The Impact of Individual Differences on the Interviewing Success and Post-Training Performance of Investigative Interviewers

by

Davut Akca

A thesis submitted to
School of Graduate and Postdoctoral Studies in partial
fulfillment of the requirements for the degree of

Doctor of Philosophy in Forensic Psychology

Forensic Psychology/Faculty of Social Sciences and Humanities

University of Ontario Institute of Technology

Oshawa, Ontario, Canada

August 2019

© Davut Akca, 2019

THESIS EXAMINATION INFORMATION

Submitted by: **Davut Akca**

Doctor of Philosophy in Forensic Psychology

Thesis title: The Impact of Individual Differences on the Interviewing Success and Post-Training Performance of Investigative Interviewers

An oral defense of this thesis took place on August 6, 2019 in front of the following examining committee:

Examining Committee:

Chair of Examining Committee	Dr. Steven Downing
Research Supervisor	Dr. Joseph Eastwood
Examining Committee Member	Dr. Matthew Shane
Examining Committee Member	Dr. Christopher O'Connor
University Examiner	Dr. Brian Cutler
External Examiner	Dr. Kyle Scherr, Associate Professor, Central Michigan University

The above committee determined that the thesis is acceptable in form and content and that a satisfactory knowledge of the field covered by the thesis was demonstrated by the candidate during an oral examination. A signed copy of the Certificate of Approval is available from the School of Graduate and Postdoctoral Studies.

Abstract

In this thesis, whether and how personality characteristics affect the performance in investigative interviewing and the efficacy of training was examined in a three-step research design. In Study 1, the structure of a 50-item aptitudes scale, a modified and extended version of the Police Interviewing Competences Inventory (PICI), was assessed using a general population sample (N = 300), and a four-dimensional aptitudes scale was created. The four dimensions found were named as Humane (13 items), Communicative-Insisting (13 items), Self-controlled (9 items), and Careful-Tenacious (10 items). In Study 2, student participants (N = 154) completed the aptitudes and the Five Factor Model (FFM) scales, and then interviewed witnesses who watched a mock robbery crime video. Interviewer performance was assessed based on the amount of details they could elicit, the perception of the witness, and researcher ratings of behaviours and question usage. Three dimensions of the FFM were correlated with the success measures: Agreeableness with witness perception and appropriate questioning, Extraversion with researcher ratings and inappropriate questioning, and Openness with researcher ratings. Only the Communicative-Insisting dimension of the aptitudes scale predicted high researcher ratings. In Study 3, we used a policing student sample (N = 38) to investigate the impact of training on the interview performance and also to analyze how training effect interacts with personality measures when predicting the performance of participants. Overall, training increased the performance of participants in most of the success measures. The Humane dimension of the aptitudes scale and the Openness/Intellect dimension of the FFM predicted training efficacy. The post-interview performance of the participants was predicted by the Openness/Intellect, Agreeableness, Extraversion, and Neuroticism (negatively) dimensions of the Big Five and the Careful-Tenacious

and Communicative-Insisting dimensions of the aptitudes scale. Findings might help police departments to identify potential successful interviewers and develop new training policies.

Keywords: investigative interviewing, aptitudes, Big Five, interview performance, training efficacy.

AUTHOR'S DECLARATION

I hereby declare that this thesis consists of original work of which I have authored. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I authorize the University of Ontario Institute of Technology to lend this thesis to other institutions or individuals for the purpose of scholarly research. I further authorize University of Ontario Institute of Technology to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research. I understand that my thesis will be made electronically available to the public.

The research work in this thesis that was performed in compliance with the regulations of UOIT's Research Ethics Board under the REB files #14560, #14561, and #14954.

_____ Davut Akca

STATEMENT OF CONTRIBUTIONS

The studies described in Chapter 2 and Chapter 3 has been published in the *Police Practice and Research* with the co-authorship of my supervisor Dr. Joseph Eastwood and the manuscript is currently under review:

Akca, D. & Eastwood, J. (2019). The Impact of Individual Differences on Investigative Interviewing Performance: A Test of the Police Interviewing Competences Inventory and the Five Factor Model. *Police Practice and Research*.

The study described in Chapter 4 is currently under consideration for publication and will be submitted to a journal once the manuscript is finalized.

I performed the majority of the experimental work, data analysis, and writing of both of the manuscripts. Study 2 and Study 3 were conducted in the ALERT lab operated by Dr. Joseph Eastwood. The undergraduate Research Assistants working in this lab helped in coding the data in both studies and conducting the experiments in Study 3.

