsifying information the participants used significantly fewer words, fewer unique words and larger TTRs, smaller perceptual-cognitive activity measures, and fewer past tense verb forms.

Considerations of lexical diversity have proven to have utilitarian value in several contexts. For example, the lexical diversity of different messages have been analyzed to determine whether they had a common source or to specify communication origin in a process referred to as "stylometry" (e.g., see Arens & Meadow, 1957; Gudjonsson, 1992; Miron, 1984; Morton, 1978). Such methods have also been used to successfully determine whether threats were genuine or hoaxes. For example, it was successfully predicted that the Symbionese Army, who kidnapped Patty Hearst, would die in a suicidal confrontation with law enforcement officers (Miron & Pasquale, 1978).
Certain principles of statement analysis have been elucidated and formalized by Avinoam Sapir, an ex-Israeli police polygrapher. Sapir teaches these concepts to law enforcement personnel in a seminar entitled Scientific Content Analysis (SCAN). Although some criminal investigators have adapted Sapir's concepts, there has been no scientific assessment of their worth (Lesce, 1990). It is not known what aspects of SCAN best distinguish true and false statements by suspects or the overall utility of the technique. Following are some of Sapir's (generally unsubstantiated) principles which were tested in the present study:

4.1 Ratio of Component Elements. Sapir argues that the ratio of three main elements of a suspect's statement can be indicative of deception. In general, statements are composed of three elements - an introduction, a central issue, and a conclusion. According to Sapir, people intending to make a truthful statement will generally recount a lengthier conclusion than introduction while people making a deceptive statement show the opposite pattern. Sapir argues that because humans tend to be egocentric they evaluate events primarily in terms of how they are affected by them. Thus, truthful people are not
concerned with being perceived as credible but instead focus solely on the event's personal impact. Witnesses to a fatal car accident may state as much information about how the wreck halted traffic and how waiting for the ambulance made them late for work as the wreck itself. In contrast, deceptive people are primarily concerned with appearing credible. Sapir argues that they devote a great amount of time describing how it was logical and rational for them to have been where they were at the time of the accident. Since they did not personally experience the incident, there is little affect present. Accordingly, they truncate their conclusions. Sapir asserts that this factor singularly identifies deceptive statements with roughly 85% accuracy.

4.2 Unnecessary Connectors. Unnecessary connectors include words such as "started", "continued", "completed", "after", "afterwards", "the next thing I remember", "then", and "and then after that" which are employed by a speaker to selectively pass over incriminating information. Sapir argues that a person who is falsifying information will employ a higher frequency of unnecessary connectors. The most commonly used connector is "then after that" which is used to buy time for "mental processing"; a recent event being honestly recalled should require little such processing. Sapir asserts that connectors often indicate that details have been deliberately edited. For example, "I had dinner and I went to the store" indicates that nothing
happened in the interim whereas "I had dinner and then after that I went to the store" is an intimation that something of significance may have transpired in the hiatus. An increased frequency of the use of connectors should arouse suspicion and lead to further investigation into the particular temporal segment being referred to.

4.3 Use of Pronouns. According to Sapir, the use of pronouns during speech also reveals a considerable amount of subtle information. Sapir argues that any deviation from the first person, past tense, or active voice in describing an event is significant in indicating deception. This perspective coincides with that of Macdonald and Michaud (1987) who asserted that a shift by a suspect from talking in the first person to the third person indicates that the suspect wants to distance himself from the interrogation and the crime. Recollections from genuine memory usually begin with the pronoun "I" whereas someone making a false statement is more likely to employ the pronoun "you". For example, a person making an honest statement might remark "I could see the blood dripping from the back of the truck" whereas a deceptive person would likely say "You could see the blood dripping from the back of the truck."

Nonetheless, as discussed previously, Schooler, Clark, and Loftus (1988) have found that true memory accounts for an event contain the pronoun "I" at a lower frequency than false memory accounts (using the Loftus postevent
misinformation paradigm). They also found that false memory accounts included more "verbal hedges" (e.g., "It seems to me" or "I believe") which was a replication of an earlier finding (see Schooler, Gerhard, & Loftus, 1986), explicated in the framework of Johnson and Ray's (1981) theory of reality monitoring.
Chapter 5
SUMMARY OF THEORETICAL FRAMEWORK AND THE PRESENT STUDY

In the preceding discussion it has been argued that, in seeking principles of verbal deception in the interrogation context, it is necessary to consider theories from a variety of perspectives. Approaches emphasizing memory accuracy suggest that systematic differences in several qualitative and quantitative aspects of memory narratives can reliably indicate their veracity (i.e., the Undeutsch Hypothesis, Reality Monitoring). Perspectives on deception consider differences in truthful and untruthful statements due to emotional, motivational, and cognitive factors unique to the interrogation context. Finally, the assertions of an ex-Israeli polygraph operator and police officer named Avinoam Sapir are discussed. Sapir argues that the ratio of component elements, the presence of unnecessary connectors, and deviations in pronoun usage in interrogation transcripts are indicative of deception. Together, these approaches form an integrated model of statement analysis which can be developed for use in certain criminal investigations.

This thesis represents an empirical evaluation of the measures discussed previously. In this investigation, the accounts of truthful, partially deceptive, and completely deceptive mock suspects were compared on eighteen verbal attributes derived from these approaches. The methodology and dependent measures employed reflect a holistic approach
to statement analysis considering the emotional-motivational and cognitive processes associated with criminal interrogations.

The major circumscribed objectives of this thesis research were: (1) to investigate the utility of the various known verbal correlates of deception for assessing the credibility of criminal suspects (2) to investigate the utility of some previously unresearched statement analysis principles in assessing the credibility of criminal suspects. Minor goals of this study included: (1) to introduce a methodological variation on the traditional mock theft paradigm aimed at maximizing experimental realism and ecological validity (2) testing the hypothesis that interrogation probes will result in differential verbal behavior by truthful and deceptive criminal suspects. The ultimate goal of this and future research is to develop a model of credibility assessment for use with certain criminal suspects based on verbal indicators of deception.
Chapter 6

METHOD

6.1 **Participants.** Sixty participants (N=60) were recruited from the University of British Columbia undergraduate population under the premise of participating in an experiment addressing security effectiveness. The incentive for participating was five dollars for a one hour session with the "possibility of earning more." Although it was originally anticipated that equal numbers of males and females would be used, this proved impossible given that mostly women responded to my advertisement (forty-four women, sixteen men). The first language of all participants was English.

6.2 **Procedure.** Upon arriving at the laboratory, each participant was asked to sign a consent form. Next, they were falsely informed that the Psychology Department had just hired an un-uniformed security guard because of a recent rash of thefts in the building. Further, unbeknownst to the guard, our laboratory group had been contracted the task of testing the effectiveness of this new measure. In three of the four conditions, the participants' task was to proceed to a specified office in the Psychology Department and attempt to locate a one hundred dollar bill hidden in a folder on a bookshelf in the room. They were given ten minutes to obtain and pocket the cash, restore the room to
its former condition, and return to the laboratory. While acquiring the cash they were told to be as careful as possible to escape detection from the guard. Participants in Condition 1 were asked to obtain a folder (the same folder used for the theft conditions but without the money) from the same location for a professor.

Each participant was provided a key to the office and directions to proceed to its location. It was reiterated just prior to their exit that they would have only ten minutes to return to the laboratory and to be very careful.

Upon returning to the laboratory, participants gave either the money (Conditions 2, 3, and 4) or the folder (Condition 1) to the experimenter. Next, they were informed that the second part of our contract with the security company was to test the prowess of their interviewers at interrogating suspects and establishing which are guilty and innocent. Thus, all participants were told that they would be interrogated by someone who was aware that he/she would be questioning both truthful and deceptive suspects and whose agenda was to establish the truth based solely on the interview. Further, this interviewer had been informed that he/she may hear some truthful confessions, false confessions, truthful denials, and false denials.

