Calhoun

Institutional Archive of the Naval Postgraduate School

Calhoun: The NPS Institutional Archive

Theses and Dissertations   Thesis Collection

2015-03

Deception detection process and accuracy: an examination of how international military officers detect deception in the workplace

Calhoun is a project of the Dudley Knox Library at NPS, furthering the precepts and goals of open government and government transparency. All information contained herein has been approved for release by the NPS Public Affairs Officer.

Dudley Knox Library / Naval Postgraduate School
411 Dyer Road / 1 University Circle
Monterey, California USA 93943

http://www.nps.edu/library

NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

DECEPTION DETECTION PROCESS AND ACCURACY: AN EXAMINATION OF HOW INTERNATIONAL MILITARY OFFICERS DETECT DECEPTION IN THE WORKPLACE

by

Boris Kun
Will Whaley

March 2015

Thesis Advisor:	Lisa Lindsey
Second Reader:	Benjamin Roberts

Approved for public release; distribution is unlimited

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE		Form Approved OMB No. 0704–0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington, DC 20503.

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE March 2015	3. REPORT TYPE AND DATES COVERED Master's Thesis
4. TITLE AND SUBTITLE DECEPTION DETECTION PROCESS AND ACCURACY: AN EXAMINATION OF HOW INTERNATIONAL MILITARY OFFICERS DETECT DECEPTION IN THE WORKPLACE		5. FUNDING NUMBERS
6. AUTHOR(S) Boris Kun and Will Whaley		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A		10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. IRB Protocol number ____NPS.2014.0051-IR-EP7-A____.

12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited	12b. DISTRIBUTION CODE

13. ABSTRACT (maximum 200 words)

This thesis replicates recent diagnostic utility studies to determine whether the original methods are (1) generalizable to a new population and (2) useful in identifying specific questioning strategies relevant to international militaries. Previous research shows that people are, on average, only slightly better-than-chance at detecting deception. In 2006, Personality and Social Psychology Review published "Accuracy of Deception Judgments" in which Charles F. Bond Jr. and Bella DePaulo identified that meta-analysis yields an across-study average accuracy rate of about 54%. New research has shifted from the historical cue-based deception detection paradigm in favor of the idea of diagnostic utility. Specifically, this new line of research provides a basis for demonstrating that the design of specific questions is vital in determining deceptive individuals. Currently, the research conducted thus far provides levels of deception detection accuracy significantly greater than the usual slightly-better-than-chance results that is characterized by historical research. Our findings from quantitative Study 1 demonstrated that international military officer participants detected deception at 70.8% for experts and 63.8% for non-experts. Finally, the authors' qualitative Study 2 identified that participant's claim to have utilized third-party information, physical information, and verbal/nonverbal clues most often when detecting deception in previous situations. These findings are in line with historical research.

14. SUBJECT TERMS Deception Detection, Fraud, Diagnostic Utility, Questioning Method, International Officers, Officers, Workplace Deception, Fraudulent Enlistment, Recruiting, Training, Executive Development, Cost Reduction Solutions			15. NUMBER OF PAGES 73
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT UU

NSN 7540–01-280-5500

Standard Form 298 (Rev. 2–89)
Prescribed by ANSI Std. 239–18

THIS PAGE INTENTIONALLY LEFT BLANK

DECEPTION DETECTION PROCESS AND ACCURACY: AN EXAMINATION OF HOW INTERNATIONAL MILITARY OFFICERS DETECT DECEPTION IN THE WORKPLACE

Boris Kun
Ensign, United States Navy
B.S., San Diego State University, 2013

Will Whaley
Major, United States Marine Corps
B.A., University of Arizona, 2003

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN MANAGEMENT

from the

**NAVAL POSTGRADUATE SCHOOL
March 2015**

Authors: Boris Kun

                  Will Whaley

Approved by: Lisa Lindsey
                  Thesis Advisor

                  Benjamin Roberts
                  Second Reader

                  William R. Gates
                  Dean, Graduate School of Business and Public Policy

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

This thesis replicates recent diagnostic utility studies to determine whether the original methods are (1) generalizable to a new population and (2) useful in identifying specific questioning strategies relevant to international militaries. Previous research shows that people are, on average, only slightly better-than-chance at detecting deception. In 2006, *Personality and Social Psychology Review* published "Accuracy of Deception Judgments" in which Charles F. Bond Jr. and Bella DePaulo identified that meta-analysis yields an across-study average accuracy rate of about 54%. New research has shifted from the historical cue-based deception detection paradigm in favor of the idea of diagnostic utility. Specifically, this new line of research provides a basis for demonstrating that the design of specific questions is vital in determining deceptive individuals. Currently, the research conducted thus far provides levels of deception detection accuracy significantly greater than the usual slightly-better-than-chance results that is characterized by historical research. Our findings from quantitative Study 1 demonstrated that international military officer participants detected deception at 70.8% for experts and 63.8% for non-experts. Finally, the authors' qualitative Study 2 identified that participant's claim to have utilized third-party information, physical information, and verbal/nonverbal clues most often when detecting deception in previous situations. These findings are in line with historical research.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

ACFE	Association of Certified Fraud Examiners
CEE	Center for Executive Excellence
CHDS	Center for Homeland Defense & Security
CiC	Content in Context
DOD	Department of Defense
IRB	Institutional Review Board
NPS	Naval Postgraduate School
PII	Personally Identifiable Information
RAN	Royal Australian Navy
SCW	Strategic Communication Workshop
SPEAR	Strategic Planning for Execution: Assessment & Risk
SUE	Strategic Use of Evidence

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

THIS PAGE INTENTIONALLY LEFT BLANK

I.   INTRODUCTION

The body of work concerning interpersonal deception detection has, for decades, focused on the verbal and nonverbal cues of those being questioned. Regardless of the setting or sample, the methods of questioning subjects have not ventured beyond what the subject emits as cues. Over the course of the last decade, however, a new school of thought has emerged. Although there is an abundance of collected works on deception detection, this is the first thesis of its kind at the Naval Postgraduate School to study in detail the use of content- and context-based questioning within a military population. The issues discussed in this thesis have wide-ranging applicability both in the military and accounting/acquisition fraud environments. In evaluating how accurately military officers detect deception and what information is relied upon to make veracity judgments, the authors' overall research will be replicating previous studies conducted by Levine, Blair, and Clare (2014) and Park, Levine, McCornack, Morrison, and Ferrara (2002) on a previously unexamined population: international military officers. The successful completion of this thesis will not only add to the body of knowledge on the topic of deception detection but has the possibility for further replication in classified subsets such as defector interrogation and counterintelligence.

A.   BACKGROUND/PROBLEM

The predominance of nonverbal cue-related research in the field of deception detection has led to stagnation in terms of accuracy improvement. Using the Levine et al. (2014) methodology of diagnostic utility, this thesis will show replicability and applicability within a military context. Furthermore, through applying the principles of Park et al.'s (2002) research, this thesis will examine how the population determines deception.

1.   Deception Detection Accuracy

Until recently, research into deception detection has focused primarily upon interviewees' nonverbal cues. Decades of research on these nonverbal cues show we are barely better than chance, garnering nothing more than a meta-analysis accuracy rate of 53.46% (Bond, 2006). Moreover, Bond, Levine, Park and other researchers have noted

that most people believe in the fallibility of physical human nature, the crossing of legs, the twitching of an eye, the drumming of fingers, and that such subtle nonverbal cues provide the critical clues to identifying when deception is occurring. Indeed, modern pop culture, including popular television programs, shows behavior that demonstrates leakage or "tells" during high-stakes poker matches or a criminal suspect's subtle behaviors during interrogation by law enforcement officers. DePaulo's most recent meta-analysis looking at 208 deception detection accuracy studies tells us the opposite: When relying on nonverbal cues, people are not much better than chance (50%) in any given situation when attempting to determine if somebody is lying. In the literature, an accuracy ceiling of 65% exists (Levine et al., 2014).

The focal point of Study 1 is replicating research conducted by Levine et al. (2014), which indicates that more useful strategies can lead to better deception detection accuracy. Specifically, Levine et al. argue that what has been lacking for decades is attention paid to the specific question strategies used in interview situations. One important concept related to question strategies is diagnostic utility. Diagnostic utility is a scalable level upon which an individual uses information to form a correct conclusion. This scale ranges both positive and negative in that negative information would be viewed as deceptive. As noted by Levine et al. (2014) in their sixth experiment, which is being replicated in this research study (and which involves diagnostic utility), Levine and his colleagues were able to raise accuracy levels in excess of 70%.

## 2.	How People Really Detect Lies

The prevailing deception detection literature and experimentation have been based on the immediacy of nonverbal cues exhibited by an individual and interpreted by a recipient. Though this concept retains validity, it is inherently flawed as it fails to address the additional information a recipient possesses when he/she makes a determination of authenticity. As examined by Park, Levine, McCornack, Morrison, and Ferrara (2002), most individuals in lie-determination settings rely upon information gathered prior to the detection of the lie. This elongated timeline includes "information sources such as information from third parties, the consistency of statements with prior knowledge, the

consistency of messages with physical evidence, and confessions," which when taken into account with nonverbal cues build an individual's truth/lie judgment (Park et al., 2002, p. 144). It is this combined effect of nonverbal and subjective historical evidence that forms the basis of Study 2.

3.   The Current Studies

Study 1 replicates the sixth study in Levine et al.'s (2014) program of research using U.S. military officers to determine whether or not a replication with a different population would yield similar results and, thus, determine the generalizability of their findings. Study 2 replicates the Park et al. (2002) research into the applicability of historical subjective information used in the detection of lies and if such information varies when applied to a different population. The results of both studies will help to better shape the fields of accession, recruiting, training and development, interrogation, associated fields of defector questioning, flag officer promotion board questioning, and fraud detection in international military contexts.