ACKNOWLEDGEMENTS

This doctoral thesis would not be completed without the supports and helps of the people around me. First, I would like to thank my dear parents, my wife Sema, and lovely kids Tarik and Leyla for their existence, support, patience, and understanding during this long journey. I have been very lucky to have a supervisor like Dr. Joseph Eastwood as he always helped me whenever I needed and guided me towards the right direction while I was struggling with the difficulties of this doctoral study. Thanks so much Joe for your guidance and support! I am also thankful to my thesis committee members, Dr. Brian Cutler, Dr. Matthew Shane, Dr. Christopher O'Connor, and Dr. Kyle Scherr, for their contributions and feedback on my dissertation. I would like to thank to my fellow lab members in the ALERT lab, Christina Connors, Mark Snow, and Quintan Crough, as well for their helps and collaboration in the research studies that we have conducted together. I appreciate the undergraduate research assistants, Samuel Musiala, Joey Rao, Rachel MacDonald, Amrutha George, Katrina-ray Villeneuve, Monika Szota, Robert O'donnell, Hamdi Jimale, and Amanda Whan, who helped me during the lab studies and I wish all of them a successful career in the rest of their lives. I would also like to express my appreciation to the Durham Regional Police Service for the financial assistance they have provided throughout my doctoral study and the faculty members of the Forensic Psychology Department, Dr. Leigh Harkins, Dr. Amy Leach, Dr. Lindsay Malloy, Dr. Kimberley Clow, Dr. Shannon Vettor, and Dr. Karla Emeno, for their academic mentorship and guidance. Finally, special thanks to my Master's thesis supervisor Dr. Nawal Ammar and committee member Dr. Carla Cesaroni from the Criminology department as they always helped and guided me in my future career plans.

"Every cloud has a silver lining."

Contents

Thesis Examination Information ii
Abstract iii
Author's Declaration v
Statement of Contributions vi
Acknowledgements vii
List of Tables x
List of Abbreviations and Symbols xi
Chapter 1: Introduction 1
1.1. Investigative Interviewing 1
1.2. Poor Interviewing Skills 2
1.3. Police personality and performance from a historical perspective 4
1.4. Does personality matter in investigative interviewing tasks? 7
1.5. Police Interview Competences Inventory (PICI) 9
1.6. Definitions and Related Theories 13
1.7. What makes a good interviewer? 17
1.8. The Current Thesis 20
Chapter 2: 22
Examining the structure of the aptitudes scale (Study 1) 22
2.1. Method 22
Participants 22
Materials 22
Procedure 24
Data Analysis 24
2.2. Results 24
2.3. Discussion 27
Chapter 3: 31
The relationship between personality and interview performance (Study 2) 31
3.1. Method 32
Participants 32
Materials 32
Procedure 34

3.2. Data Coding and Reliability 35
3.3. Data Analysis 38
3.4. Results 40
3.5. Discussion 41
Chapter 4: 50
Training Effect and Personality (Study 3) 50
4.1. Training Effect on Investigative Interview Performance 50
4.2. Study 3 55
4.3. Method 56
Participants 56
Materials 57
Procedure 58
4.4. Data Coding 60
Reliability and Inter-rater Agreement 61
4.5. Data Analysis 62
4.6. Results 64
4.7. Discussion 71
4.7.1. Training effect 71
4.7.2. Personality-Training Effect Relationship 73
4.8. Concluding thoughts 80
Chapter 5: General Discussion 83
Future research 90
References 93
Appendix A: Witness Perception Scale 105
Appendix B: Researcher Rating Scale 108
Appendix C: Aptitudes Scale (Adopted from DeFruyt et al., 2006) 113
Appendix D: Demographics Questionnaire 116

LIST OF TABLES

CHAPTER 1

Table 1. Distribution of competences to five PICI dimensions……p.10

CHAPTER 2

Table 2. Four-dimensional structure of the aptitudes scale…..…….....p.25

Table 3. Reliability analysis for the aptitude dimensions…..……...p.26

Table 4. Principal Components Analysis Factor Loadings……...p.26

CHAPTER 3

Table 5. Inter-rater reliability for researcher ratings…….p.38

Table 6. Descriptive statistics for dependent variables ……p.39

Table 7. Correlations between success measures……..p.39

Table 8. Correlations between personality measures and success measures……....p.40

Table 9. Correlations (Aptitude Dimensions & FFM Factors)…….p.41

Table 10. Multiple regression models to predict interview success measures…….p.42

CHAPTER 4

Table 11. Descriptive Statistics for the Dependent Variables…….....p.63

Table 12. Correlations Between the Dependent Variables……..……p.64

Table 13. Difference in the Success Measure Scores (Pre-Post Interview)…..……..…..p.65

Table 14. Correlations between Big Five Dimensions and Success Measures (pre and post)..p.66

Table 15. Correlations between Aptitudes and Success Measures (pre and post)……p.67

Table 16. Correlations between Aptitude Dimensions and Success Measures…….p.68

Table 17. Correlations between Big Five Dimensions and Success Measures…….p.69

Table 18. Correlations between 10 Big Five Aspects and Success Measures……..p.70

Table 19. Correlations between 10 Big Five Aspects and Post-Interview Success……..p.71