6.3 Instructional Manipulation. The main experimental manipulation in this study involved the directions provided to participants preceding the interrogation. Participants
were randomly assigned to one of four experimental conditions (n=15 per group: 11 females, 4 males): 1. Truthful alibi condition 2. Partial deception condition 3. False alibi condition 4. Truthful confession condition. All participants were given fifteen minutes to prepare their accounts before entering the interrogation room. The participants from each condition were provided the following instructions:

1. **Truthful Alibi condition**: These participants had been assigned an innocuous task in the crime room (to obtain a brown folder for a professor). Prior to the interview they were informed that they were suspects in the theft of a one hundred dollar bill and would be interrogated as such. They were told to simply relate the complete truth about anything that was asked by the interviewer (i.e., relate their truthful alibi). If they were able to convince the interviewer that they were reporting truthfully they would receive an extra five dollars from the experimenter.

2. **Partial Deception condition**: These participants were instructed to provide the interviewer with an account containing some truth and some falsehood. They were asked to relate a truthful account with the exception that the theft of money was to be replaced by an unsuccessful attempt to retrieve "a folder" belonging to a specified professor. These participants were asked to deny any knowledge of money existing in the room. Further, they were informed that if
the interviewer believed their account they would receive an extra five dollars at the culmination of the interview.

3. **Complete Deception (false alibi) condition:** These participants were asked to provide the interviewer with a completely falsified account in the form of an alibi. They were instructed to create a plausible story accounting for their whereabouts and actions in the ten minute time period during which the money was stolen. Additionally, under no circumstance should they admit their deception. Again, participants were told that if the interviewer believed that they were recounting a truthful account, they would receive an extra five dollars at the end of the interrogation.

4. **Truthful Confession condition:** These participants were asked to provide the interrogator with a completely truthful account admitting their involvement in the theft if requested. Further, they were informed that the interrogators were expecting some false confessions. Again, the monetary incentive for successfully convincing the interviewer was offered.

6.4 **Interviewing.** Each participant was initially questioned in the fashion of the Step-Wise Interview (see Yuille, Hunter, Joffe, & Zaparniuk, 1993). Basically, this semi-standardized interview procedure involved eliciting a free recall account in which the participant was asked to state everything he/she could recall from the time he/she exited the laboratory until the time he/she returned. This
was followed by a series of more directive questions intended to clarify and elaborate the information provided in the free recall account. These questions consisted mainly of open questions (e.g., "You mentioned seeing a desk. Can you tell me more about that?") but occasional specific questions (for a more complete description of the interrogation protocol refer to Appendix B).

**Interrogation Probes**

The Step-Wise Interview was ensued by a series of interrogation probes designed to simulate a police interrogation and test participants' responses to negative feedback. This consisted of three statements expressing disbelief by the interviewer: (1) "Do you have any knowledge of the money that is missing from Room 2408?" (2) "Frankly, I'm a little skeptical about your explanation. Are you being completely honest?" (3) "Come on now, it's really obvious that you are lying to me. Why don't you just admit it?". All interviews were video- and audiotaped.

Five senior or graduated Psychology students acted as interviewers in this study. They were trained in this semi-structured interview via reading materials (see Appendix B) and practice with critical feedback from the experimenter who had previous training and experience with the technique; e.g., Porter, Yuille, & Bent, in press). They initially received two one and a half hour training sessions in which they observed a sample interview, received an audiotape containing the sample interview, and the handout detailing
the interview steps. Working in pairs, the interviewers subsequently taped their practice interviews and provided them to the experimenter who gave exhaustive feedback. This process continued until the quality of the interviews was considered acceptably high.

With the culmination of the interviews/interrogations, all participants were thanked and paid the monetary bonus. They were then completely debriefed regarding the purpose of the research and the reason for the experimenter's deception (see Appendix A). They were also implored not to discuss the nature of this study with acquaintances.

6.5 Transcribing. All interviews were transcribed verbatim from the audiotapes onto computer diskettes. All utterances made by the participants were recorded in the transcripts. In those cases where a participant's speech was inaudible on the audiotape, the videotaped interview was examined. If the utterance in question was still considered ambiguous or inaudible by the transcriber, it was recorded as "inaudible" on the transcript. Finally, hard copies of all interviews were generated.

6.6 Dependent Variables and Scoring Procedures
a. Amount of Detail Reported: The accounts were subjected to an evaluation procedure based on the protocol developed by Yuille and his colleagues (e.g., see Yuille & Cutshall, 1986; Yuille & Porter, 1993). The reports were partitioned
into single units of information. All details were scored as correct, incorrect, or unclassifiable. This was done by comparing the details to the videotape of the event for each participant. Each detail received one point if it contained a specific unit of information. For example, the sentence "I leaned on the beige desk" contains one action detail and two descriptive details. Because previous research with adults has typically found similar patterns with action-oriented and descriptive details (e.g., Yuille & Tollestrup, 1990; Yuille, Tollestrup, Marxsen, & Porter, under review; Yuille, Davies, Gibling, Marxsen, & Porter, in press) possible differences were not tested in this study. Information that is vague or nonspecific was given one half point (e.g., "The room was large"). A detail was considered unclassifiable if it may have occurred but verification was not possible (e.g., "I felt a cool breeze"). The information deemed unclassifiable was not be entered into this analysis. This variable was frequency coded in lieu of the more subjective SVA rating process usually used in forensic practice.

b. Laconic-Garrulous Continuum: The length of the accounts was measured by counting the number of words in each account. For this variable, all filled pauses were deleted and a spelling check generated prior to a computer word count.
c. **Lexical Diversity** (Type-token ratio): An analysis of lexical diversity as measured by the TTR was conducted with the use of a computer program developed for this purpose.

d. **Frequency of Verbal Hedges**: This criterion was measured using Loftus' definition of hedging ("I believe", "It seems to me", "I think", "I figure"). A computer search was conducted for each of these phrases and the frequencies recorded.

e. **Frequency of Filled Pauses**: The number of filled pauses were tallied by recording the number of times the participant used one of the following sounds: "ummm", "uhhh", "wellel", "hmm", etc. The transcriber had been instructed to record filled pauses with the one of these four items which sounded most like the utterance. This allowed a computer search to later be conducted generating frequencies for each of these items.

f. **Frequency of Unnecessary Connectors**: The frequency of unnecessary connectors were tallied. These include the words "after", "afterwards", "next", "the next thing I remember", "then", and "then after that". These phrases had to be employed during a description of a participant's actions related to the event in question (i.e., from the time the participant exited the lab until the time he/she returned). Hence, a description of the initial interaction with the experimenter using a "and then after that" would not constitute an unnecessary connector. A computer search was generated for the six above phrases and an examination
of the context of the utterance was made by the coder. If the connector was employed in describing actions pertinent to the central event it was tallied.

g. **Frequency of Pronoun Usage Deviation**: All deviations from the first person (e.g., "You could see a bunch of books on the desk" instead of "I could see a bunch of books on the desk") and past tense (e.g., "I'm going into the office") in describing the event were recorded.

h. **Number of Self-References**: All self-references were tallied (i.e., I, me, my, mine, myself) through a computer search.

i. **Ratio of Component Elements**: A coding system was devised to trichotomize the accounts into three elements — an introduction, a central issue, and a conclusion. Coding was only conducted on the initial, uninterrupted free recall account. The number of words employed by each participant to describe events up to the point of entering the "crime room" were counted and considered the "introduction" to the central issue. The number of words used to describe events subsequent to exiting this room (up to the culmination of the narrative) were counted and labelled the "conclusion". The ratio of the number of words in the conclusion over the number of words in the introduction was calculated for each participant. Data from participants in Condition 3 (false alibi) were not coded for the present analysis because the events described generally did not include the crime room and were thus were not comparable to the other conditions.
Two raters with a recent extensive three-day training in SVA scoring by a recognized expert in the technique judged the presence of each of the following SVA criteria and assigned a rating (0 – 3 [strongly absent – strongly present]) or tallied the number of pieces of evidence for each criterion contained in the accounts. Where possible the more objective method of counting specific pieces of evidence for each criterion was employed (note that the following SVA criteria were not applicable to the event used in the present study: reproduction of conversation, attributions of the perpetrator's mental state, pardoning the perpetrator, details characteristic of the offense, contextual embedding, descriptions of interactions, accurately reported details misunderstood, self-deprecation; for a more detailed explication see Steller, 1989; Steller & Kohnken, 1989; Yuille, 1990).

j. Coherence: Basically, this criterion is present when the various parts of the account fit together in a coherent and consistent fashion. A rating of 0 – 3 was assigned to each account.

k. Spontaneous Reproduction: The actual recounting of the event should occur in a spontaneous, unstructured fashion.
Rigid structure would result in a low rating on this criterion (0 - 3).