B.   PURPOSE

The purpose of this research is to conduct the experimental replication and documentation of further improvement in deception detection accuracy over that of previous findings along with the parallel study of domestic military officers. The nature of this topic dictates the use of both qualitative and quantitative opportunities to further expand known knowledge of deception detection. Quantitatively, the focus will be on strategic questioning methods and resulting deception detection accuracy. Qualitatively, this research looks at when, how, and with what information deception is actually detected in the workplace.

1.   Study 1: Content and Contextual Questioning Effects on Accuracy

Unlike the bulk of previously conducted research, Study 1 aims to remove the focus on nonverbal cues. Question effects detail how the questioning of an individual may potentially impact the outcome of deception detection. It is specifically the manipulation of such question effects through the use of applied content and contextual

questioning that the authors believe can and will lead to an increase in diagnostic utility. Diagnostic utility is the overarching conceptual idea that information has varying amounts of utility, both positive and negative. It is the application of diagnostically useful statements that the authors measure as a means of deception detection.

2.   Study 2: How People Really Detect Lies

The purpose of Study 2 is to examine what additional information the selected population relies upon in truth/lie determinations and the associated time horizon with regards to deception realization. This study differs from traditional works in that, rather than focus on the questions or questioning method, it instead places emphasis on the individual detector's backgrounds and historical subjectivity.

C.   OBJECTIVES

As our primary objective, this research study will provide a thorough and current review of the issues regarding deception detection and its applicable utility within the international military communities, Department of Defense, and other governmental agencies. Study 1 will seek to determine the role of specific question strategies and expertise (previous interview or interrogation training/experience) on deception detection accuracy by replicating Levine et al.'s (2014) research on diagnostic utility. Study 2 will seek to determine what subjective historical methods individuals use in deception detection by replicating Park et al.'s (2002) work on what information people use other than the verbal and nonverbal behaviors of the liar when determining deception.

D.   RESEARCH QUESTIONS

This thesis and associated studies are, in essence and design, a replication of Levine et al.'s (2014) and Park et al.'s (2002) studies to determine if both researchers' results are generalizable beyond college students and law enforcement and useful in identifying specific questioning strategies relevant to international military activities and coalition operations to include the fields of accession, recruiting, training and development, interrogation, associated fields of defector questioning, flag officer

accession questioning, and fraud detection in international military contexts. Specifically, this thesis aims to answer the following two research questions:

- Utilizing the diagnostic utility methods of content and contextual questioning presented in the Levine et al. (2014) study, are international military officers able to distinguish deception with greater accuracy than the previous meta-analysis mean of 54%?

- Utilizing the Park et al. (2002) method, what types of information do international military officers report using when detecting lies in the workplace?

E. SCOPE

This research project is to cover the examination of the abilities of international military officers (N = 41) to detect deception when presented with high-stakes interviews where some interviewees lied and others told the truth. Specifically, Study 1 examines the role of specific questioning strategies and the self-reported expertise of the participants (expert versus inexpert) in accuracy levels. Study 2 examines what, if any, additional subjective information participation is used in determining deception and, also, over what timeline veracity judgments occurs. This thesis merely looks to replicate both the Levine et al. (2014) and Park et al. (2002) studies and determine whether or not they are generalizable to a different population, namely international military officers. Also, the results of these studies might shed light on appropriate high-stakes deception settings such as defector questioning and detection of fraud in international military procurement and acquisition settings.

F. ASSUMPTIONS AND LIMITATIONS

(1) Assumptions

As explained in both the Levine (2014) and Bond (2006) works, the primary assumptions of this research include that recent findings are not flukes and are instead related to changes in the research and empirical findings of older limited works. Additionally, this research study assumes international military officers are a distinct and different population from what has previously been studied. This uniqueness is a result of training received by military personnel, the inherent cultural emphasis on truthfulness,

and risks associated with high-stakes deception. Lastly, another assumption is that Study 2 will replicate the work of Park et al. (2002), given that Lindsey, Dunbar, and Russell. (2012) is the only other existing replication of such results using a workforce sample.

(2)   Limitations

Previous deception studies assume participants will display a truth bias in that human beings have shown over time a belief in the innate honesty of others (Park et al., 2002). This explains the feelings of betrayal often exhibited by those who have experienced lies. For the purpose of this research, the authors assume the participants will operate with a truth bias. Truth bias, however, is not manipulated or measured to determine if that is a replicable finding in the specific population studied. This thesis will not address any psychological factors such as mood, temperament, or any other conditions that might affect respondents' participation. No financial constraints limited the authors, as data collection included only volunteer participation.

G.   ORGANIZATION

This thesis is organized into two separate but mutually reinforcing studies.

Chapters II–V comprise Study 1. Study 1 is the replication of Levine et al.'s (2014) study of diagnostic utility—or questioning effects—on the role of expertise on detection deception accuracy. Chapter II, specifically, is the literature review that introduces the history of deception detection research and related literature. Chapter III provides the Study 1 methodology, including information on participants, materials, and the procedure. Chapter IV presents the statistical results of Study 1, and Chapter V expands on those results through detailed discussion of the findings and their limitations.

Chapters VI–IX comprise Study 2. Study 2 is the replication of Park et al.'s (2002) study of how people really detect lies in their everyday lives. Chapter VI is the literature review that introduces the premise of Park et al.'s (2002) research on how people actually detect lies in their interactions. Chapter VII provides the Study 2 methodology, including information on participants and the procedure. Chapter VIII

presents the results of Study 2, and Chapter IX expands on those results through detailed discussion of the findings and their limitations.

Lastly, in Chapter X, the overall summary, conclusions, and recommendations for further deception detection research will be presented.

THIS PAGE INTENTIONALLY LEFT BLANK

II. STUDY 1: BACKGROUND HISTORY / LITERATURE REVIEW

Few things are as fundamentally human as the quest to accurately ascertain the veracity of one's intentions. The earliest codified documents are rife with tales of and punishments for lying, be it Indian Sanskrit Vedas or the Greek physician Erasistratus, humanity's ordeal with the nature of truth remains a constant (Trovillo, 1939).

A. HISTORICAL BACKGROUND OF DECEPTION DETECTION

The historical underpinnings of deception detection begin with the initial reliance on what Trovillo (1939) notes as superstition and the concept of the Ordeal.

> It is significant that, with few exceptions, the historical accounts of deception-detecting from the days of Christ, through the Middle Ages, are the history of the Ordeal. Superstition so swayed the minds of people that it was the rule for them to ask for the Ordeal to prove their innocence. The accuser was not looking, evidently, for suspicious clues in the face or actions of the individual, for apparently the psychology of deceit did not exist. Even the religions of Europe, as late as the 16th Century, taught that proof of innocence or guilt would be furnished from on High in a variety of mystical modes. People did not consider that proof lay within or on the surface of the suspect himself (Trovillo, 1939, p. 850).

The Ordeal method of deception detection is easily understood when examined in the context of the early witch trials and the Inquisition of the 14th and 15th centuries where individuals were forced to participate in torturous tests where the outcome of the event was the prime determinate of the presence of deception.

As religious fervor subsided and use of the scientific method began to flourish, so too did the study of emotional states and their bearing on deception detection (Trovillo, 1939). One of the earliest such works was conducted by Mosso, an Italian physiologist. Mosso's work focused on the effect of fear with regard to deception, specifically the fear one experiences in being detected. Mosso's measurement of blood flow as it circulates and pools in the body led to the first crude attempts to measure the physiological effects of deception (Trovillo, 1939).

What these two examples illustrate is the ever shifting framework through which deception detection has been viewed. Recent, cutting-edge research indicates that diagnostic utility—or the degree to which information is useful as prompted through strategic questioning—is key to understanding how humans detect deception. Prior theories, however, hinge upon the psychological and physiological states of the person lying and the resulting nonverbal cues that could be "read" to detect deception. The recent work on diagnostic utility questions the usefulness of these nonverbal cues that might or might not yield diagnostically useful information about whether somebody is lying. Such reliance on nonverbal cues has yielded detection accuracy rates that are not much different than chance (Bond & DePaulo, 2006). As such, the concept of diagnostic utility presented in this thesis is based on Levine et al.'s (2014) work and differs from the classical interpretation of diagnostic utility in that it includes the use of Park et al.'s (2002) additional reliance on subjective historical information. Specifically, Levine et al. (2014) indicate that diagnostic utility, in relation to deception detection, is the consideration of contextual message content including plausibility, correspondence with known facts, the correlation of the two, and these are what influence one's true ability to detect deception.

B.   REVIEW OF PREVIOUS DECEPTION DETECTION ACCURACY STUDIES

Upon review of the last four decades of deception detection research, a common theme emerges across all studies: A set of individuals is recruited as message initiators (liars or truth tellers), a separate group acts as the sample and is tasked to determine the veracity of the initiator's message with accuracy being calculated as the proportion of correct judgments made by the sample and based on nonverbal low-risk deception settings (Park et al., 2002). Furthermore, the meta-data accuracy rate presented by these studies consistently falls around 57% (Kraut, 1980) and always between 45% and 70% (e.g., Kalbfleisch, 1994; Miller & Stiff, 1993; Vrij, 2000). Over the past 40 years, including 208 studies as noted by Bond and DePaulo (2006), this belief has been near collectively held in the literature (e.g., Burgoon, Buller, Ebesu, & Rockwell, 1994;

DePaulo, Kirkendol, Tang, & O'Brien, 1988; Feeley, deTurck, & Young, 1995; Kalbfleisch, 1994; Millar & Millar, 1995; Stiff & Miller, 1986; Vrij, 1994).