LIST OF ABBREVIATIONS AND SYMBOLS

CPI	California Psychological Inventory
BFAS	Big Five Aspects Scale
FFI	Five-Factor Inventory
FFM	Five Factor Model
GATB	General Aptitude Test Battery
IPI	Inwald Personality Inventory
IRI	Interpersonal Reactivity Index
MMPI	Minnesota Multiphasic Personality Inventory
OAP	Occupational Aptitude Patterns Theory
PICI	Police Interviewing Competences Inventory

Chapter 1: Introduction

Investigative interviewing is one of the most fundamental features of criminal investigations, and the ability to interview victims, witnesses, and suspects effectively is a key skill for law enforcement officials (Baldwin, 1992; Milne & Bull, 1999; Snook, Eastwood, Stinson, Tedeschini, & House, 2010). Real-world analyses of police interviewers from around the world, however, have observed consistently poor interviewing practices (Baldwin, 1993; Fisher, Ross, & Cahill, 2010; Milne and Bull 1999; Smets, 2009, Snook et al., 2010). The majority of recommendations for rectifying this issue to-date have focused on improving the skills of interviewers through evidence-based interview training. However, little attention has been paid to the relationship between the personality characteristics of interviewers and their effectiveness in conducting investigative interviews. In the current research, the aptitudes and personality traits which may affect the success in investigative interviewing, and how they interact with interview training were examined using an experimental research design.

1.1. Investigative Interviewing

Interviewing is defined as “a conversation with a purpose” (Hodgson, 1987, p2). In almost every aspect of their job, police officers talk to people to gather some kind of information. Especially when the necessary information cannot be obtained by other means, such as with historical crimes, interviewing with people becomes particularly crucial (Memon & Bull, 1999). There are a variety of situations in which the police talk to people, including roadside interactions, helping people on the streets, job interviews with prospective police applicants, suicide attempts, hostage negotiations, and controlling rallies (Schollum, 2005). Although all of these interactions are used to gather information from the interviewees for different purposes, only the information

gathering activities within an investigation process are considered as 'investigative interviewing' (Ord, Shaw & Green, 2008; Schollum, 2005).

In order to resolve criminal cases thoroughly and to ensure a due process in the following stages, the interviewing of witnesses, victims, and suspects needs to be accomplished appropriately (Maguire, 2003; Schollum, 2005). In that sense, Gudjonsson (2002) explains the main purpose of investigative interviewing as eliciting "reliable, accurate and legally admissible information" from suspects, victims, and witnesses to find the facts within the context of a crime investigation (as cited in Smets, 2009, p.18). The information gathered through investigative interviewing is estimated to constitute eighty percent of all evidence gathered throughout a typical crime investigation, while the "real and documentary evidence" (i.e. physical evidence such as objects and documents gathered to prove or disprove the guilt in an investigation) constitutes the remaining twenty percent of all evidence (Yeschke, 2003, p. 47). Therefore, the quality of interviewing directly affects the judicial quality of an investigation file, which is important for the decision-making processes in judiciary (Smets, 2009).

1.2. Poor Interviewing Skills

An investigative interview is a dynamic social interaction between the interviewer and the interviewee, and it is a challenging task because interviewees might not be willing to share information for different reasons (e.g., embarrassment, guilt; Griffiths, Milne, & Cherryman, 2011; Yeschke, 2003). To overcome the challenges in interviews and obtain the most extensive amount of accurate and reliable information from suspects, victims, and witnesses, interviewers require variety of specialised skills and aptitudes (Schollum, 2005). The need for a higher quality of investigative interviewing has been emphasized in the literature based primarily on the research

findings demonstrating the poor interviewing capacities of police officers and the use of coercive and accusatory tactics in interviewing (Baldwin, 1993; Fisher, Ross, & Cahill, 2010; Milne and Bull 1999; Smets, 2009; Snook et al., 2010).

In one large scale analysis of real-world police interviews, Baldwin (1993) analyzed four hundred video tapes and two hundred audio tapes of interviews conducted by the police officers in the United Kingdom. His overall evaluations about the quality of the interviews indicated that the police officers were poor interviewers. Baldwin concluded that during the interviews, police officers frequently had an assumption of guilt, asked leading questions, interrupted the interviewee, lacked confidence, failed to control the interview, and sometimes acted aggressively. These behaviours are considered as the characteristics of a poor interviewer and do not result in complete, accurate, and reliable information being obtained (Milne & Bull, 1999). In addition, such an interviewing style is associated with false confessions, miscarriages of justice and wrongful convictions (Meissner et al., 2014; St-Yves 2009).