1. **Unexpected Complications During the Incident:** This refers to the number of unexpected occurrences complicating the main task of the participant as reported in the recall account. This did not include being unable to locate the folder (participants in Condition 2 had been instructed to give this falsehood).

m. **Presence of Unusual Details:** This criterion refers to the spontaneous reporting of details concerning the event that are realistic but unusual (e.g., reporting seeing a book by one of his/her favorite authors) to the coder. The number of such details was tallied for each account.

n. **Presence of Peripheral Details:** This refers to a report containing details that are not directly related to the target event (e.g., temperature of the room). A rating of the 0-3 was assigned to each account from Conditions 1, 2, and 4. Condition 3 was omitted from this analysis because preliminary trials indicated that it would be very difficult to define details "peripheral" to the central event. For example, a few participants described simply wandering through the building looking around as their alibi.

o. **Related External Associations:** This criterion refers to spontaneous references to an incident related to the event but did not occur within the event (e.g., a participant might report that the office reminded him/her of Dr.
Yuille's office). The number of such details was tallied for each account.

p. Accounts of Subjective Mental State: This criterion refers to the number of times a participant refers to his/her thoughts, intentions, or emotions during the event.

q. Spontaneous Corrections: This refers to the number of times the participant corrects him/herself in the course of recounting the event.

r. Admitting Lack of Memory: This criterion refers to the number of times the participant admits that he/she cannot recall an aspect of the event (generally in response to directive questions from the interviewer). This includes responses to inquiries such as "Is there anything else you can remember?". Multiple admissions regarding the same aspect of the event were tallied as a single admission.
Chapter 7  
RESULTS  
All reported statistical analyses were conducted with SYSTAT 5.01 software.

7.1 Reliability Checks. Inter-coder reliability checks were conducted for all dependent variables. For variables coded by computer tallying, test-retest reliability was measured by conducting a second computer search (to control for any input errors in the initial search). For all other variables, eight interviews (two from each condition) were randomly selected for dual coding (see Kohnken, 1985). For the SVA variables two coders recently trained in SVA scored each of these interviews. Pearson's product moment correlations were calculated for each variable followed by between-subjects t-tests to ensure that the average frequencies for the two coders did not differ (see Bordens & Abbott, 1991, p. 155). None of these t-tests showed significant mean differences between the two coders. The Pearson r reliability scores revealed that, with the exception of one criterion, inter-coder reliability was acceptable ($r > .80$) as found in most previous studies examining SVA (e.g., Steller, 1989; Zaparniuk, Yuille, & Taylor, under review). As expected, all variables scored via computer-based coding showed 100% test-retest reliability (Variables b, c, d, e, f, h) including Variable f (Unnecessary Connectors) which contained a minor element
of coder judgment. The 100% agreement for Variable q (Spontaneous Corrections) is misleading. As indicated in Table 1, the mean frequency of this variable in the accounts from all conditions was very low. A visual inspection of the data from the sample interviews used for the reliability check revealed that in all cases the coders found no evidence for this variable resulting in all "0" scores and thus perfect agreement. The one variable for which agreement was not acceptable was Variable p (Accounts of Subjective Mental State) showing $r = 0.67$.

7.2 Preliminary Analyses. Initially (prior to coding for the other variables), it was necessary to compare the mean account lengths (Variable b) in the four conditions. If there were no differences, coders could proceed with the process of examining frequencies for various variables (e.g., frequency of self-references). If there were disparities, frequency coding would not have been appropriate and an alternative approach would have been selected (e.g., Likert scale). A one-way ANOVA was conducted on Variable b. As Table 1 suggests (see Appendix for corresponding standard deviations), no significant differences in the number of words reported were found among the four groups, ($F[3, 56] = 0.264, p > .05$). Thus, the frequency coding approach was validated.

For Variables "i" (ratio of component elements) and "n" (presence of peripheral details) only three of the four
conditions were examined. These data were not included in the MANOVA. Instead, a one-way three-group ANOVA was conducted for each of the two variables. As Table 1 indicates no mean differences were detected for either Variable "i" ($F[2,42] = 0.23, p>.05$) or Variable "n" ($F[2,42] = 1.02, p>.05$).
Table 1.

Mean scores for the four conditions on all dependent variables.

<table>
<thead>
<tr>
<th>EXPERIMENTAL CONDITION</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DEPENDENT VARIABLE</strong></td>
<td>Truthful Alibi</td>
<td>Partial Deception</td>
<td>False Alibi</td>
<td>Truthful Confession</td>
</tr>
<tr>
<td>A. PERSPECTIVES ON MEMORY ACCURACY</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Details***</td>
<td>50.51</td>
<td>36.03</td>
<td>31.34</td>
<td>55.10</td>
</tr>
<tr>
<td>Words</td>
<td>828.62</td>
<td>751.91</td>
<td>861.14</td>
<td>901.10</td>
</tr>
<tr>
<td>Coherence**</td>
<td>1.73</td>
<td>0.80</td>
<td>0.87</td>
<td>1.93</td>
</tr>
<tr>
<td>Sp. Reproduction</td>
<td>1.60</td>
<td>1.27</td>
<td>1.33</td>
<td>1.93</td>
</tr>
<tr>
<td>Un. Complications</td>
<td>0.73</td>
<td>1.00</td>
<td>0.87</td>
<td>1.00</td>
</tr>
<tr>
<td>Un. Details</td>
<td>1.47</td>
<td>1.20</td>
<td>1.00</td>
<td>1.73</td>
</tr>
<tr>
<td>Peri. Details</td>
<td>1.00</td>
<td>1.53</td>
<td>-</td>
<td>1.13</td>
</tr>
<tr>
<td>R.E. Associations</td>
<td>0.83</td>
<td>1.62</td>
<td>2.53</td>
<td>0.60</td>
</tr>
<tr>
<td>S.M. State</td>
<td>1.27</td>
<td>2.13</td>
<td>1.13</td>
<td>1.33</td>
</tr>
<tr>
<td>Spont. Corrections</td>
<td>0.33</td>
<td>0.40</td>
<td>0.20</td>
<td>0.60</td>
</tr>
<tr>
<td>A.L. of Memory*</td>
<td>4.73</td>
<td>1.87</td>
<td>2.13</td>
<td>4.40</td>
</tr>
<tr>
<td>B. PERSPECTIVES ON DECEPTION</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal Hedges</td>
<td>2.87</td>
<td>3.33</td>
<td>4.53</td>
<td>4.80</td>
</tr>
<tr>
<td>Filled Pauses</td>
<td>21.15</td>
<td>23.90</td>
<td>27.62</td>
<td>27.33</td>
</tr>
<tr>
<td>Self-references</td>
<td>70.70</td>
<td>61.91</td>
<td>83.32</td>
<td>83.73</td>
</tr>
<tr>
<td>Type-token Ratio</td>
<td>0.28</td>
<td>0.29</td>
<td>0.29</td>
<td>0.28</td>
</tr>
<tr>
<td>C. SAPIR'S CRITERIA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Un. Connectors</td>
<td>6.73</td>
<td>5.73</td>
<td>5.93</td>
<td>6.07</td>
</tr>
<tr>
<td>Pron. Deviations</td>
<td>7.07</td>
<td>9.33</td>
<td>8.67</td>
<td>6.33</td>
</tr>
<tr>
<td>Comp. Elements</td>
<td>0.53</td>
<td>0.51</td>
<td>-</td>
<td>0.43</td>
</tr>
</tbody>
</table>

*Many criteria fall under multiple categories.

*p<.05

**p<.001

***p<.0001
7.3 Multivariate Analysis. The design of the experiment is a four group (levels of deception) between-subjects design with eighteen dependent variables. With this number of variables, simple analyses of variances (ANOVAs) could result in an elevated Type 1 error rate. Consequently, as in previous SVA studies (e.g., Joffe, 1992; Landry & Brigham, 1992), a one-way multivariate analysis of variance (MANOVA) was conducted using the remaining fifteen dependent variables. There was a significant main effect for condition (Wilk's lamda = 0.13; $F_{[48,122]} = 2.51, p < .0001$). Subsequently, univariate F-tests were conducted for each of the criteria. These were succeeded by multiple comparisons employing the Tukey method in cases where significant group differences were detected.