What deviated from this trend was the work by Levine et al. (2014) that focused on the previously described intrinsic value of diagnostic utility. Two complementary approaches to deception detection involving diagnostically useful information comprise the current pertinent body of knowledge concerning the subject. The Strategic Use of Evidence (SUE) method hinges on an interviewer possessing a form of useful evidence pertinent to the line of questioning without the interviewee knowing, in hopes the subject will inadvertently make false statements as compared to the evidence (Clemens, Granhag, & Strömwall, 2013). Further analysis regarding the way in which the evidence is presented during the interview leads to what has been described earlier as the Content in Context (CiC) technique. Both methods set an arbitrary baseline with presumably useful obtained, or obtainable, knowledge that allows the interviewer to measure the variable feedback (Levine et al., 2014). Providing useful knowledge and background specifics are critical to both methods and at the heart of this current study.

C.   STUDY DESIGN

This study was a 3x2 mixed design with the three sets of questions presented in Table 4 as a repeated factor, the two levels of expertise (expert versus inexpert) as an independent groups variable, and detection accuracy as the dependent variable. Expertise was operationally defined using participants' answers to the following question: Have you ever conducted interviewing or interrogation as a regular part of your job? Participants who answered "No" were coded as inexpert; those who answered "Yes" were coded as experts. An additional question asked was, "Have you ever received formal interviewing or interrogation training?" Only 20% of participants had received such training; therefore, the sample size was too small to make meaningful comparisons in the current study using this question as a proxy for expertise. Participant answers (their truth/lie judgments) were scored for accuracy by adding the number of correct judgments and dividing by the total number of judgments.

11

Each participant watched and rated 12 videotaped interviews of different students denying cheating. The 12 interviews used in the current study were the same interviews that were used in Levine et al.'s sixth study. For each of the question sets, two lying cheaters' interview segments were shown. For each deceptive interview, a corresponding honest interview was selected by matching on sex, race, and approximate physical appearance. Thus, there were two honest non-cheaters and two lying cheaters, all of whom denied cheating, interviewed with each of the three question sets.

D.   RATIONALE FOR STUDYING DECEPTION DETECTION ACCURACY OF INTERNATIONAL MILITARIES

Initial thoughts on deception detection in the military are guided toward intelligence and counter-intelligence specialty fields. As previously noted, however, the humanistic quality of deception means detection practices are relevant in all settings of interpersonal interaction. As a government entity, the military must always maintain ethical practices and morally sound principles, which in turn facilitate a strong truth bias. It is very difficult for individuals interacting within the military system, where it is reasonably assumed that all participants are honest brokers, for those same military members to then exit the system and work with individuals whose motivations are self-serving and not bound by the same politico-social contract. Additionally, differences exist in national and social cultures throughout the world. Deception is thus easily found in varying forms in varying specialties. Although not all specialties or circumstances require deception detection training on the level of enemy combatant interrogation, deception detection is a useful tool in areas such as contracting fraud, military law enforcement, and varying degrees of leadership where high-stakes deception can occur and where a healthy skepticism and base of training may aide mission accomplishment.

Buller and Burgoon's (1996) research on Interpersonal Deception Theory for the U.S. Army Research Institute serves as the primary academic linkage between this work's analysis of deception detection theory and the military. His four-year examination inadvertently parallels the same issues as the research questions posited in this work. Buller and Burgoon (1996) focus on issues related specifically to intelligence gathering and explores defector deception detection and the application or influence of psycho-

cultural lenses on strategic posturing/positioning as a focus of further research. This is further addressed in Skidmore and Ortiz's (2014) parallel research study on domestic officers.

Broadhurst and Cheng's (2005) research on The Effects of First and Second Language on Lie Detection Ability demonstrate how language barriers affect an observer's ability to detect lies. They identified that observers were better able to detect liars speaking in a second language rather than their first language, but they found it more difficult to identify those telling the truth in a second language versus their first language (Broadhurst & Cheng, 2005). While our research only focuses on detection in participants' second languages, it should be noted that participants'' English as a second language aptitudes may cause differences between our results and the results of Skidmore and Ortiz's (2014) simultaneous deception detection accuracy research.

E.   RESEARCH QUESTIONS

Study 1 research is based upon the qualitative replication of the Levine, Blair, and Clare (2014) methodology. Given that utilizing the diagnostic utility methods of content and contextual questioning presented in the Levine et al. (2014) study, are international military officers able to distinguish deception with greater accuracy than the previous meta-analysis mean of 54% and, if so, to what degree? If the study holds true to the findings of Levine et al. (2014), then the authors should find a mean accuracy in excess of 54% and improvement most likely greater than 70%, with a corresponding minimal effect for expert judgments. Should the replication prove false, further research will be required to examine which portion of the study does not hold true in the chosen sample and to what amount modification will need to occur. The next chapter fully details the methodology used in Study 1.

THIS PAGE INTENTIONALLY LEFT BLANK

III. STUDY 1: METHODOLOGY

In order to replicate the Levine et al. (2014) sixth experiment in an appropriately clinical method, the authors first sought approval of the Naval Postgraduate School's Institutional Review Board for the Protection of Human Subjects (IRB). The authors completed the mandated IRB ethically based training and all additional reviews and subsequently received approval to initiate research. All research occurred on campus in specially designated rooms designed to best replicate the atmosphere of the initial study. Daily findings and survey materials were maintained under secure conditions, no personally identifiable information (PII) was collected and no leakage of demographics occurred.

A. PARTICIPANTS

Participants were 41 international military officers studying in resident programs at the Naval Postgraduate School. Tables 1, 2, and 3 summarize each participant's country, rank and branch of service, respectively. Participants' ages ranged from 25–45 ($M = 34.68$, $SD = 5.298$) and years of military service ranged from 3–25 ($M = 14.63$, $SD = 6.541$). There are a total of 17 countries that the participants volunteered from to participate in this study. Due to a low sample size and some countries and regions being represented by only one participant, a comparative analysis based on language, behavioral, and cultural norms was not able to be conducted. Also, 85.4% were male; 82.9% identified themselves as Caucasian/White; 9.8% were Asian/Pacific Islander; 7.3% were Hispanic/Latino(a) (see Table 3). In terms of expertise, 26.8% indicated they had received formal interviewing or interrogation training, and 29.3% reported that they had conducted interviewing or interrogation as a regular part of their jobs. All participants volunteered their participation and none received any form of compensation for their involvement in this research. Participation was anonymous with the only demographic information being gender, age, race, rank, years of military service, military branch, and level/use of interrogation training.

Participants By Country			
Country	Frequency	Percent	Cumulative Percent
Ukraine	5	12.2	12.2
Brazil	3	7.3	19.5
Greece	1	2.4	22
Australia	7	17.1	39
Norway	1	2.4	41.5
Columbia	1	2.4	43.9
Singapore	2	4.9	48.8
Japan	1	2.4	51.2
Turkey	10	24.4	75.6
Germany	2	4.9	80.5
Bulgaria	1	2.4	82.9
Sweden	1	2.4	85.4
Chile	1	2.4	87.8
Romania	1	2.4	90.2
Netherlands	1	2.4	92.7
Indonesia	1	2.4	95.1
El Salvador	1	2.4	97.6
Switzerland	1	2.4	100
Total	41	100	

Table 1.   Participants by Country

Rank	Frequency	Percent
O2	8	19.5
O3	11	26.8
O4	12	29.3
O5	8	19.5
O6	2	4.9
Total	41	100

Table 2.   Participant Military Rank

Race/Ethnicity	Frequency	Percent
Caucasian/White	34	82.9
Asian/Pacific Islander	4	9.8
Hispanic/Latino(a)	3	7.3
Total	41	100

Table 3.   Race/Ethnicity

B.   MATERIALS

Twelve videotaped interview segments—all used in the Levine et al. (2014) examination of diagnostic utility and expertise in their sixth study—were used as the experimental stimuli in the current study. Specifically, each segment showed a university student being interviewed about cheating in a research-related trivia game. All students in the 12 segments denied cheating, but half (6) actually cheated. Thus, the 12 video segments resulted in six honest students telling the truth (they denied cheating and they, in fact, did not cheat), and six dishonest students lying (they cheated, but they denied it in the interview). See Levine et al. (2014, especially pp. 270–272) for a detailed description of how the experimental videotapes were created. The 12 interview segments represented three different diagnostic utility questioning strategies (see Table 4).

First Question Set
1. Did you find the trivia questions difficult?
2. Was team work much of a factor? How so, please explain.
3. In looking at your score, you did better than other groups. Does this surprise you?
4. How would you explain your success?
5. Did cheating occur when I left the room?
6. Why should I believe you?
Second Question Set
1. How much experience have you had working in teams?
2. Have you played trivial pursuit games before?
3. How often do you play? [follow-up] Why do you/ don't you play them much?
4. In the current game, did you find the questions difficult? [follow-up] If you were going to scale it on a scale of 1 to 10, what would you say they were in terms of difficulty?
5. Was teamwork much of a factor in your performance?
6. What would you say if I told you that you did better than other groups?
7. How would you explain your performance?
8. Did you cheat when the experimenter left the room?
9. Are you telling me the truth?
10. What would your partner say if I asked them the same question?
Third Question Set
1. Tell me, in as much detail as you can, what happened during the trivia game?
2. How well did you do on in the trivia game?
3. Which questions did you and your partner get right?
4. For the answers you got right, explain how you knew the right answer?
5. In detail, what happened when the experimenter left the room?
6. Did any cheating occur?
7. When I interview your partner, what will they say about cheating?
8. Did you and your partner discuss cheating?
9. If someone did cheat, what should happen to them?