The low quality of interviewing persists even with the cooperative interviewees such as victims and witnesses. Witnesses often play an important role in the resolution of criminal cases through their knowledge about the crime being investigated. Nevertheless, police rarely benefit from these sources of information sufficiently due to the lack of proficiency at interviewing (Fisher, Ross, & Cahill, 2010). Snook and Keating (2010) suggest poor interviewers make some basic interview mistakes that reduce the quality of interviewing such as interrupting the responses of the interviewees, not allowing them to give their free narratives, over-talking during the interview, and asking unproductive questions (i.e. inappropriate closed-ended questions, leading questions, multiple questions, forced-choice questions, and opinions or statements) rather than

productive questions (i.e. open-ended questions, probing questions, and appropriate closed-ended questions).

Fisher et al. (2010) reviewed the extant research on witness interviewing ability of police officers in various countries such as Germany (Berresheim & Weber, 2003), the U.S. (Schreiber & Fisher, 2005), Canada (Snook & Keating, 2010), Finland (Korkman et al., 2006), Sweden (Cederborg et al., 2000), and Norway (Fahsing & Rachlew, 2009). Based on this review, Fisher et al. (2010) concluded that the interviewers use "intuitive or natural" methods instead of evidence based interview techniques (p. 57). They usually ask predetermined and suggestive questions, control the flow of information, and talk more than the interviewee, and these methods negatively affect the amount and quality of the information gathered during the interviews.

Taken together, the research conducted to-date suggest that the majority of police interviewers do not follow best practices that would allow for the maximum amount of information to be generated from interviewees.

1.3. Police personality and performance from a historical perspective

Police personality assessment has been used to examine the relationship between the individual differences among police officers and their job performance in different formats such as performing evaluations in the recruitment process, "fitness-for-duty" evaluations, and for the mental support to the officers who have psychological troubles (Weiss & Inwald, 2010, p.5). The use of assessment tests in the prediction of police performance dates back to the early 20th century. Assessment tests used in the early 20th century to predict police personality aimed primarily to evaluate either the intelligence level or the mental health of the officers before recruiting them in the police departments (Weiss & Inwald, 2010). For instance, Terman (1917) used the Stanford-

Binet scale, which is one of the earliest intelligence scales, to predict the future performance of police officers. In later decades, two types of inventories have been used in testing police personality: psychopathology tests (e.g., Minnesota Multiphasic Personality Inventory [MMPI]) and normal personality tests (e.g., Five Factor Model [FFM]) (Aamodt, 2010). Built on these two types of tests, specified inventories for police officers were created in late 20th century. The first personality inventory that was specifically designed to be employed for the selection of public safety officers was created by Robin Inwald in 1979 (Inwald Personality Inventory-IPI) and it involves 25 personality scales (Weiss & Inwald, 2010). Most of the scales of the IPI are related to socially deviant attitudes such as alcohol abuse and family conflicts, and in that sense, the inventory aims primarily to eliminate unsuitable candidates rather than selecting those who are more likely to excel in policing (Sanders, 2008). In fact, the studies that examine the predictive validity of IPI found that the inventory is successful in predicting negative or problematic behaviours such as incidences of absence and disciplinary issues among public safety officers (Sanders, 2008; Shusman, Inwald, & Landa, 1984).

In the 1980s, police personality assessment inventories started to focus more on the personality factors than on psychopathology (Weiss & Inwald, 2010). Following the trend in most of other professions, personality tests that measure the personality traits of ordinary people in daily life started to be used more often in police personality assessment in the last few decades (Aamodt, 2010). The Five Factor Model, which was created by Costa and McCrae (1992), is one of the most prominent personality tests and classifies personality traits into five major groups: Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness to Experience. Based on the model, several observations and studies across the world have been conducted, and these studies have indicated that the model is comprehensive and applicable to a variety of contexts (Costa &

McCrae, 1992). The model provides a common framework for researchers to study and summarize the personal characteristics of people in the five basic dimensions which are mentioned above.

Research on general job performance has shown that certain sub-scales of the Five Factor Model can predict the success in certain tasks or skills (Aamodt, 2010). For instance, openness to experience predicts academy performance, measures of emotional stability predict disciplinary problems, and measures of conscientiousness predict supervisory ratings of performance (Murphy & Dzieweczynski, 2005). Barrick and Mount (1991) conducted a meta-analysis of 117 studies on the relationship between Big Five personality traits and performance measures (i.e., job proficiency, training proficiency, and personnel data) designed for different jobs such as managers, police officers, and salespersons. Their findings indicated that Conscientiousness predicted success in all performance criteria in all professional groups; Extraversion was correlated with all criteria for social interaction, manager, and sales jobs; and Openness to Experience and Extraversion predicted training proficiency in all professions in the analysis.