Criteria Showing Significant Effect.

Variable a (Amount of Detail Reported). Bartlett's test for homogeneity of group variances indicated homogeneity ($X^2[3] = 2.42, p > .05$). As Figure 1 indicates, there was a highly significant main effect for condition on the number of details reported, $F(3, 56) = 8.48, p < .0001$. Followup multiple comparisons revealed that participants in the truthful alibi condition (Condition 1) reported more details ($M = 50.5$) than participants in the partial deception condition (Condition 2; $M = 36.0, p < .05$) and the false alibi condition (Condition 3; $M = 31.3, p < .01$). Similarly, participants in the truthful confession group
(Condition 4) provided more details (M = 55.1) than participants in Condition 2 (p < .01) and Condition 3 (p < .001). Mean scores from the two truthful conditions (1 and 4) did not differ nor did mean scores from the deceptive conditions (2 and 3).

**Variable j (Coherence).** Bartlett's test indicated homogeneity of group variances (X[3] = 0.71, p > .05). There was a significant main effect for condition on ratings of coherence, F(3, 56) = 6.22, p < .001 as revealed in Figure 2. Subsequent Tukey comparisons revealed that accounts by Condition 1 participants (truthful alibis) were rated as being more than twice as coherent (M = 1.73) as Condition 2 (partial deception) accounts (M = 0.80, p < .05) and Condition 3 (false alibi) accounts (M = 0.87, p < .05). Similarly, reports given by Condition 4 (truthful confession) participants were rated as more than twice as coherent (M = 1.93) as those provided by participants in both Condition 2 (p < .01) and Condition 3 (p < .05). Again, mean ratings for coherence did not differ between the truthful conditions (1 and 4) or the deceptive conditions (2 and 3).
Figure 1. Number of details reported across conditions.
Figure 2. Coherence ratings across conditions.
Variable r (Admitting Lack of Memory). Bartlett's test indicated homogeneity of group variances ($X[3] = 2.73$, $p>.05$). There was a significant main effect for condition on Variable r, $F(3,56) = 3.00$, $p<.05$ (see Figure 3). Tukey tests indicated that Condition 1 participants admitted to lacking memory for an aspect of the target event more frequently ($M = 4.73$) than did Condition 2 participants ($M = 1.87$, $p<.05$). No other pairwise differences attained statistical significance although absolute mean differences between the truthful groups and deceptive groups were large (see Table 1).

7.4 Relative Discriminating Ability of Criteria Showing Significant Effects. A discriminant analysis was carried out to assess the relative contribution of the three criteria showing significant effects in the MANOVA to discriminations between the accounts of truthful and untruthful suspects. Because no differences were found between the truthful groups (Conditions 1 and 4) or the deceptive groups (Conditions 2 and 3) on any criteria the data were pooled for the discriminant analysis to examine the ability of the criteria to globally distinguish truthful and deceptive reports.

As expected, the initial multivariate test statistic was highly significant, reiterating the initial MANOVA ($W_{ilk's \, lamda} = 0.54; \, F[3,56] = 15.75, \, p<.0001$). Similarly, the three univariate F-tests were all highly significant.
Figure 3. Number of admissions to lack of memory across conditions.
The analysis using the three criteria showing significant effects as predictors of membership in the truthful and deceptive conditions yielded a highly significant discriminant function \( \chi^2(3) = 34.56, p < .0001 \). As can be seen from the canonical coefficients presented in Table 2, this discriminant function suggests that the three criteria examined distinguished the truthful from deceptive accounts about equally.

The discriminant function correctly classified 47 of the total sample of 60 (30 truthful and 30 deceptive) accounts according to group membership. Classification success was approximately equal for both groups. Specifically, 23 of the 30 (76.7%) truthful accounts and 24 of the 30 (80.0%) deceptive accounts were correctly classified according to the discriminant function.

7.5 Followup Exploratory Analyses. Although the MANOVA indicated that the truthful and deceptive groups differed on three criteria, it was not clear at what point in the interrogations these differences were emerging. For example, in the partially deceptive reports it was possible that the deception clues were emerging chiefly before the lying behavior occurred (typically near the end of the free recall accounts) or following the initial lying. To answer this question, 2 x 2 (truthful/deceptive condition x phase of interrogation) between-within factorial analyses were conducted using the data from Condition 1 (truthful alibi)
Table 2.

Discriminant analysis of criteria showing significant effect for classification of truthful and deceptive accounts.

<table>
<thead>
<tr>
<th>Discriminant function: $\chi^2(3) = 34.56, \ p&lt;.0001$</th>
</tr>
</thead>
</table>

**Standardized Discriminant Function Canonical Coefficients**

<table>
<thead>
<tr>
<th>Number of Details Reported</th>
<th>0.525</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coherence Ratings</td>
<td>0.619</td>
</tr>
<tr>
<td>Lack of Memory Admissions</td>
<td>0.569</td>
</tr>
</tbody>
</table>
and Condition 2 (partial deception). The previous multiple comparisons indicated that these were the only two groups who differed significantly on all three criteria. Nonetheless, the data for the criterion Coherence were not amenable to the present exploration in that they were in the form of global ratings rather than by frequency and could not be dichotomized into interrogation stages. For the other two criteria (number of details and lack of memory admissions) the data were coded in terms of when in the interrogation they had been elicited - during the free recall phase or during the questioning phase, constituting the repeated-measures aspect of the present analysis.

The first 2 X 2 factorial was conducted on the number of details reported. As expected, the between-subjects component replicated the results of the MANOVA with more details being reported by truthful participants ($F[1,28] = 4.48, p<.05$) overall. As well, there was a main effect of phase of interview ($F[1,28] = 15.18, p<.001$) with more details being reported overall during free recall ($M = 26.45$) than directive questioning ($M = 17.48$). No interaction effect was found ($F[1,28] = 0.215, p>.05$) with truthful and deceptive suspects showing similar reporting patterns. Specifically, truthful suspects reported a mean of 30.27 details during free recall (60.0% of total details) and 20.23 (40.0%) during directive questioning. Similarly, deceptive suspects reported a mean of 22.63 details during free recall (60.6% of total details) compared to 14.73
details (39.4%) during directive questioning. Thus, the finding that deceptive suspects report fewer details seems to reflect a general trend occurring throughout the interrogation rather than being associated with a particular time in the interview or resulting from a particular questioning format.

The 2 X 2 analysis on the criterion "lack of memory admissions" revealed findings comparable to the first analysis. However, in this case the analysis on the between-subjects variable was not significant ($F[1,28] = 2.89, p>.05$, likely because of reduced statistical power. There was a highly significant main effect of phase of interview ($F[1,28] = 23.17, p<.0001$) with more lack of memory admissions occurring during the directive questioning phase of the interrogation ($M = 1.78$) than free recall ($M = 0.67$). No interaction effect was found ($F[1,28] = 0.532, p>.05$) with truthful and deceptive suspects showing very similar reporting patterns. Specifically, truthful suspects admitted lacking memory on a mean of 0.87 occasions during free recall (28.9% of total admissions) and 2.13 (71.1%) occasions during directive questioning. Deceptive suspects admitted lacking memory on a mean of 0.47 occasions during free recall (25.0% of total admissions) compared to 1.40 admissions during directive questioning (75%). As with number of details reported, the finding that deceptive suspects are less willing to admit lacking memory for some
aspect of the event appears to reflect a general trend occurring throughout the interrogation.