Table 4.   Questions Asked in the Three Sets of Cheating Tapes
(from Levine et al., 2014, p. 268)

18

C. PROCEDURE

Participants entered a lab setting at which point they read and completed a consent document. After consenting to participate, each respondent individually watched a series of 12 videotaped interview segments lasting approximately two minutes apiece. After each segment, the participant paused the video for as long as necessary to make a truth-lie judgment regarding the individual in the video (see Appendix for full questionnaire for Studies 1 and 2). Participants also answered a series of demographic questions.

D. DATA ANALYSIS APPROACH

Data was analyzed using mixed-model multivariate analysis of variance where question strategy was the repeated factor and expertise was the between factor, with accuracy as the dependent variable. Effect sizes are also reported.

The majority of our sample identified as Caucasian/White as shown in Figure 1. If we include unknown and international officers (Naval Postgraduate School, 2013), this is in line with Naval Postgraduate School ethnicity/race demographics for 2013. The small sample size for races other than Caucasian/White does not allow for statistically significant analysis to be conducted.
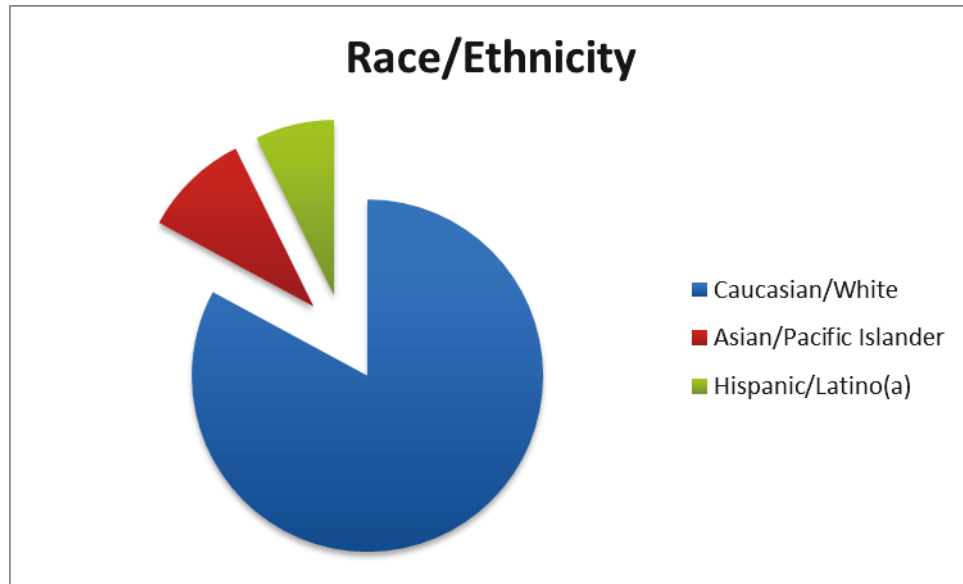
Race/Ethnicity

Race/Ethnicity	Frequency	Percent
Caucasian/White	34	82.9
Asian/Pacific Islander	4	9.8
Hispanic/Latino(a)	3	7.3
Total	41	100

Figure 1.   Race/Ethnicity of International Officers

As with race/ethnicity, the gender disposition is in keeping with the aforementioned demographics report. The increase also holds true when examining the NPS population (Naval Postgraduate School, 2013).

Gender of Participants

Gender	Frequency	Percent
Male	35	85.4
Female	6	14.6
Total	41	100

Figure 2.   Gender of Participants

The sample was comprised of a diverse range of military ranks centered predominately on mid-career officers, O–3 and O–4, which is in keeping with the overall demographics of the school (Naval Postgraduate School, 2013). Within the sample of international officers at NPS, we had a modest distribution of participants from various ranks and years of service (Figure 3)—participants with a mean age of 35 years old and a standard deviation of about five years. This can be because countries want to make sure they have commitment from their service members and are sending seasoned officers to NPS. With a somewhat large standard deviation, however, it also can mean that some foreign allies want to invest in their officers' development. This is further reinforced with the mean service of 15 years with a standard deviation of 6 ½ years. This would be a good comparative to the domestic study participants. This can have a severe impact on the biases that individuals might have with gain or little experience when it comes to deception detection.

21

Rank of Participants
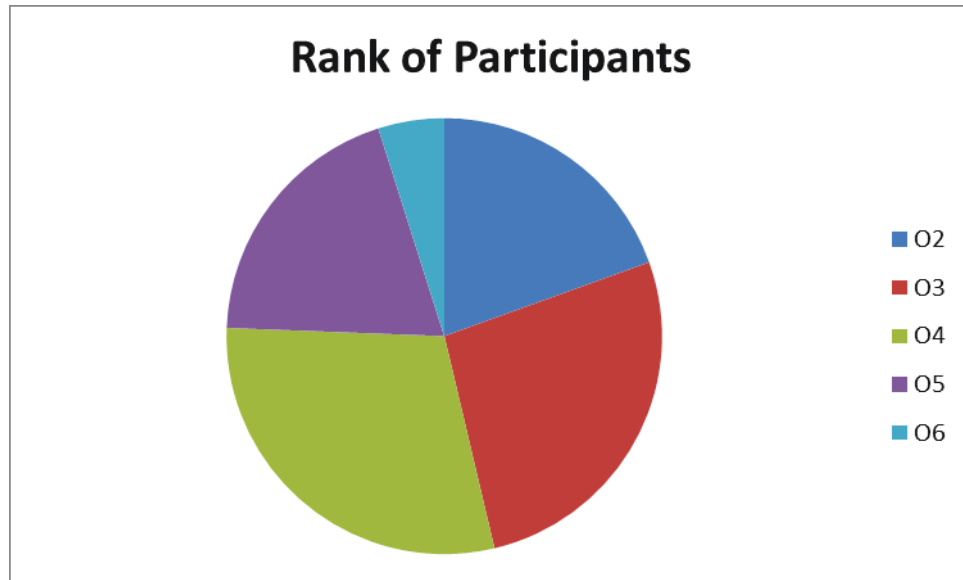
Rank	Frequency	Percent
O2	8	19.5
O3	11	26.8
O4	12	29.3
O5	8	19.5
O6	2	4.9
Total	41	100

	N	Mean	Std. Deviation	Minimum	Maximum
Age	41	34.68	5.298	25	45
Years of Service	41	14.63	6.541	3	25

Figure 3.   Rank of International Officer Participants

Among the sample we had a diverse mix of participants from various branches of service from international students. The majority of participants served in the Army as Figure 4 illustrates.

Branches of Services Among Participants
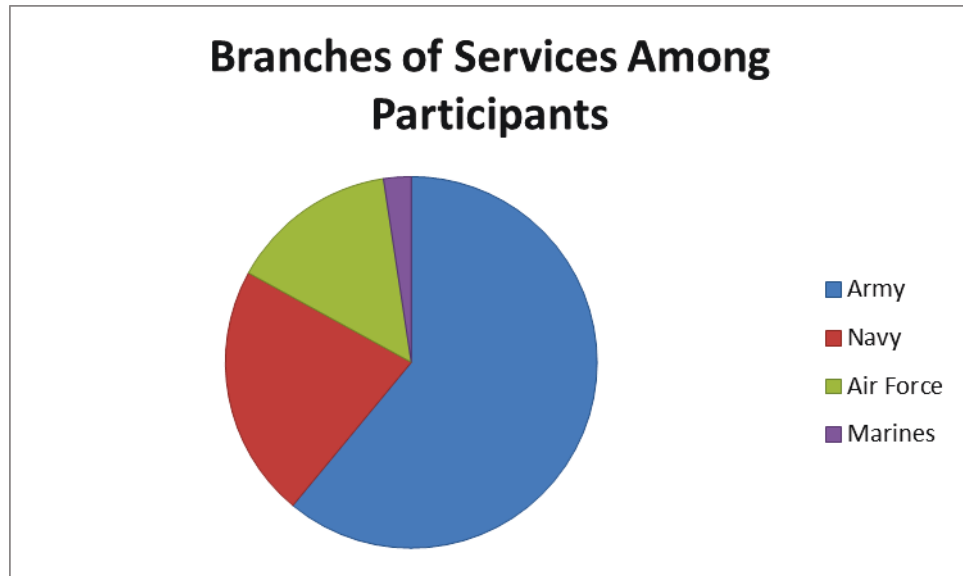
Branch	Frequency	Percent
Army	25	61
Navy	9	22
Air Force	6	14.6
Marines	1	2.4
Total	41	100

Figure 4.   Branches of Service Among International Officers

As shown in Figure 4, the sample contained a diverse mixture of military branches which loosely follows the represented population (Naval Postgraduate School, 2013).

THIS PAGE INTENTIONALLY LEFT BLANK

IV. STUDY 1: RESULTS

The data were analyzed with a 3x2 mixed Analysis of Variance with the three question sets as a repeated factor, the two levels of expertise (expert versus non-expert) as an independent groups variable, and detection accuracy (percent correct) as the dependent variable.

Consistent with Levine et al. (2014), this study replicated the strong main effect for questioning strategy, $F(2, 78) = 12.59$, $p < .001$, $\eta^2 = .24$, and a main effect for expertise, $F(1, 39) = 4.71$, $p < .05$, $\eta^2 = .01$. One should note that the main effect for expertise found in the current study is similar in magnitude to Levine et al.'s finding ($\eta^2 = .01$). Although Levine et al. found a statistically significant question type x expertise interaction, the current study did not (the effect size for Levine et al.'s interaction finding was small with $\eta^2 = .01$). Specifically, the question by expertise interaction in the current study was not statistically significant, $F(2, 78) = 0.44$, $p = 0.65$.