In line with the findings on general job performance, research has shown that personality traits and the performance of police officers are also significantly related (Forero, Gallardo-Pujol, Maydeu-Olivares, & Andrés-Pueyo, 2009; Goldberg, 1993; Tormini 1997). A meta-analysis of 78 studies examining the personality-performance relationship in law enforcement settings revealed that personality traits assessed based on Personality Inventory (MMPI), California Psychological Inventory (CPI), or Inwald Personality Inventory (IPI) are moderately but significantly related to job performance or training performance (Varela, Boccaccini, Scogin, Stump, & Caputo, 2004). The CPI was found to be the strongest predictor of police performance among the three inventories. Black (2000) tested the influence of Five Factor Model personality traits in the performance of 284 police recruits. Their performance was assessed based on seventeen different practical and

academic tasks and skills. The analysis revealed that the performance was positively correlated with Conscientiousness ($r = 0.27$, $p < .001$) and Extraversion ($r = 0.16$, $p < .001$) dimensions and negatively correlated with Neuroticism dimension ($r = -0.16$, $p<.001$) of the Five Factor Model. When these dimensions were incrementally entered in a multiple regression analysis with cognitive ability to predict performance, Conscientiousness dimension added significant incremental validity ($R = 0.42$).

1.4. Does personality matter in investigative interviewing tasks?

Despite the substantial literature on job performance in policing, however, the specific relationship between personality and interviewing performance of police officers has received little attention from psychologists. Sear and Stephenson (1997) examined how Five Factor Model personality dimensions are correlated with the interviewing performance of nineteen police officers in four different real-life interviews that they conducted. The performance of the officers were rated based on thirteen interviewing skills (i.e. engage, legal requirements, reason for interview, routines, first account, topic selection, probing, questioning, summarizing and linking, listening and conversation, challenging, resistance, and closure). The analysis showed that there were no significant relationship between personality factors and the interviewing performance scores. However, when the officers were divided into two groups based on their interviewing scores as high scorers (those who score between 94-148) and low scorers (those who score between 82-89), a significant difference between the two groups on the Openness personality factor was seen. Specifically, those who score higher in interviewing had a lower measure in Openness (mean 2.05) than the others (mean 2.90). They concluded that the results of the study might have been influenced by the fact that the participants were a highly selected and homogeneous group of police

officers who held “a cold, calculating, and dominating approach to others” (p.32). Despite these results, Sear and Stephenson (1997) insisted that research should continue to seek to identify “good interviewers, assess their characteristics, attitudes, and policing styles”, and emphasized the need for reliable measures of job-related personality traits in the field of investigative interviewing (p.33).

Ono, Sachau, Deal, Englert, and Taylor (2011) examined whether both training and job performance of 131 law-enforcement agents from the US Air Force Office of Special Investigation in investigative interviewing were associated with their personality characteristics such as cognitive ability, Big Five personality traits, and emotional intelligence. The performances of the officers were assessed by their advisors. They found that only the Conscientiousness dimension of the Big Five was moderately correlated with the performance of the participants in a 17-week training program. Thirty-eight of the participants were re-evaluated a year after the training. Their cognitive ability and emotional intelligence were positively correlated and the high scores in the Neuroticism dimension of the Big Five was negatively correlated with their job performance. Wachi et al. (2016) asked two hundred and seventy-one Japanese police officers who are experienced in interviewing serious criminals to complete the NEO Five-Factor Inventory (FFI), the Interpersonal Reactivity Index (IRI)), and a scale to assess their interviewing style. They examined the relationship between the personal characteristics of the officers and their interviewing styles and found a positive relationship between the “perspective taking” domain of the IRI and the active listening technique. They also found that those who scored higher on Agreeableness tended more to employ the rapport building technique.

Overall, there are limited research studies examining the impact of personality on investigative interviewing performance. The few studies outlined above mainly used Five Factor

Model as a personality measure and reached inconclusive findings on the relationship between its dimensions and interview performance.

1.5. Police Interview Competences Inventory (PICI)

DeFruyt, Bockstaele, Taris, & Van Hiel (2006) were the first researchers who attempted to create an inventory of interviewing competences of police officers and examine the underlying structure of these competences. The list of 66 competences that are likely to be related to interviewing success were created by the authors through a literature review and by asking the views of experienced police interviewers. In the study, two hundred and thirty experienced interviewers, who were attending a course, rated their own estimated performance in 20 different interviewing vignettes. Then, they evaluated their characteristics based on the 66 interviewing competences. When DeFruyt et al. (2006) ran a principal component analysis based on the self-rated competences to discover the latent structure of the inventory, they found five underlying dimensions of interviewing loaded by forty competences which explain 44.43% of the variance among the participants. Namely, the underlying dimensions were ‘Careful-tenacious’, ‘Controlled-non-reactive’, ‘Dominant-insisting’, ‘Communicative’ and ‘Benevolent’. Each dimension refers to a certain set of competences and describes a different type of interviewing approach. Table 1 shows how forty interviewing competences were distributed among the five PICI dimensions. The final product of the analysis was called Police Interview Competences Inventory (PICI) and consisted of forty competences of interviewing. This instrument was created to enable police officers to assess their own values, skills, personal features, and personality traits that are necessary to conduct effective investigative interviews (Smets, 2009). The second analysis run by DeFruyt et al (2006) was to examine the association between self-rated competences and self-estimated

performance of the police officers. The results demonstrated that each of the five dimensions were highly associated with the self-estimated performance of the police officers in interviewing, although there were some differences based on the type of interview (i.e. suspect, witness, or victim interview).