The final exploratory analysis was conducted on Variable f (frequency of unnecessary connectors). Although no effects of condition were found in the MANOVA on this criterion, it was possible that an effect may have been masked by analyzing the accounts globally. According to Sapir, the unnecessary connector will surface more frequently as a lie is being related rather than at other points in an account. To address this possibility, a subsample of data from the accounts of suspects in Conditions 1 (truthful alibi) and 2 (partial deception) was examined. Only the portions of the accounts relating events in the crime room itself (in the case of Condition 2 this would represent the deceptive portion of the accounts) were examined for the frequency of unnecessary connectors. It seemed likely that unnecessary connectors would appear more frequently by deceptive suspects when recounting events occurring in the crime room. It was found that deceptive suspects used a mean of 1.47 unnecessary connectors when relating events in the crime room compared to a mean of 0.93 by honest suspects. Although this indicates a considerably higher rate of unnecessary connector usage in the deceptive accounts (approximately 60%), an independent-samples t-test did not yield significance ($t_{[28]} = 1.11, p>.05$).
7.6 Effects of Interrogation Probes. A second one-way MANOVA was conducted using only the information in the accounts following the three interrogation probes to examine whether such probes affect the behavior of honest and deceptive suspects differentially. Only the eleven dependent variables which had been frequency-coded (and thus could be divided into pre- and post-probes) were included in this analysis (a, b, d, e, f, g, h, l, o, p, and r). The MANOVA did not yield significant results (Wilk's lamda = 0.51, $F[33,136] = 1.05$, $p>.05$) although univariate F-tests did indicate a similar pattern of results to the inclusive analysis with regard to Variable a (Amount of Detail), $F(3, 56) = 3.19$, $p<.05$. 
Chapter 8
DISCUSSION

From error to error one discovers the entire truth. - Sigmund Freud

This thesis represented the first empirical test of the general hypothesis that systematic differences exist between the accounts of honest and deceptive suspects in the interrogation context. This hypothesis was founded on several theoretical approaches subsumed under one of: perspectives on memory accuracy, perspectives on deception, or the proposals of Avinoam Sapir. Employing a methodology aimed at maximizing external validity, it was demonstrated that certain verbal clues, chiefly derived from the SVA technique, clearly distinguish truthful and deceptive reports. Dishonest suspects engaging in verisimilitude gave less detailed and less coherent accounts. Further, they were much less likely to admit failure to recall some aspect of the target event. These three criteria successfully classified 76.7% of the truthful accounts and 80.0% of the deceptive reports, each contributing about equally to the discrimination. As well, the differences found for these criteria seem to reflect global reporting patterns unrelated to the phase of the interrogation.

These results corroborate the findings of other recent investigations looking at the utility of statement analysis (e.g., Joffe, 1992; Landry & Brigham, 1992). In the past, there has been a marked over-reliance on nonverbal clues to
deception by both forensic researchers and practitioners (McCornack, 1992), with a virtual disregard for the possible utility of verbal clues. This is especially problematic given that in a considerable proportion of forensic cases the sole evidence available to investigators is the words of the complainant and the accused. The situation has been improving with the advancement of Statement Validity Analysis (SVA) with children and some empirical research on verbal clues with adults in the mid-1980s, most notably studies by Zuckerman (e.g., Zuckerman et al., 1981; 1985) and Miller and colleagues (DeTurck et al., 1985; 1986). Nonetheless, there have been few previous attempts to integrate the various theoretical perspectives in the area or to devise research with the issue of external validity in mind (Miller & Stiff, 1993).

One theoretical bent considered in this study was founded on memory accuracy. More specifically, the Undeutsch Hypothesis states that accurate memory accounts (reports based on true experience) will differ qualitatively and quantitatively from inaccurate memory accounts (reports based on events not actually experienced). Reality Monitoring posits that accurate ("perceptual") memory accounts contain more external sensorial descriptive information whereas inaccurate accounts contain more references to cognitive processes and subjective information (and thus more self-references, verbal "hedges", and references to cognitive processes).
In sum, the results of the present study buttress the general Undeutsch Hypothesis but do not support the use of the Reality Monitoring model for credibility assessment with suspects. One of the more intriguing and potentially consequential findings yielded in this thesis lends credence to Undeutsch's claim that accounts based on true experience will contain a richness of detail not characteristic of fictitious accounts. Specifically, participants relating a completely false alibi provided many fewer details than truthful participants describing an event within the same time frame, consistent with the Zuckerman et al. (1985) meta-analysis. However, the Undeutsch Hypothesis fails to illuminate the foundation of a related finding; the partially deceptive suspects also reported fewer details than the both truthful groups. The striking aspect of this discovery is that these groups were, with a minor exception (folder/money), relating exactly the same occurrence (which had actually been experienced). Obviously, this difference is not associated with the memories themselves (assuming guilty suspects didn't repress information!) but instead with the act of providing deceptive information to an interrogator. Similarly, accounts by partially deceptive suspects were rated as less coherent than those by truthful suspects. These findings are more congruous with perspectives on deception (discussed later).

As in other recent explorations (e.g., Joffe, 1992), the theory of Reality Monitoring did not garner support in
this study. In accordance with this theory, it had been hypothesized that statements by participants in the truthful conditions would contain less references to cognitive operations, fewer verbal hedges ("I believe", "I think"), and less self-references than those of the participants providing a false alibi. None of these criteria showed significant effects. On the other hand, an examination of the absolute mean differences indicates that potential effects may have been masked by sizable variances (and hence elevated Type 2 error rate). For example, deceptive participants provided approximately twice as many references to subjective mental state as both truthful conditions, as predicted by Reality Monitoring. As well, the five accounts with the highest frequency of verbal hedges were by deceptive suspects. In particular, one suspect providing a false alibi employed twenty-three hedges, approximately twice as many as any other suspect in the study. Again, further research is needed to clear up this ambiguity. Of course, it is possible that the Johnson/Schooler variables are indeed effective discriminators of genuine (externally generated) and false (internally generated) memories for which they were originally formulated. All truthful and deceptive accounts in the present study were externally generated; similarly all dishonest suspects were aware of their deceptions. An interesting possibility is that Reality Monitoring may have utility in cases in which the
suspect's or witness' memory is unintentionally erroneous or false (e.g., resulting from a previous suggestion).

A second major theoretical perspective considered in this thesis underscores the betrayal of deception in language due to affective and cognitive factors. Emotional-motivational approaches posit that fear of detection results in arousal and associated verbal deception "leakage". Cognitive approaches are founded on the tenet that deceptive accounts are distinctive quantitatively and qualitatively because of rehearsal, information processing strategies, and cognitive load.

One of the classic indices of deception associated with affective factors is a high frequency of pauses and filled pauses during speech (Ekman, 1992; Kohnken, 1985; Zuckerman et al., 1985). Ekman (1992) argued that pauses are indeed the most common type of verbal deception clues. Nonetheless, as in other investigations (e.g., Stiff & Miller, 1986), the present study did not support the use of the pause as a criterion for credibility assessment. Truthful and deceptive suspects showed a very similar, generally high propensity for filled pause usage. The experience of being interrogated and motivation to convince the interrogator of their veracity appears to have elicited many filled pauses in all suspects, honest or deceptive. Given this ubiquitous nature, the filled pause is probably not an appropriate criterion for assessing credibility in the interrogation context. This is consequential in light
of previous research indicating that people rely heavily on the filled pause in making judgments of veracity (Stiff & Miller, 1986).

Fitting best within the framework of emotional-motivational theory are the revelations concerning the coherence of the reports. Oliver Wendell Holmes once observed that "People who honestly mean to be true really contradict themselves much more rarely than those who try to be consistent." In the present investigation, it was found that both truthful groups provided more coherent, logical accounts than both deceptive groups. The principal evidence to suggest affective underpinnings emerges in the less coherent reports of the partially deceptive suspects. If reduced coherence were related only to cognitive factors (e.g., difficulty in recalling and maintaining previously rehearsed false information), it would be expected in completely untruthful reports (i.e., false alibis) but not in reports containing a very small proportion of deceptive statements (i.e., partially deceptive group). Again, it is likely that heightened arousal from lying produced a general tendency for relative incoherence across the accounts.

The intriguing possibility that this criterion may be useful in forensic settings with adult suspects has been unveiled in this study. However, a strong word of caution is demanded; forensic suspects or witnesses may very well show different coherence levels than the individuals in this type of study. For example, as noted by Gudjonsson (1992)
high proportion of innocent and guilty suspects are in a state of confusion, shock, and incoherence during the police interrogation for several reasons including the nature of the crime, intoxication, sleep deprivation, or psychopathology. Similarly, a genuine adult victim and an adult feigning victimization may both provide relatively incoherent accounts. Nonetheless, such suspects and victims may show a similar global pattern to the participants in this thesis; that is, the reports of honest individuals may show relatively higher levels of logical structure than dishonest individuals. Further research is needed to evaluate this hypothesis. As well, previous investigations (e.g., Joffe, 1992) have indicated that coherence may not be a particularly effective discriminator of truth and deception in children's accounts. It is possible that in younger children the experience of lying may produce less pronounced affective reactions because of an under-developed moral "conscience" against lying (Kohlberg, 1984; Lewis, 1993).