The cell means are presented in Table 5. Across experts and non-experts, accuracy was 33.7% (95% CI = ± 4.4%), 52.1% (95% CI = ± 5.3%), and 67.3% (95% CI = ± 3.8%) for question sets 1, 2, and 3, respectively. Accuracy means in the Question Set 1 and Quest Set 3 cells were significantly different from both 50–50 chance and the 54% meta-analysis mean at $p < .01$. Interestingly, accuracy means for Question Set 2 were exactly 50–50 chance for non-experts and consistent with the meta-analytic mean of 54% for experts.

Question Set	Set One	Set Two	Set Three
Expert	31.3% (18.8%)	54.2% (23.4%)	70.8% (23.4%)
Non-expert	36.2% (28.0%)	50.0% (33.4%)	63.8% (21.7%)

Table 5.   Mean Accuracy (and Standard Deviations) by Condition, Study 1

The data were further analyzed using a 3x2 mixed Analysis of Variance with the three question sets as a repeated factor, two levels of training (whether or not participants had ever received formal interviewing and interrogation training) as the independent groups variable, and detection accuracy (percent correct) as the dependent variable.

Consistent with this study's previous results, Levine et al.'s findings were replicated such that a strong main effect for questioning strategy, $F(2, 78) = 10.77$, $p < .001$, $\eta^2 = .22$, and a main effect for training, $F(1, 39) = 4.71$, $p < .05$, $\eta^2 = .11$, were found. The question by training interaction was not statistically significant, $F(2, 78) = 0.17$, $p = 0.85$.

The cell means are presented in Table 6. Across trained and untrained participants, accuracy was 32.4% (95% CI = ± 4.5%), 48.0% (95% CI = ± 5.3%), and 64.4% (95% CI = ± 3.9%) for question sets 1, 2, and 3, respectively. Accuracy means in all but one cell were significantly different from both 50–50 chance and the 54% meta-analysis mean at $p < .01$. Accuracy means for Question Set 2 with no formal training were consistent with the meta-analytic mean of 54% for experts.

Question Set	Set One	Set Two	Set Three
Formal Training	27.3% (23.6%)	40.9% (20.2%)	61.4% (28.2%)
No Formal Training	37.5% (26.1%)	55.0% (33.1%)	67.5% (19.9%)

Table 6.   Mean Accuracy (and Standard Deviations) by Condition, Study 1

The following chapter provides discussions, findings, limitations, and recommendations based on analysis.

V. STUDY 1: DISCUSSION, FINDINGS, LIMITATIONS AND RECOMMENDATIONS BASED ON ANALYSIS

The results of the study, as hypothesized, replicated the previous findings illustrated in Levine et al.'s (2014) research by increasing the mean accuracy by question set increase in utility of content and context. The importance of the findings, similar to that of Skidmore and Ortiz's (2014) simultaneous research conducted on domestic officers, is that participants demonstrated an increase in deception detection accuracy regardless to expertise disclosure as previously shown in Table 4.

A.   DISCUSSION AND FINDINGS

Historical deception detection methods utilizing verbal and nonverbal cues demonstrated meta-data analysis mean of 54% (Bond & DePaulo, 2006). The researchers' replication of content and contextual questioning methods demonstrated an increase in expert participant accuracy to 70.8% and non-expert participant to 63.8%. Even though the replication using international military officers illustrated statistical difference for non-experts, there was a higher mean accuracy utilizing content and contextual questioning compared to Levine et al. (2014) results. No statistical difference for experts compared to Levine et al. (2014) was identified for expert participants as illustrated in Table 7.

Question Set	Set One	Set Two	Set Three
Experts	29.2% (22.9%)	64.3% (24.1%)	72.4% (22.2%)
Student (Inexpert)	39.2% (24.1%)	66.6% (24.9%)	72.9% (22.7%)

Table 7.   Mean Accuracy (and Standard Deviation) by Condition
(from Levine et al., 2014)

Comparing international officers mean accuracy to the results of the simultaneous research of domestic officers by Skidmore and Ortiz (2014), we identify lower mean accuracy among international military officers as illustrated in Table 8.

Question Set	Set One	Set Two	Set Three
Expert	36.0% (23.2%)	66.2% (22.9%)	75.7% (22.6%)
Inexpert	39.0% (26.8%)	69.5% (23.3%)	81.0% (22.3%)

Table 8.   Mean Accuracy (and Standard Deviation) by Condition
(from Skidmore & Ortiz, 2014)

Although international military officers did not demonstrate the similar results as Levine et al. (2014) when it comes to levels of training, international military officers demonstrated a higher mean accuracy than historical mean accuracy of 54% (Bond & DePaulo, 2006).

Mean accuracy among O4 officers (Table 9) had the greatest increase in resemblance to Levine et al.'s 2014 study. The authors identified junior international officers and senior grade international officers had the widest dispersion of mean accuracy, possibly due to limited exposure to western culture and language aptitudes as discussed in Broadhurst and Cheng (2005). Question Set 3 mean accuracies utilizing content and contextual diagnostic utility questioning illustrates its applicability in deception detection benefits over historical meta-analysis across ranks.

Rank	Mean	Trimmed Mean (5%)	Standard Error	Lower Bound (95%)	Upper Bound (95%)
Question Set 1 Accuracy					
O2	.3125	.2917	.12275	.0222	.6028
O3	.4091	.4129	.06098	.2732	.5450
O4	.3542	.3380	.08404	.1692	.5391
O5	.3438	.3264	.06576	.1882	.4993
O6*	.1250	-	.12500	-	-
Question Set 2 Accuracy					
O2	.4688	.4653	.13724	.1442	.7933
O3	.5455	.5505	.10010	.3224	.7685
O4	.4167	.4074	.07107	.2602	.5731
O5	.6563	.6597	.09375	.4346	.8779
O6*	.5000	-	.2500	-	-
Question Set 3 Accuracy					
O2	.5938	.5903	.09375	.3721	.8154
O3	.6364	.6376	.06182	.4986	.7741
O4	.7708	.7731	.04825	.6646	.8770
O5	.5938	.5903	.09375	.3721	.8154
O6*	.6250	-	.12500	-	-

* Only (2) O6 Participants in this Sample

Table 9.   Rank Accuracy by Question Set

B.   LIMITATIONS

The sample in this research of international military officers did not take into account cultural/social differences that do exist based on regional/national and religious norms. Nor did the research take into account English language aptitude other than that the international military officers scored adequately to be accepted at NPS via the Test of English as a Foreign Language (TOEFL). Furthermore, the sample size (N=41) is fairly small considering the degrees of freedom (DOF). It was, however, directly representative of Naval Postgraduate School international officer population (2013 Factbook, Naval Postgraduate School, 2013). Lastly, the restriction of a single contextual nature of the lies examined in the experiment by international military officers is not representative of the high-stakes situations of the international military environment.

C. RECOMMENDATIONS BASED ON ANALYSIS

This methodology should be further utilized and replicated to gain a more robust understanding of the effects of content- and contextual-oriented questioning when it comes to international officers. Because of the limitation of the Naval Postgraduate School sample, we recommend replication of this research in Study 1 be conducted in other regional settings and further refined to incorporate local cultural and language norms. This research demonstrates that deception detection training be reviewed/redesigned to include content and contextual methodology techniques. The next chapter will provide the background and literature review for Study 2.

VI.   STUDY 2: BACKGROUND HISTORY / LITERATURE REVIEW

The focus of Study 2 is the qualitative factor of deception detection versus the quantitative issue of accuracy as in Study 1. Study 2 seeks to determine what factors, in addition to nonverbal cues and leakage, individuals use when making a truth/lie judgment. As Park et al. (2002) argue that there are four false assumptions that previous research has relied on in believing the primacy of verbal and nonverbal behaviors of the messenger—"questions researchers have asked, the research designs used to study deception detection, the directions the literature has taken, and the theories used to predict and explain the results" (p. 147)—Study 2 directly addresses the issues of the first and second assumptions by broadening the field of questions asked in relation to the underpinning reasons for the truth/lie judgment and by modifying the study to attempt replication of the Park et al. (2002) findings.

The only other study of workplace deception detection specifically examined the relational aspect of the power dynamic to that of deception detection (Lindsey, Dunbar, & Russell, 2011). The study consisted of a sample of 214 employed individuals of which 55% were in a management/supervisor status, of which ($n = 96$) stated they had engaged in deception in the workplace (Lindsey et al., 2011). Remarkably, "no lies were uncovered through the interpretation of nonverbal cues, rather they were discovered after the fact through evidence or confessions" (Lindsey et al., 2011, p. 74). This effect of after-the-fact evidence and complete disregard for nonverbal variables coupled with the workplace power dynamic lends great weight to attempting a replication of the study within the highly charged culture of the military, where both power and stakes are greatly increased.

Further support of investigating the qualitative rationale for veracity judgment is provided by Park et al., who argue in *How People Really Detect Lies* (2002) that the majority of previous studies relied too heavily on the nonverbal cues of interviewees in addition to factors such as:

(1)   Sources and judges in deception detection experiments were most often unacquainted and that detection accuracy might be higher if judges had relational or idiosyncratic knowledge of the message source.

(2)   Participants had only rarely been allowed to interact face-to-face.

(3)   Lies are often sanctioned (encouraged to a degree) by the researcher. This argument holds that liars telling sanctioned lies should be less aroused than those telling unsanctioned ones. Consequently, unsanctioned lies should be more easily detected than sanctioned lies, and detection accuracy might be better if more researchers studied unsanctioned lies.

(4)   Predominance of testing under conditions of everyday versus high-stakes lies.

Awareness of these limiting assumptions must be noted when conducting deception research. The risk associated with the factors listed can be mitigated through the use of content- and context-based questioning regimens.