Table 1. Distribution of competences to five PICI dimensions (DeFruyt et al., 2006)

Careful-tenacious	Controlled-non-reactive	Dominant-insisting	Communicative	Benevolent
thorough zealous concentrated careful persistent driven industrious determined	quiet calm self-controlled able to handle pressure thick-skinned able to keep one's head cool able to put things in perspective patient	offensive having a tongue of one's own not being tongue-tied rigid talkative authoritative taking action assertive	having feeling communicative quick to understand' fluent in social contacts able to respond quickly and appropriately having good intuition able to observe him/herself persuasive	complaisant goodhearted being tenderly able to act gently having sympathy empathetic having the ability to calm other people understanding

The interviewing behaviours referred through the PICI dimensions were briefly depicted by DeFruyt et al (2006). Specifically, the 'Careful-tenacious' interviewers had a systematic and thorough approach when they were interviewing, and they examined and scanned all details step-by-step by focusing on details. Those who scored lower on this dimension were less systematic, fell short in planning their interview, could be easily distracted during the interview, and their work sometimes resulted in superficial and incomplete reports. The second dimension of the PICI, 'Controlled-non-reactive' described the abilities to maintain non-reactivity, handle pressure, and cope with stress during an interview. High scorers in this dimension could successfully control their emotions when they were interacting with interviewees even if they were provoked verbally

or physically. Those who have difficulties in controlling their emotions and were quickly provoked scored low in this dimension.

The third dimension of the PICI was labelled as 'Dominant-insisting' by DeFruyt et al (2006) and it referred to the coercive interviewing style. High scorers in this dimension acted dominantly during the interview, held a tense approach, used direct-confrontation, and interrupted the interviewee more frequently than low scorers. In contrast, low scorers enabled interviewees to give their account freely and did most of the talking, interrupted them less, and held a non-coercive approach. 'Communicative' interviewing behaviours generated the fourth dimension of the PICI. Interviewers who scored high on 'Communicative' dimension could easily contact with interviewees and treated them empathetically. Low scorers, on the other hand, could not easily take the perspectives of others and maintained a reserved approach in their interactions. The last component of the PICI instrument was labelled as 'Benevolent' and it referred to a non-prejudiced, altruistic, and benevolent interviewing approach. Although the holders of this approach were perceived as "naïve" by their colleagues, they could reach successful results by easily building rapport thanks to their soft attitudes towards the interviewees.

The relationship between the five dimensions of PICI and Five Factor Model (FFM) personality traits were also analysed by DeFruyt et al. (2006), and the results showed that all of the five interviewing dimensions were associated with the five personality dimensions. The main idea behind investigating the association between five factor model and the five PICI dimension, according to DeFruyt et al (2006), was that personal characteristics are the building blocks of interviewing competences and the results of their analysis verified this assumption. In that sense, the five personality dimensions (Neuroticism, Extraversion, Openness to Experience, Agreeableness and Conscientiousness) were considered the predictors of interviewing behaviors

that were unique to an interviewer. The analysis showed that all interviewing dimensions were associated at least one of the five personality factors. Namely, DeFruyt et al. (2006) found, Extraversion and Conscientiousness factors were positively correlated with the 'Careful-tenacious', 'Dominant-insisting', and 'Communicative' dimensions of the PICI. Neuroticism, on the other hand, was negatively correlated with these three dimensions. That is to say, the police officers who had higher scores on Extraversion and Conscientiousness and lower scores on Neuroticism might be more competent in interviewing. Agreeableness was positively correlated with 'Benevolent' dimension, and the combination of Agreeableness and Neuroticism predicted a 'Controlled-non-reactive' interviewing behavior. Those who scored lower on Agreeableness had a 'Dominant-insisting' approach, while higher scores on Openness indicated 'Controlled-non-reactive' and 'Communicative' interviewing dimensions (DeFruyt et al, 2006).

Although an important first step in the area, the PICI research of DeFruyt et al. (2006) had two important limitations that might have affected the findings. First, they measured only the self-estimated interviewing performance of police officers. Such a measurement could be misleading, because the self-ratings might not reflect the reality given the limitations of the self-report questionnaires. When people are asked to rate their own performance in a given task, they might be affected by a kind of self-serving bias, which refers to one's attribution of positive events to their own character and while blaming external factors for negative events (Cucina, Vasilopoulos, & Sehgal, 2005). Second, the interviewers rated their estimated performance based on the vignettes which were fictional cases previously designed by the researchers. Therefore, the officers did not actually perform the interviewing task, but imagined that they did so. This imaginary performance might not have sufficiently reflected the real interviewing skill levels of the officers as used in a real case.