In general, cognitive approaches to deception are highly functional in the interpretation of the present results. Enigmatically, the much less detailed accounts did not contain fewer words giving credence to the proverb "Words should be weighed and not counted." To restate, the liars spoke as much but the information communicated in their speech was significantly less. A possible explanation for this finding was formulated by Kohnken (1985) who argued
that increased cognitive load should result in a reduction in the informational content of speech. In other words, deceptive suspects have the added burden of thinking about the story they are reporting which detracts from the quality of the information reported. Whereas deceptive suspects recall words previously rehearsed and worry about "account maintenance", truthful suspects engage in the less complex cognitive task of recalling details from a memory image. It is also conceivable that the deceivers reiterated details over the course of their accounts. The scoring procedure employed involved coding repeated details as a single detail. Details which had been rehearsed to "keep the story straight" may have appeared repeatedly in these reports. Miller and Stiff (1993) argued that deceptive statements contain fewer details because "unnecessary elaboration increases the opportunity for contradiction and the likelihood of detection" (p. 64). An alternative affect-based interpretation is that the deceptive subjects repeated words due to anxiety (as found by Beattie, 1981). However, this seems less plausible given that the lexical diversity scores did not differ between the truthful and deceptive groups.

Information processing theories received mixed support in this study. McCornack's Information Manipulation Theory (IMT), although somewhat nebulous in nature, seems to be applicable to these findings. As predicted, deceptive suspects failed to fulfill the informational requirements of
the interaction (i.e., provided fewer details) and provided information in a less orderly, coherent fashion. It is not clear whether the details provided by deceptive suspects were "less relevant" to the central issue being discussed (the theft) than honest suspects. Further research is required to test IMT's utility in explicating credibility assessment principles.

Other cognitive approaches were not bolstered in this study. Responses following negative feedback from interrogators through the use of interrogation probes did not appear to effectively discriminate truthful from untruthful suspects, contrary to the superficial-encoding hypothesis. Nonetheless, it is still possible that such probes might aid in credibility assessment with certain criminal suspects. For example, the present author recently read the interrogation transcripts from an ongoing American homicide case in which they were highly effective. Immediately following probes very similar to those employed in this study, the prime suspect drastically altered his account of the event, from witnessing someone else shooting the victim to admitting that he had pulled the trigger accidentally. Lexical diversity was also unrelated to deception in this study. One explanation is that elevated lexical diversity results from anxiety associated with the experience of being interrogated rather than anxiety due to guilt or deception (Dulaney, 1982). On the other hand, systematic changes in lexical diversity within suspects'
accounts were not examined which might be a useful endeavor in the forensic context (Carpenter, 1990).

The finding that suspects providing a partially deceptive account were less willing to admit lack of memory than those providing a truthful alibi certainly requires further investigation. It could be associated with the depiction of confidence and certainty by deceivers attempting to appear credible. This seems a cogent strategy given that a strong determinant of perceived credibility is confidence (e.g., Leippe & Romanczyk, 1989; Wells, 1993). On the other hand, meta-analyses have indicated a very weak positive relationship between accuracy and confidence in witnesses (who are attempting to provide accurate information). The current study suggests that there may be a strong negative relationship between confidence and accuracy in criminal suspects (who are being consciously inaccurate). Nonetheless, this explanation remains purely speculative in that the proposed confidence/lack of memory acknowledgment relationship has never been established. A second possibility relates to the finding that deceptive suspects provided less detailed accounts to interviewers. Lack of memory admissions typically occurred in response to directive questioning from interrogators rather than during the free recall stage of the interviews. The provision of fewer details by deceptive suspects may have resulted in fewer directive questions being posed by the interviewers and hence fewer opportunities for lack of memory admissions.
Adding to this quandary is the fact that suspects relating completely falsified accounts did not employ fewer admissions. Why was the effect evidenced in suspects lying only about the theft itself but not in suspects providing entirely deceptive information? One can only surmise that for the completely deceptive suspects lack of memory admissions functioned to facilitate account maintenance (i.e., it is easier for a suspect to maintain a story if he/she purports to recall less about it), whereas the partially deceptive suspects were being more heavily influenced by one of the factors discussed above. Future research should address this mysterious finding.

Sapir's proposals concerning differences in pronoun deviation, component elements, and unnecessary connectors did not fare well in this study. However, their merits have not been singly obliterated by this experiment. For example, both deceptive groups showed substantially (but not significantly) more pronoun deviations than the truthful groups. More research should address this interesting possibility that attention to pronoun usage may aid in credibility assessment. Secondly, and more importantly in my opinion, the unnecessary connector has not really had its "day in court" either. Variables like the unnecessary connector are extremely difficult to evaluate through research simply because of the possibility of their rarity; it may be that an unnecessary connector surfaces in exceptional forensic cases but with momentous consequence
when detected by an experienced investigator. Even in the present study, although mean differences were miniscule, a glaring unnecessary connector was employed by a particular suspect (Condition 2) attempting to conceal information relating to the theft. During the free narrative account, he stated:

I had to look for a folder that was supposed to be in a...in a box but I couldn't find it and then...uhhh...I looked around a bit more. Then after that I just went out of that office, made sure the door was locked again, went back the same way...

With little more than anecdotal evidence (Sapir has provided many such examples), a conservative stance must be taken. It would be advisable to emphasize to law enforcement personnel that unnecessary connectors may, in certain cases, be indicative of the concealment of information, suggesting a topic requiring clarification or more in depth exploration but certainly to never rely singularly on this variable.

Three of the ten (30%) SVA variables investigated in this thesis (coherence, sufficiency of detail, and admitting lack of memory) significantly discriminated the truthful and untruthful accounts. This performance is encouraging and gives reason for the continued quest for reliable clues to credibility in forensic contexts. It is especially promising given related findings with adult witnesses (e.g., Brigham & Landry, 1992 found that 71% of the fourteen SVA variables they examined were effective). Indeed, the performance of SVA in this study may be considered remarkable in that the methodology may not have been
conducive to the successful employment of SVA. Raskin and Esplin (1991) recommend that in designing generalizable experimental research addressing SVA the event used should be involving, have a component of lost control, and a negative tone. The present study lacked a component of lost control with its intention to examine the utility of SVA with suspects rather than victims or witnesses. The methodology employed sought to create components of involvement, anxiety during the commission of the "offense", motivation to relate a deceptive or truthful account successfully, and realistic negative feedback in the interrogation context. As earlier discussed, SVA was originally developed for use in cases of child sexual abuse (Undeutsch, 1982) for which certain criteria may be context-dependent. Overall, this thesis and previous studies seem to paint a picture of SVA as a potent general credibility assessment technique whose criteria are differentially useful across contexts.

One of the more striking serendipitous findings in this study mirrored, to a lesser extent, Zimbardo's classic prison studies from the early 1970's (e.g., Zimbardo, 1972). As with the guards and inmates in his experiments, the interrogators and suspects in this study appear to have become heavily enmeshed in their respective roles. An examination of the first couple of interviews showed the interviewers becoming visibly aggravated with the participants when they suspected they were being lied to
Correspondingly, the participants reacted defensively and angrily to this negative feedback (e.g., one truthful subject stated "I brought the folder up and now you're accusing me of lying. What's the problem with you? Jesus, people these days!"). It was necessary to reiterate to the interviewers to treat all suspects with parity resisting the urge to evaluate the honesty of the suspects and to maintain a relatively polite, unpassioned tone. As well, the majority of suspects reported feeling very wary of each person encountered en route to the crime room with the suspicion that he/she might have been the guard. Some participants reported feeling guilty but exhilarated following the commission of the theft. Although such reports may lend support to the validity of this study, they also show how amazingly expeditious such a simulated role "was absorbing [the participants] as creatures of its own reality" (Zimbardo, 1972, cited in Myers, 1990).

8.1 Limitations of the Current Study. In this investigation, an effort was made to overcome several methodological shortcomings evident in previous studies addressing credibility assessment. Specifically, the simulated crime and interrogation were relatively involving and anxiety-provoking and as analogous to the forensic context as ethically permissible. Both honest and deceptive suspects were motivated to appear truthful and were supplied
time to prepare their accounts. The intimated presence of a guard and motivation of the "administration's interrogators" served to further increase the validity of this study. As described earlier, this validity seems to have been attained given the comments of the participants during the interrogation and debriefing phases.