B.   RESEARCH QUESTION

Utilizing the Park et al. (2002) method, what types of information do U.S. military officers report using when detecting lies in in the workplace? If the study holds true to the original as set forth by Park et al. (2002), the authors expect to find that the vast majority of individuals report using subjective historical information or after-the-fact data rather than nonverbal cues to detect deception. The following chapter discusses the methodology used for Study 2.

VII. STUDY 2: METHODOLOGY

Replication of Park et al.'s (2002) study occurred in an appropriately clinical manner, beginning with the authors' approval of the NPS Institutional Review Board for the Protection of Human Subjects (IRB). The authors completed the mandated IRB ethically based training and all additional reviews, and subsequently received approval to initiate research. All research occurred on campus in specially designated rooms designed to best replicate the atmosphere of the initial study. Daily findings and survey materials were maintained under secure conditions, and no leakage of personally identifiable information (PII) or demographics occurred. Upon completion of the study, the questionnaires were divided, and all qualitative information was independently coded by two coders.

A. PARTICIPANTS

Participation in Study 2 was fully voluntary and comprised the identical sample used in Study 1, international military officers (n=41) serving as students at the Naval Postgraduate School. Tables 1, 2 and 3 and Figures 1, 2, 3, and 4 discussed in Chapter III for Study 1 provide the full details and breakdown of the participants' demographics.

B. PROCEDURE

After completing the truth-lie judgments in Study 1, respondents were asked a series of open-ended questions following Park et al.'s (2002) protocol (see Appendix. Consent Form and Questionnaire). Specifically, participants were told to recall a recent work-related situation in which they discovered that someone lied to them, and they were asked to remember as much as they could about what happened.

Participants were asked to write a detailed description of the event:

(1) Recall as much as you can about the situation in which the person originally lied to you. In as much detail as possible, describe the event where you were lied to: Where did it happen? What was the lie about? If you can, be sure to write down the exact thing that the person said to you.

Next, respondents were asked a series of questions related to some of the details surrounding the situation:

(2)   How long ago did this event (the lie) originally take place?

(3)   What was/is the relationship between you and the person who lied to you?

(4)   Now, think about how you found out that the person lied to you. Describe in as much detail as you can the events surrounding your discovery of the lie: How exactly did you find out that the person lied to you?

2.   Coding of Qualitative Data

All questionnaires were collected from participants upon conclusion of their session and independently coded by two coders. The coding scheme was created by the authors based on N = 41 data collected from willing international military officers serving in resident student capacity at the Naval Postgraduate School. Participation was fully voluntary and anonymous and comprised the identical sample used in Study 1. Upon coding completion, the inter-coder reliability (Kappa) was calculated and any discrepancies were resolved via discussion between coders and the primary thesis advisor. This data are further reported in the results section. The qualitative codebook comprises the following questions and their associated Kappa. The question addressing "how long ago the lie was originally told" was coded using months as the unit of measure with a resulting Kappa of .99. The question of linking "relationship" included none specified, superior/immediate boss, superior/above immediate boss, subordinate, child, spouse, immediate family member (brother, sister, mom, dad), peer/friend, teacher/caregiver, senior in rank (but no command relationship), and other. Relationship coding resulted in a Kappa of .91. The "discover method" coding options included none listed, third party information, physical information, solicited direct confession, unsolicited direct confession, at-the-time verbal and/or nonverbal behavior, inconsistencies with prior knowledge, combination of two or more, and other. The "discovery method" Kappa was .84. The question addressing the lapse of time between when the lie was told and when the subject discovered the lie was coded as no answer provided, immediate detection, less than one hour, less than one day, less than one week,

less than one month, less than one year, and more than one year. The Kappa for time lapse was .72. The second coder transcribed the listing of discovery method examples as presented in Chapter IX. The following chapter will discuss the results of Study 2.

THIS PAGE INTENTIONALLY LEFT BLANK

VIII. STUDY 2 RESULTS

Study 2 respondents were asked to recall a work-related situation in which an individual lied to the participant; data was not collected or analyzed for 21 (51.2%) respondents who choose to recuse themselves of the question. Additionally, one respondent noted deception by his/her child and one respondent noted the deception as a teacher/student relationship. As the current study is interested in workplace-related deception, this interfamilial and teacher/student deception was eliminated from further analysis. Therefore, the sample for the relational research question is N=41 while the rest are N=39 to compensate for the removal of the non-workplace deception respondents.

For those who chose to answer, 18 (46.2%) of the initial sample, the most common discovery methods were Third Party Information (15.4%) and Physical Information (12.8%). The least common valid discovery methods were unsolicited direct confession (2.6%) and inconsistencies with prior knowledge (2.6%). Table 10 delineates these findings while Table 11 presents examples of the discovery methods.

DISCOVERY METHOD	f	%
None Listed	21	53.8%
Third Party Information	6	15.4%
Physical Information	5	12.8%
Unsolicited Direct Confession	1	2.6%
Verbal/Nonverbal Behavior	3	7.7%
Inconsistencies with Knowledge	1	2.6%
Combination	2	5.1%
Total	39	100%

Table 10.   Frequency of Recalled Workplace Lie Discovery Methods, Study 2

DISCOVERY METHOD	Response Example
Third Party Information	A colleague lied about getting married to a local woman while on an operation in the Balkans. His new spouse told me, in confidence, that they were married.
Physical Information	Going further into the investigation, the scanner in question was found at the indivduals residence.
Unsolicited Direct Confession	One sailor failed to show up for work because he said his mother was sick. After expressing verbally that I would help him get his mother help on several occasions, he finally told me that his mother was not sick and was late to work because he had not awakended after a long party.
Verbal/Nonverbal Behavior	They kept looking the other way, being very unsettled during the questioning.
Inconsistencies with Knowledge	I told a colleauge about my trip to San Diego. He told me he had went to San Diego two weeks before and went to Sea World. Two weeks later he asked me how long of a drive it was from San Diego? So I asked him how he traveled to San Diego, in which he said that he had not been to San Diego before.
Combination	A subordinate didn't come back from leave on time because he said he broke his hand. I checked his facebook page, talked to his friends, and spoke with his friends doctor which proved .

Table 11.   Examples of Discovery Method Categories

The second portion of the questionnaire regarded the amount of time, in months, that had passed since the respondent had been told the lie. The range provided was 0–180 months with the mean being 13.8735 and a standard deviation of 37.19455 as shown in Table 12.

	N	MINIMUM	MAXIMUM	MEAN	STD. DEVIATION
How Long Ago	39	0	180	13.8735	37.19455
Valid N	39				

Table 12.   "How Long Ago" Descriptive Statistics

Analysis was then conducted on the relationship of the participant and the liar as displayed in Table 13. Respondent disclosure analysis shows the primary relationship occurred between participants and subordinates (19.5%) with the least reported relationship between participants and superiors/immediate bosses (2.4%).

RELATIONSHIP	ƒ	%
None Listed	21	51.2%
Superior/Immediate Boss	2	4.9%
Subordinate	8	19.5%
Child	1	2.4%
Peer/Friends	6	14.6%
Teacher/Caregiver	1	2.4%
Senior (No Command Relationship)	1	2.4%
Other	1	2.4%
Total	41	100%

Table 13.   Frequencies of Participant and Liar Relationship

The final question related to the relationship between the respondent being told the lie and the subsequent realization and discovery of the lie that had been told. In 31 (79.5%) of the cases, participants either did not answer the question or the answer provided was too vague for objective discernment. Of the lies told, six (15.4%) were the product of immediate truth judgments. Of note, the possibility of bias arises in the qualitative results as the participants were cued toward applicable near-term, work-related lies. The remainder of the timing data can be found in Table 14.

LIE TIME LAPSE	ƒ	%
No Answer Provided	31	79.5%
Immediate Detection	6	15.4%
< One Day	1	2.6%
< One Week	1	2.6%
Total	39	100%

Table 14.   Frequencies of Lie Time Lapse

In the highly charged culture of the military, where lies carry significant consequences, it is expected there will be a high participant and subordinate discovery method response. Park et al. (2002) supports this with "…it can be argued that accuracy should be higher for high stakes lies because there should be more nonverbal leakage when the stakes are high" (p. 146). This is increased if consequences are shared by the respondent. Examinations of further distinguishing discovery methods are detailed in Chapter IX.

IX. STUDY 2: DISCUSSION, FINDINGS, IMPLICATIONS AND RECOMMENDATIONS BASED ON ANALYSIS

What is notable about this study's findings is that international military officers' third highest discovery method was verbal/nonverbal clues illustrated in Table 10. This coincides with the historically held idea that the deception detection basis is that of verbal and nonverbal cues, but contradicts Skidmore and Ortiz (2014) simultaneous research conducted on domestic officers.

A. DISCUSSION AND FINDINGS

As hypothesized, this discovery method results closely resembled those of Park et al. (2002) and Lindsey et al. (2011): Third party and physical information were the most prevalent discovery methods in our work, as well as simultaneous research conducted by Skidmore and Ortiz (2014) on domestic officers. Tables 15 and 16 are provided for comparison.