1.6. Definitions and Related Theories

Hoekstra and Van Sluijs (2003) argue that competences are likely to be developed through training and personal experience, however, the actual change in job performance due to this development is limited with the dispositions of the individual. In other words, the development of the aptitudes that are related to interviewing success might be predictable based on the personal traits of people. Therefore, the differences in interviewer style and the reason why some police officers perform better in interviewing can be explained through the personality traits and aptitudes. Personality is defined as "a dynamic organization, inside the person, of psychological systems that create the person's characteristic patterns of behavior, thought and feelings" (Graver & Scheier, 1996, p. 5). Based on this definition, Smets (2009) argue that a person's behavior in an interview might vary according to his or her personality characteristics. That is to say, different interviewers might respond in various ways to similar interview situations and they are consistent in their behavior over time. Therefore, there are some intrinsic motives that drive interview behavior, shape how the interviewer thinks and behaves during the interview, and influence the way we perceive the interviewer.

Murphy and Dzieweczynski (2005) argue that personality tests such as Five Factor Model consist of broad personality traits such as conscientiousness and agreeableness, and these measures may not be sufficient to predict "the functional relations" between personality and performance as they can disregard the situational factors that affect the performance in specific tasks. The personality measurement tools that include narrow traits, rather than the broad ones, might be more successful in predicting specific sets of behaviours, because such traits rely on explicit description of personality that considers various factors such as the time, place, and the role of the individual

in a society or group (Barrick & Mount, 2005). Barrick and Mount (2005) suggested that global personality traits like the FFM factors are useful in theoretical explanations, however narrower trait constructs are needed to predict specific behaviors at work. In that sense, specific personality measurement tools needs to be developed based on job-related criteria which can be identified through examining the necessities of specific tasks like investigative interviewing.

Differential aptitude theory suggests that the success in a certain job and the capacity to acquire the training related to that job can be predicted by incorporating measures of several specific aptitudes (Schmidt & Hunter, 2004). Based on this theory, specific aptitude measures such as quantitative, verbal, or spatial ability tests can be designed to understand the expected ability demands of the job being studied and these tests can be used to predict the success in that job (Brown, Le, & Schmidt, 2006). Individual differences have been used to select and classify the personnel according to the needs of specific tasks since the Second World War. During the war, people were assigned to various divisions of the army based on their aptitudes. Since then, designated tests have been used to measure the individual differences among people for recruitment or assignment purposes in different professions (Graham, 1998). General Aptitude Test Battery (GATB), for instance, is one of the most popular tools used to identify specific aptitudes that are required to select the appropriate personnel out of the job market in the U.S. (U. S. Department of Labor, 1970). In Canada, seventy-five percent of the organizations use some kind of aptitude test before deciding which person fits the job (Cronshaw, 1991).

Graham (1998) examines the relationship between specific aptitudes and job performance through the term ‘congruence’, which was defined by Gati (1989) as “the size of the gap between the profile of an individual’s characteristics and that of his/her occupation or job” (p. 182). The narrower the gap, the more congruent a person for a job or specific task in a job, because each

professional occupation necessitates specialized ability patterns. Therefore, specific aptitudes, in addition to general cognitive ability, are needed to achieve the goals of a job successfully (Prediger, 1987). Graham (1998) gives some examples of aptitudes related to certain jobs such as higher numeric ability for engineers, higher clerical ability for secretaries, and higher manual dexterity for tire inspectors. In that sense, research should focus on the relationship between the aptitudes of professionals and their success in specific tasks of their jobs.

Gustafson and Mumford (1995) suggested that some professional occupations require a match between the personality and the necessities of the works to be accomplished. They invited researchers to identify the patterns of aptitudes which predict the person-job congruency. Job satisfaction might be an aspect of such a congruency. Dawis and Lowquist's (1993) Theory of Work Adjustment suggest that "satisfactoriness" is an important indicator of success and the job performance in the work environment. In addition, people are more likely to be satisfied with their job when their aptitudes are congruent with the requirements of the job.

According to Gottfredson's (1986) Occupational Aptitude Patterns (OAP) Theory, aptitudes form patterns and these patterns are related to the requirements of specific jobs. She reached this result by categorizing 460 jobs into 13 job clusters and applying the General Aptitude Test Battery (GATB) scale for each cluster. As a result, Gottfredson (1986) found that the aptitudes that predict success in specific jobs cluster and create patterns that differentiate some jobs from others. Gottfredson (1986) suggested that when the general cognitive ability is held constant, job performance can be better predicted through the possession of particular aptitudes, and measuring these aptitudes might help to distinguish the people who have required abilities for a job. By using Gottfredson's (1986) OAP theory, Graham (1998) identified aptitude scores for the highest job performance for each cluster of jobs and used these scores to predict job performance in

comparison to general mental ability. As a result, she found that specific aptitude information does not add a substantial amount of predictive validity when it was taken into account with general mental ability. Instead, the latter was found to be the main predictor of success in different clusters of jobs. This finding might be a result of a limitation of the study which is the usage of job groups as a whole to analyze the relationship between the aptitude scores and job performance. The problem here might be that Graham (1998) used Gottfredson's (1986) Occupational Aptitude Patterns map as a guide for grouping various jobs into clusters and analyzed the impact of aptitudes on the clusters as a whole.