Despite these methodological strengths, others have argued that the accuracy of credibility techniques can never be measured in the laboratory because the emotional characteristics of the suspects and witnesses are too dissimilar from those involved in real-life investigations (Arntzen, 1983; Iacono & Patrick, 1987). Although this perspective has been criticized as revealing "a lack of understanding of empirical research needs and methods" (Steller, 1989, p. 144), it certainly highlights a valid concern with generalizability.

In considering the generalizability of this research, the issue of motivation becomes salient. It is clear that the monetary incentive offered participants to be convincing in an experimental setting cannot approximate the motivation present in forensic suspects - freedom. Not only is the simulation problematic quantitatively but also qualitatively; in the present study a dynamic of negative reinforcement was created (i.e., "If you fail to convince your interrogator of your veracity you will not receive extra money") whereas in real life a punishment contingency is more prevalent (i.e. "If you fail to convince your
interrogator of your veracity you will go to jail"). It is not clear if or how this motivational difference might affect language behavior during the interrogation.

Intention to deceive must also be considered in exploring the issue of generalizability. In some cases, innocent suspects may come to believe that they committed a crime contrary to traditional legal doctrines (e.g., "The presumption upon which declarations are evidence is that no man would declare anything against himself unless it were true" - Judge Eyre in Hardy's Trial, 1794, cited in Arcaro, 1990). This was exemplified in the Christie case described earlier; Timothy Evans became so convinced by police that he had killed his wife that he eventually recalled the act. Subsequent to Evans' execution, it was proven that Christie had solely committed the homicide. Similarly, although not examined in the present study, voluntary false confessions sometimes occur. This phenomenon was exemplified in the case of serial murderer Henry Lee Lucas who, in the early 1980's, falsely confessed to several hundred murders supposedly for the increased media attention he received with each confession (Cox, 1991). It is not known whether the verbal clues examined in this study would have utility in these types of cases. In this study, it was found that truthful confessions tend to be quite detailed. It is likely that false confessions will be less detailed than truthful confessions unless the false confessor has a great deal of knowledge about the crime scene from other sources.
Another aspect of this study which must be examined in considering external validity is the population employed. Undergraduates are certainly not representative of the criminal class of society. For example, research has consistently demonstrated that middle-class youths show lower rates of criminal behavior than working-class and lower-class youths (e.g., Elliot & Huizinga, 1983; Wolfgang, Figlio, & Sellin, 1972). Evidently, undergraduates have less experience being interrogated which would likely affect their behavior in this context.

As well, undergraduates may show different intelligence and verbal abilities than most criminals, affecting their language behaviors in general. The results in this thesis may not extend to psychopaths, who constitute 15-25% of the criminal population (Hart & Hare, 1989). Exciting recent research by Hare and colleagues indicates that psychopaths process and communicate language on a different level than most people (Day & Wong, 1993; Hare & Jutai, 1988; Louth, Williamson, Alpert, & Hare, in preparation; Williamson, Harpur, & Hare, 1991). For example, the Louth et al. study evinced that psychopaths speak with lower amplitude but with a faster rate than nonpsychopaths and do not distinguish between emotional and neutral words as did the nonpsychopaths. With a greater insensitivity to the emotional connotations of language, it is probable that dissimilarities would also emerge in the variables examined in the present study. Future research should address the
possibility that psychopaths show fewer verbal differences when being honest and deceptive than nonpsychopaths because this insensitivity (cf. Raskin & Hare, 1978). It should also be noted that pathological lying, deception, and manipulation are some of the key diagnostic features of the psychopath (e.g., Hare, Forth, & Hart, 1988) leading to experiential adeptness in lying and corresponding behaviors in the interrogation.

Finally, cultural distinctions may also restrict the results of this thesis which employed North American English-speaking participants. Aside from obvious cross-cultural language variations, cultural norms may also dictate a suspect's behavioral repertoire in the interrogation context (e.g., see Vrij & Winkel, 1991). Similarly, sex differences were not examined in this thesis because of a female participation bias.

One of the problems associated with the above factors which must be considered when evaluating generalizability is the potential complexity of their effects; their presence might increase the frequency of certain deception indicators and decrease the frequency of others. For example, it might be that a psychopath experienced with the interrogation process would recount a (rehearsed) coherent, detailed false alibi, but frequently admit lacking memory for aspects of the event to avoid further probing by interrogators. It is evident that much research is needed to clarify these possible relationships.
8.2 Directions for Future Research. As noted by several forensic researchers (e.g., Davies, 1990; Yuille, Davies, Gibling, Marxsen, & Porter, in press; Porter, Yuille, & Bent, in press; Yuille, 1993), only when we have converging results from diverse methods of inquiry should we generalize to forensic contexts. The present research was intended as a preliminary examination of verbal clues to deception in which ground truth is established. Before intercalating my results into the forensic interrogation context much further inquiry is essential, especially field studies. Cross-validation studies are crucial with this type of applied research. In certain criminal cases it is possible to determine with a fair degree of certainty whether a suspect had been lying or telling the truth during an interrogation (e.g., in those cases in which the suspect subsequently confesses or another suspect confesses). As observed by Miller and Stiff (1993) law enforcement agencies frequently record investigative interviews with of suspects, witnesses, and victims as well as pre-polygraph and polygraph examinations, providing fertile ground for researchers interested in deception. The results of the present study must be applied to forensic populations to test the ecological validity of this experimental paradigm and to examine whether these findings would have utility in actual investigations. Importantly, in the one published field study applying statement analysis to an actual case,
quantity of details and coherence were two of main criteria employed in assessing the credibility of a suspect (Yuille & Cutshall, 1989).

Unfortunately, the finding that dishonest suspects tend to give less detailed and coherent accounts in which there are fewer admissions to lack of memory does not tell us how to apply this knowledge to a particular case. Within-subjects experimental designs are essential to examine whether verbal characteristics change as a function of veracity within individual suspects. If so, one way to utilize this knowledge would be to subtly ask a suspect to relate an event other than the crime being investigated during the interrogation (J.C. Yuille, personal communication, July 5, 1994). Patterns within this account would then be compared to those in the account relating to the crime (e.g., an alibi). Similarly, there is a need for establishing the base rates of any verbal characteristics examined for credibility assessment purposes (e.g., level of detail and coherence, frequency of lack of memory acknowledgments) in the forensic context. As well, virtually no research has addressed the base rate of lying in the criminal interrogation. This knowledge is important for practitioners of credibility assessment in any applied context.

As previously mentioned, verbal clues to deception may be differentially useful across contexts. As well, there may be a host of verbal clues to deception not examined in
this thesis. For example, after being asked if she had seen any evidence of a theft while retrieving the folder, one participant responded "No. . . I didn't see any envelope with a hundred dollar bill on it" expressing knowledge only a guilty suspect would possess. Although unclassifiable with the set of criteria coded in this study, attention to such verbal slips may obviously give clues to credibility. Another intriguing example from an actual case is reported by Avner Less who interrogated the infamous S.S. officer Adolf Eichmann, captured in 1960 (Less, 1983). Less stated: "As time went on, I noticed that each time Eichmann said 'Never! Never! Never, Herr Hauptmann!' or 'At no time! At no time!' he was lying. That was always a cue for me to ask my colleagues for additional material with which to probe the sensitive spot" (Less, 1983, p. xxi). The next formidable but attainable step for researchers in credibility assessment is to seek which verbal clues are effective with whom and under which circumstances.
REFERENCES


DECEPTIVE LANGUAGE


Appendix A

DEBRIEFING

Thank you for your participation. Following is a brief description of the rationale behind the project in which you have just participated:

My area of research is the use of verbal clues for the purpose of credibility assessment or "lie detection" with criminal suspects. What is different about the language of a person telling the truth and a person being deceptive during an interrogation? Although fieldwork is planned with a criminal suspect population, this analog study represents the preliminary controlled research (in which we always know who is lying and who is not).

In this study, participants were divided into four conditions: truthful alibi, partial deception, false alibi, and truthful confession. All participants proceeded to the crime room where most (except those in the truthful alibi condition) took a $100 bill. On returning to the lab, a monetary incentive was offered to motivate participants to lie successfully to the interviewer. The interviewers used a "Step-Wise" interview format as well as interrogation probes (e.g., "It is obvious that you are lying to me. Why don't you just admit it?").