Discovery Method	ƒ	%
None Listed	-	-
Third Party Information	62	32.0%
Physical Information	35	18.0%
Solicited Direct Confession	7	3.6%
Unsolicited Direct Confession	16	8.2%
Verbal/Nonverbal Behavior	4	2.1%
Inconsistencies with Knowledge	4	2.1%
Inadvertent Confession	4	2.1%
Combination	60	30.9%
Other	2	1.0%
Total	194	100.0%

Table 15.   Frequencies of Recalled Lie Discovery Methods
(after Park et al., 2002)

Discovery Method	ƒ	%
None Listed	36	43.4%
Third Party Information	14	16.9%
Physical Information	12	14.5%
Solicited Direct Confession	2	2.4%
Unsolicited Direct Confession	4	4.8%
Verbal/Nonverbal Behavior	0	0.0%
Inconsistencies with Knowledge	5	6.0%
Combination	9	10.8%
Other	1	1.2%
Total	83	100.0%

Table 16.　　Frequencies of Recalled Lie Discovery Methods
(from Skidmore and Ortiz, 2014)

B.　LIMITATIONS

Study 2 participant response rate was only 46.2% of the total sample of international military officers, preventing the authors from making robust inferences about deception detection in the workplace. Individuals exhibited difficulty in expressing deception detection in the workplace due to English as second language aptitude deficiencies. Furthermore, there were unaccounted cultural normalcies that inhibited the ability of an unknown number of international military officers to openly express workplace situations that involve deception detection.

C.　RECOMMENDATIONS BASED ON ANALYSIS

Future Study 2 replication could be further improved by allowing participants the ability to describe deception detection in the workplace events utilizing their primary language. Nonetheless, the authors' analysis concluded that third party and physical information were the primary methods for detecting deception as previous research indicates. Therefore, organizations and institutions should continue to seek out third party and physical information to develop content and contextual questioning of suspected individuals. The next chapter discusses the summary, conclusions, and areas for further research.

X. SUMMARY, RECOMMENDATIONS FOR FURTHER RESEARCH AND CONCLUSIONS

This study successfully demonstrated Levine et al. (2014) replicability and strength of diagnostic utility usage over cue-based deception detection with a sample of diverse backgrounds of international military officers. Rendering that, future research can be focused and have widespread applicability. This study further illustrates the goal of improving consistency through both quantitative and qualitative reasoning is obtainable over historical methodologies and methods as Park et al. (2002) findings demonstrated.

A. SUMMARY

The purpose of this study was to replicate Levine et al. (2014) quantitative methodology and qualitative approach of Park et al. (2002) within a niche sample of international military officers at Naval Postgraduate School to gain insight into the transferability and gauge future modification and improvements to further content and contextual questioning approach to the greater body of knowledge of deception detection. For Study 1, quantitative analysis, strategically grouped questioning was employed, and deception detection accuracy was retrieved from participants' observations of recorded interviews. For Study 2, qualitative analysis, the focus was on participants' experiences with deception in the workplace, specifically asking descriptive information of where, what, why, when, and how individuals were able to recognize or identify deception. The unification and utilization of quantitative and qualitative approach for the enhancement of methodology within this study will garner a robust understanding of the applicability and inferences made possible by content and contextual questioning towards the deception detection body of knowledge. The depth and breadth enrichment of this research towards the greater body of knowledge of deception detection will be discussed in the sections to follow.

The results of Study 1, a quantitative replication methodology of the Levine et al.'s (2014), demonstrated that diagnostic utility and judgment expertise are replicable with international military officers' sample. The sample ($N = 41$) is somewhat small,

43

however, but it is highly represented of the demographic international officer population within Naval Postgraduate School. Given that the results produced increased accuracy in deception detection by the authors' study regardless of a participant's expertise level demonstrates the replicability and need to greater research content and contextual questioning methodology when it comes to deception detection. This is supported by the fact that non-expert participants' accuracy increased to 63.8% and experts participants accuracy increased to 70.8% overall without discerning language and/or cultural exposure and/or proficiency versus replicable meta-data analysis mean of 54% (Bond & DePaulo, 2006). Even though no concise breakaway within this sample is demonstrated and not a large enough sample was obtained to discern by region, nationality, language, or culture, it is highly conceivable that these characteristics play a role and need to be further researched. Nonetheless, with all the variability that exists, the authors are able to demonstrate the relevance and replicability without boundaries of the Levine et al.'s (2014) methodology in Study 1.

The results of Study 2, a qualitative approach developed by Park et al. (2002), within the international military officers illustrated similar findings to that of Skidmore and Ortiz (2014) domestic officers. In that third party and physical information were the primary independent methods for detecting deception in the workplace, it is interesting to note that within our sample verbal and nonverbal cues was significant as a discovery method, keeping in line with historical body of knowledge of deception detection. Second, similarity within the samples was that deception was recognized within immediate or involved relationships in the workplace as identified by participants. The knowledge gained by Study 2 helps create a more robust understanding of workplace interaction and dynamics. The authors' analysis of the sample data lends way to further incorporate organizational and morale ethos to garner smooth lines of communication by limiting deception noise by incorporating training and development of coalition and international work centers.

B. RECOMMENDATIONS FOR FURTHER RESEARCH

The authors encountered several limitations working with the current sample of international military officers at Naval Postgraduate School. The authors' parallel study with Skidmore and Ortiz (2014) rendered some analysis limited or infeasible at this time due to numerous issues, one of the biggest for both studies being that of self-selection. The authors recommend additional studies of international military officers be undertaken so that data can be merged for further analysis to possibly identify regional, national, and ethnicity/race cultural norms that inhibit or bias individuals accuracy in deception detection. Furthermore, the videos utilized from Levine et al.'s (2014) methodology incorporates a trivia high-stakes game of northeastern university students, which renders its own challenges for nonwestern or un-immersed culturally participants. A highly accepted recommendation is that language proficiency or adaptation to native language be addressed to exemplify content and contextual questioning strategy in deception detection. Before going forth with developing custom or specific modifications to methodology, data should be gathered on diverse backgrounds of participants' abilities to accurately detect deception with current tools to determine the strength of content and contextual questioning in lieu of bias and other barriers. Psycho-cultural effects are hard to gather or assess from the sample the authors obtained, which further hinders the ability to have a wide applicability or correlation analysis or inference to be conducted. Thus, being that, this data should be further expanded into a larger data pool for greater robustness. The authors' ability nonetheless to replicate the diagnostic utility method of questioning on the sample in the parallel study from the population available is example of the fact that this body of knowledge can be tailored for specific stakeholders, especially business and national security, military, and intelligence communities. The authors are in agreement with Skidmore and Ortiz (2014) that obsolete methods of nonverbal variables and questioning techniques need to be reevaluated by various institutions and organizations like the Association of Certified Fraud Examiners.

(1) Accession, Recruitment, and Training

Fraudulent enlistment/non-medically qualified attrition costs the DOD millions of dollars annually. According to usmilitary.com, referencing 2006 DOD recruitment

statistics, the average cost per recruit is nearly $50,000 through basic training across all services (About.com, 2006). A United States Government Accountability Office study from 1997, which is seemingly outdated but does provide relevant factual data, identified that in the fiscal year 1994, of the 176,400 newly accessed recruits, more than 7,000 were discharged within their first six months of service because they were found to be non-medically qualified due to preexisting medical conditions that they concealed from recruiters and medical staff during processing (USGAO, 1997, p. 33). Furthermore, the research identified that more than 3,500 accessions were discharged within their first six months of service because they falsely represented or concealed their eligibility for military service (USGAO, 1998, p. 4).

Applying 2006 recruitment costs to those discharged in fiscal year 1994 shows that the DOD spent approximately 525 million dollars on avoidable fixed and variable costs. Training service recruiters, military entrance processing station staff members, and medical screening teams in the content and contextual questioning methods when interviewing new accessions could greatly reduce the number of fraudulent enlistments and non-medically qualified attrites; also, it could potentially have significant savings for the DOD's manpower budget. Follow on to training and development is the possible need to reevaluate current methods, through a needs assessment, in order to identify gaps within the intelligence community, within the militaries of various nations on the applicability and utility to bridge the gaps of interrogations. More importantly, it will create a more effective and robust gathering, analysis, and application of data and information to gain strategic advantage over an adversary utilizing content and contextual questioning over the historic nonverbal or heuristic approaches. Looking back at the last decade of conflict with a non-state actor enemy and further identified deficiencies of interrogations, especially those of enhanced or alternative means, have proven not to be reliable; rather, detrimental to states that practice techniques that are outside the accepted global universal norms. The authors believe that with further research, various national and military organizations can adopt techniques that would increase their abilities to identify deception through new adopted core competency with diagnostic utility questioning methods.

(2)   Senior Officer Accession & Executive Training & Development

The authors identify a further advantage from content and contextual questioning: For senior leadership training and development, even more emphasis on possible enhancement in the screening of commanding officers and strategically appointed positions of executive leadership. The authors note the Naval Postgraduate School's advanced executive degrees, programs, and enrichment seminars can be greatly enhanced with the incorporation of identifying deception and diagnostic utility for the development and enrichment of risk assessment and decision making at the Center for Homeland Defense & Security (CHDS) and Center for Executive Excellence (CEE). Executive seminars like Navy Strategic Communication Workshop (SCW) and Strategic Planning for Execution: Assessment & Risk (SPEAR) offered by the CEE through the authors' researched work of diagnostic utility can further advance the courses ability to develop senior leadership to gain a greater tool chest of resources and knowledge to make informed and decisive decision making in everyday situations, especially accessions into and in strategic posts.

(3)   Fraud Detection in International Contexts

The authors' Study 2, along with the Report to the Nations on Occupational Fraud and Abuse by the Association of Certified Fraud Examiners (ACFE), identified that third party or tips, as the ACFE notes them to be, are the most common way of detected deception. The ACFE 2014 study demonstrated that 5% of revenues or budgets are lost to fraud within the 1,483 international cases analyzed. If we take the yearly budget, this amount is roughly a $150 billion cost to the U.S. each year. This is a significant cost that needs to curtail especially since those are savings that the U.S. can pass on to taxpayers or reinvest into numerous other programs. Content and Contextual questioning as identified by the authors' replication of Levine et al's (2014) study can be vital and a low cost method of identifying and assisting in international fraud, which can greatly affect reduction in costs that are passed down to consumers and stakeholders.