Eighteen variables relating to your statements in the interview are being examined. For example, it is hypothesized that deceptive suspects will use more filled
pauses than honest suspects because of cognitive factors associated with lying under pressure.

Some deceptions were used by the experimenter in this study:
1. There was no guard. There has not been a recent rash of thefts in the building.
2. This was not a "security effectiveness" study. It was a study of the language behaviors of truthtellers and liars.
3. It was not the agenda of the interviewers to establish the truth. They were asked to obtain as much information from all suspects and to treat all suspects the same. It did not matter whether the interviewer believed you or not. All participants received the extra $5.

It was necessary to employ these deceptions to increase external validity. The presence of the fictitious guard was implied to simulate the situation of an actual criminal committing a theft. The extra $5 was offered to increase participants' desire to successfully convince the interviewer of their truthfulness (the typical criminal suspect is highly motivated to do so). With the design employed, I wanted to approximate a real interrogation as closely as possible.

Because experimenter deception is a crucial component of this research it is extremely important that you not discuss this study with anyone until all subjects have been run (i.e., May, 1994). We appreciate your cooperation in this regard.
Feel free to ask any questions to either Steve Porter or Dr. John C. Yuille. If you are interested, the results of this study will likely be available to you by the end of September, 1994.
Appendix B
INTERROGATION PROTOCOL

1. INTRODUCTION

Before you begin the interview, test the tape recorder. State your full name, the date, subject's ID number and ask the subject to state his/her full name. Play back the tape to ensure that the machine is picking up both of your voices clearly. If during the interview, the subject moves significantly or lowers his/her voice interrupt the account. Politely request that he/she resume the original position or voice amplitude.

2. FREE NARRATIVE

It is important that a dynamic is established in which it is the subject who does most of the talking. State to the subject "It is my job to determine which of the suspects in this study committed the theft of a $100 bill from Room #---- downstairs. I'd like you to tell me everything you can remember, no matter how trivial or irrelevant it may seem." Start from the time you exited this lab and keep going until the time you returned to the lab. I am interested in determining your whereabouts and actions in the last fifteen minutes." Do not interrupt the subject during the free narrative, even if the subject starts to describe irrelevant details of the event or contradicts him/herself. Allow the subject to go at his/her own pace and be patient when the subject pauses. If, however, it
seems that the subject is not going to continue the account you should attempt to restart the narrative. The best method for doing this is simply to state "What happened next?" or "You were saying that (here restate the last thing the subject said). And then what happened?".

Make very brief notes regarding the content of the narrative. These notes can consist simply of the noun and verb phrases mentioned by the subject (e.g., a desk). When the subject has come to the end of the narrative it is important that you ask him/her if anything more can be remembered. If not proceed to the next step.

3. OPEN QUESTIONING

Open questions are simply requests for more information about the events, actions, and objects mentioned in the free narrative. The easiest way to do this is to refer to your notes taken earlier and asking if the subject recalls anything more about each of the noun and verb phrases (e.g., "You mentioned some books. Can you tell me anything more about those?"). Before asking open questions it is imperative that you let the subject know that "No, I don't remember" is a completely acceptable response. Never ask questions like "You told me you saw some books. Tell me more about those." Similarly, your open questioning should not be leading. Do not include information from other sources in these questions. If all previous suspects mentioned seeing a picture on the wall it is easy to make
the mistake of asking a subject who has not mentioned it if he/she can remember any more about the picture.

4. **INTERROGATION PROBE #1**

Following the open questions, express a puzzled look and state "Frankly, I'm a little skeptical about your story. Are you being completely honest?". When the subject responds positively (i.e., "Yes, of course") say "Okay, but something doesn't sound quite right." If the subject responds negatively (i.e., "No") ask what part of the account is not true. If an alternative response is encountered the interviewer should respond politely but still convey suspicion (e.g., "Why? Are you accusing me of something?"); Response - "No, I'm not accusing you of anything. It's just that I get the feeling that you are not being completely honest").

5. **SPECIFIC QUESTIONING**

If, after asking the open questions, there are still points from the subject's account that require elaboration you should now resort to specific questions (this is generally the case). It is especially important here to identify any inconsistencies in the report and ask the subject to clarify them. The best method is to simply plead confusion. You can say something like "I'm not certain I understand everything. You said that you were in the office for about five minutes but later said you spent seven or eight minutes looking for the folder in the office. Can you explain this to me?"
6. **INTERROGATION PROBE #2**

After pointing out any inconsistencies and hearing the explanations you should ask the suspect directly "Do you have any knowledge of the money which was taken from Room #---?

7. **INTERROGATION PROBE #3**

After hearing the response to Interrogation Probe #2, state "Come on now, it's really obvious that you are lying to me. Why don't you just admit it?". Don't say it in a particularly hostile manner. Be relatively polite about it. After the subject provides a response to this statement, conclude the interview. Thank the subject for his/her time and inform him/her that he/she can now leave the interview room, receive their money for participating, and be completely debriefed in the next room.
Appendix C

STANDARD DEVIATIONS FOR TABLE 1 DEPENDENT VARIABLES

<table>
<thead>
<tr>
<th>EXPERIMENTAL CONDITION</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DEPENDENT VARIABLE</strong></td>
<td>Truthful Alibi</td>
<td>Partial Deception</td>
<td>False Alibi</td>
<td>Truthful Confession</td>
</tr>
<tr>
<td>Details</td>
<td>18.90</td>
<td>13.41</td>
<td>13.57</td>
<td>13.91</td>
</tr>
<tr>
<td>Words</td>
<td>351.00</td>
<td>575.87</td>
<td>535.44</td>
<td>386.39</td>
</tr>
<tr>
<td>TTR</td>
<td>0.05</td>
<td>0.06</td>
<td>0.08</td>
<td>0.05</td>
</tr>
<tr>
<td>Verbal Hedges</td>
<td>2.30</td>
<td>2.99</td>
<td>6.10</td>
<td>3.80</td>
</tr>
<tr>
<td>Filled Pauses</td>
<td>11.52</td>
<td>25.29</td>
<td>20.80</td>
<td>14.33</td>
</tr>
<tr>
<td>Un. Connectors</td>
<td>6.76</td>
<td>6.73</td>
<td>6.96</td>
<td>3.96</td>
</tr>
<tr>
<td>Pron. Deviations</td>
<td>11.20</td>
<td>8.67</td>
<td>10.45</td>
<td>7.70</td>
</tr>
<tr>
<td>Self-references</td>
<td>33.75</td>
<td>38.05</td>
<td>48.20</td>
<td>42.50</td>
</tr>
<tr>
<td>Comp. Elements</td>
<td>0.24</td>
<td>0.32</td>
<td>-</td>
<td>0.22</td>
</tr>
<tr>
<td>Coherence</td>
<td>0.96</td>
<td>0.86</td>
<td>0.99</td>
<td>0.79</td>
</tr>
<tr>
<td>Sp. Reproduction</td>
<td>1.18</td>
<td>1.16</td>
<td>1.05</td>
<td>0.70</td>
</tr>
<tr>
<td>Un. Complications</td>
<td>0.96</td>
<td>1.00</td>
<td>0.92</td>
<td>1.36</td>
</tr>
<tr>
<td>Un. Details</td>
<td>0.72</td>
<td>0.93</td>
<td>0.68</td>
<td>0.94</td>
</tr>
<tr>
<td>Peri. Details</td>
<td>0.83</td>
<td>0.78</td>
<td>-</td>
<td>0.88</td>
</tr>
<tr>
<td>R.E. Associations</td>
<td>1.01</td>
<td>1.96</td>
<td>4.17</td>
<td>1.74</td>
</tr>
<tr>
<td>S.M. State</td>
<td>1.91</td>
<td>2.23</td>
<td>1.46</td>
<td>1.68</td>
</tr>
<tr>
<td>Spont. Corrections</td>
<td>0.62</td>
<td>0.91</td>
<td>0.41</td>
<td>0.74</td>
</tr>
<tr>
<td>A.L. of Memory</td>
<td>2.76</td>
<td>1.89</td>
<td>1.80</td>
<td>3.48</td>
</tr>
</tbody>
</table>