C. CONCLUSION

The authors' thesis demonstrates the replicability of Levine et al.'s (2014) methodology of content and contextual questioning being nondependent to an expertise method of recognizing veracity within the deception detection field, along with a qualitative approach which furthers the resilience of diagnostic utility methodology through understanding the workplace setting experienced by international military officers. The incorporation and further research as described by the authors would build a competitive strategic advantage for any stakeholder that can foster this diagnostic utility core competency over lackluster historic methods. The advent of content and contextual questioning within the international military and civilian communities and institutions furthers heuristic and holistic approaches in detecting deception to curb various inefficiencies that are caused within states and organizations. Complexities of an expanding global arena of communication and interaction, along with coalition building to combat ever dynamic insurgents worldwide, depends on the ability of belligerents to be ever adoptive and resilient. Thus, this thesis offers insight to furthering the ability of entities to better prepare, analyze, and detect deception through a demonstrated replication of a method that is not dependent on expertise and burdensome costs.

APPENDIX. CONSENT FORM AND QUESTIONNAIRE

Consent Form

You are invited to participate in a research study to measure your perceptions about others' communication. The purpose of the research is to better understand how people's perceptions of communication impact their judgments. Your participation should take about 30 minutes to complete.

You will be asked to watch short video clips, make judgments about each clip, and complete a survey about past experiences you have had with similar communication situations.

Your participation is voluntary. If you participate, you are free to skip any questions or stop participating at anytime without penalty. The alternative to participating in the research is to not participate.

Your responses are anonymous and will not be linked to your identity in any way. No personally-identifying information will be collected—the survey only asks for broad demographic information and no other identifiers from participants.

The anticipated benefit from this study is that the findings will contribute to a larger body of knowledge, and they will be used to inform coursework at NPS. You will not directly benefit from your participation in this research.

There are no known or anticipated risks associated with participation.

Results of the survey will be used responsibly and protected against release to unauthorized persons; however, there is a minor risk that data collected could be mismanaged. Only the researchers will have access to the data which will be stored on a password-protected computer.

If you have questions regarding the research, or if you experience any injury or discomfort, contact Dr. Lisa Lindsey, LLindsey@nps.edu, Principal-Investigator. If you have any questions regarding your rights as a research subject, please contact the Naval Postgraduate School IRB Chair, Dr. Larry Shattuck, 831.656.2473, lgshattu@nps.edu.

Statement of Consent. I have read the information provided above. I have been given the opportunity to ask questions and all the questions have been answered to my satisfaction. I agree to participate in this study. I understand that by agreeing to participate in this research and checking the box below, I do not waive any of my legal rights.

☐ I consent to participate in the research study.

☐ I do not consent to participate in the research study.

You will see three sets of videotaped interviews. The basic situation is always the same, but the interviewer, the person interviewed, and the questions are different.

Background: These clips are of interviews with college students who participated in a study about teamwork. Each subject had just played a trivia game with a partner for a cash prize. All participants were given an opportunity to cheat when the experimenter was called out of the room, and the answers were left in a folder within easy reach of the participants. Some participants cheated and others did not. All the people being interviewed on these tapes denied cheating.

Instructions: Watch each interview and decided if you think they cheated or not. For each interview, circle an answer indicating your opinion about whether you think that they were honest and did not cheat or that they really did cheat and are lying about not cheating.

Set 1   Video: Exline1_4clips (6 min.)

Number	Interview	Judgment (circle one)	
1	39	Honest non-cheater	Cheated and lying about it
2	45	Honest non-cheater	Cheated and lying about it
3	44	Honest non-cheater	Cheated and lying about it
4	54	Honest non-cheater	Cheated and lying about it

Set 2   Video: Exline2_4clips (11 min)

Number	Interview	Judgment (circle one)	
5	54	Honest non-cheater	Cheated and lying about it
6	57	Honest non-cheater	Cheated and lying about it
7	71	Honest non-cheater	Cheated and lying about it
8	72	Honest non-cheater	Cheated and lying about it

Set 3   Video: Exline4_4clips (12 min)

Number	Interview	Judgment (circle one)	
9	25	Honest non-cheater	Cheated and lying about it
10	18	Honest non-cheater	Cheated and lying about it
11	10	Honest non-cheater	Cheated and lying about it
12	12	Honest non-cheater	Cheated and lying about it

Finally, we'd like for you to recall a recent __work-related__ situation in which you discovered that someone lied to you. Please take a moment to think of an example and remember as much as you can about what happened. Keeping this situation in mind, please answer the following questions:

1. Recall as much as you can about the situation in which the person originally lied to you. In as much detail as possible, describe the event where you were lied to: Where did it happen? What was the lie about? If you can, be sure to write down the exact thing that the person said to you.

2. How long ago did this event (the lie) originally take place?

3. What was/is the relationship between you and the person who lied to you?

4. Now, think about how you found out that the person lied to you. Describe in as much detail as you can the events surrounding your discovery of the lie: How exactly did you find out that the person lied to you?

Please tell us about yourself (circle the correct answer or fill in the blank):

Sex: Male   Female   Age: _____   Years of Military Service: _____

Rank: _____ Branch (circle one): Army   Navy   Air Force   Marines

Have you ever received formal interviewing or interrogation training? No   Yes

Have you ever conducted interviewing or interrogation as a regular part of your job? No

Yes

Are you: U.S. Military   International Military (please specify country)

_____

Race/Ethnicity: Caucasian/White

                       African American/Black

                       Asian/Pacific Islander

                       Hispanic/Latino(a)

                       American Indian

                       Alaskan or Hawaiian Native

                       Other (please specify) _____

LIST OF REFERENCES

About.com. United States military basic training attrition rates: Discharge rates for Military Boot Camp. (2006). Retrieved from http://usmilitary.about.com/od/joiningthemilitary/l/blbasicattrit.htm

Association of Certified Fraud Examiners. (2014). *2014 Global fraud study*. Retrieved from http://www.acfe.com/rttn/images/cost-of-fraud-infographic.pdf

Bond, Jr., C., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Pscychology Review*, *10*, 214–234.

Broadhurst, R. G., & Cheng, K. H. W. (2005). The deception of detection: The effects of first and second language on lie detection ability. *Journal of Psychaitry, Psychology and Law, 12*(1), 107–118.

Buller, D. B., & Burgoon, J. K. (1996). Interpersonal deception therory. *Communication Theory*, *6*(3), 203–242.

Burgoon, J. K., Buller, D. B., Ebesu, A. S., & Rockwell, P. (1994). Interpersonal deception V: Accuracy in deception detection. *Communication Monographs*, *61*, 303–325.

DePaulo, B. M., Kirkendol, S. E., Tang, J., & O'Brien, T. P. (1988). The motivation impairment effect in communications of deception: Replications and extensions. *Journal of Nonverbal Behavior, 12*(3), 177–202.

Feeley, T. H., deTurck, M. A., & Young, M. J. (1995). Baseline familiarity in lie detection. *Communication Research Reports*, *12*(2), 160–169.

Kalbfleisch, P. J. (1994). The language of detecting deceit. *Journal of Language and Social Psychology*, *13*, 469–496.

Levine, T. R., Blair, J. P., & Clare, D. D. (2014). Diagnostic utility: Experimental demonstrations and replications of powerful question effects in high-stakes deception detection. *Human Communication Research*, 40(2), 262–289.

Lindsey, L. L. M., Dunbar, N. E., & Russell, J. C. (2011). Risky business or managed event? Perceptions of power and deception in the workplace. *Journal of Organizational Culture, Communication and Conflict, 15*(1), 55–79.

Millar, M., & Millar, K. (1995). Detection of deception in familiar and unfamiliar persons: The effects of information restriction. *Journal of Nonverbal Behavior, 19*(2), 69–84.

Miller, G. R., & Stiff, J. B. (1993). *Deceptive communication*, Newbury Park, CA: SAGE.

Naval Postgraduate School Center for Excellence Official Webpage. 01/15/2015. Retrieved from http://www.nps.edu/Academics/OtherPrograms/ExecEd.html

Park, H. S., Levine, T. R., McCornack, S. A., Morrison, K., & Ferrara, M. (2002). How people really detect lies. *Communication Monographs*, *69*, 144–157.

Skidmore, K., & Ortiz, P. (2014). Deception Detection Process and Accuracy: An Examination of How U.S. Military Officers Detect Deception in the Workplace. Naval Postgraduate School (2014)

Stiff, J. B., & Miller, G. R. (1986). "Come to think of it…:" Interrogative probes, deceptive communication, and deception detection. *Human Communication Research, 12*(3), 339–357.

Trovillo, P. V. (1939). History of lie detection. *Journal of Criminal Law and Criminology*, 30(1), 848–881.

United States General Accounting Office. (1997). *Military attrition: DOD could save billions by better screening enlisted personnel* (GAO/NSIAD-97-39). Washington, DC: Government Accountability Office.

United States General Accounting Office. (1998). *Military attrition: DOD needs to better understand reasons for separation and improve recruiting systems* (GAO/T-NSIAD-98-109). Washington, DC: Government Accountability Office.

Vrij, A. (1994). The impact of information and setting on detection of deception by police detectives. *Journal of Nonverbal Behavior. 18*(2), 117–136.

Vrij, A. (2000). *Detecting lies and deceit: The psychology of lying and the implications for professional practice*. West Sussex, UK: John Wiley and Sons.

INITIAL DISTRIBUTION LIST

1.   Defense Technical Information Center
      Ft. Belvoir, Virginia

2.   Dudley Knox Library
      Naval Postgraduate School
      Monterey, California