The professional tasks today are increasingly specialized and even the subfields in a profession necessitates distinct personality traits and skills to accomplish the work. Therefore, Graham's (1998) findings about the limited impact of aptitudes on performance when compared with that of general mental ability might be challenged if specified aptitude measures can be created and applied for each job or task separately instead of considering the job clusters as unique. In fact, Hogan (2005) states that personality traits predict job performance almost as well as measures of cognitive ability. Moreover, Goldberg (1993) suggests, even "intellectually able individuals falter on the job when their personality traits are not congruent with task requirements" (p. 32). Based on their meta-analysis of 494 studies which examine the impact of FFM personality factors on job performance, Tett, Jackson, and Rothstein (1991) supported the usage of personality traits as a measure of job performance. However, they concluded, "full potential of personality measures in personnel selection will be realized only when confirmatory research strategies employing personality-oriented job analysis become the standard practice for determining which traits are relevant to predicting performance on a given job, and when greater attention is directed to the selection of psychometrically sound and construct valid personality measures" (p.732). In

that sense, focusing on specific tasks in a job when analyzing the personality-performance relationship and identifying the related aptitudes might be a valid strategy.

1.7. What makes a good interviewer?

Investigative interviewing is one of the tasks of police officers that require special abilities and aptitudes. In fact, it is difficult to draw the frame of qualifications of a successful investigative interview because of the complicated nature of interviewing (Smets, 2009). There are several aspects that might affect the outcomes of this task such as the time and place of the interview, type of crime, the position of the interviewee in the crime, and the characteristics of the interviewee and the interviewer. For example, the interviewees might not be willing to share information for different reasons and this brings an additional challenge to the interviewer in the efforts to reach the needed information (Yeschke, 2003). Furthermore, an interview is a dynamic process of interaction between the interviewer and the interviewee, thus, it is almost impossible to predict before the interview what will happen in the interview room. During this dynamic process, the interviewer needs to consider all possibilities and find instant solutions for unpredictable incidents. To ensure they can accomplish this, the interviewer needs to pay attention to the linguistic usage, the body language, his or her own values and standards, establish a good rapport with the interviewee, and adopt an empathetic attitude towards the interviewee (Smets, 2009).

According to Milne and Bull (1991) the purpose of a good interview is two-fold: a complete and accurate account of the interviewee should be elicited and positive perception of the police in the “eyes of the public” should be created (p.72). To obtain the most exclusive amount of accurate and reliable information from suspects, victims, and witnesses, the interviewer should be open-minded and have “an insatiable curiosity” (Schollum, 2005, p.17). Clark (2016) explains “keeping an open mind” in an investigative interviewing as taking all possibilities into consideration and

staying away from all presumptions about the guilt. Such an approach will make the interview and the resulting information lead the interviewer to the proper conclusion, and prevent the interviewer from leading the interview to the conclusion that he or she believe to be correct. Maintaining a positive perception of the police in the "eyes of the public" is the other important goal of a good interview according to Milne and Bull (1991). This can be ensured through humanitarian attitudes towards interviewee which has been found to be one of the most important characteristics of a successful interview. Pearse and Gudjonsson (1999) analyzed 18 criminal investigations in which the suspects first denied the allegations against them, but after the interview they confessed their guilt. They found that the interviewees who were treated in a humanitarian and respectful manner were more likely to admit their crime. On the contrary, dominant and confrontational approaches from the interviewer resulted in more denial.

Baldwin (1993) counts some standards of good interviewing as following: "allowing suspects an unhurried and uninterrupted opportunity to state their position; listening to their responses; avoiding hurrying, coercive, or authoritarian tactics; and testing a suspect's account with fairness and integrity" (p.328). To reach these standards a competent interviewer should have some communicative and social skills such as "a calm disposition and temperament, patience, subtlety, imagination, an ability to respond quickly and flexibly, and legal knowledge" according to Baldwin (1993, p. 328).

Bull and Cherryman (1996) created a list of the most important skills for a successful interview by asking experienced detectives and reviewing the literature. The skills that were rated as most necessary by police officers were listening, preparation, questioning, knowledge of subject, flexibility, open mindedness, rapport, and compassion empathy (Cherryman & Bull, 2001).

Walsh and Bull (2011) surveyed experienced investigative interviewers and asked their beliefs about good interviewers. Those who planned, built and maintained rapport, listened, remained open-minded, calm, unrushed, and well-organized in terms of questioning and of evidence/information presentation were defined as good interviewers by the participants. On the other hand, those who are inflexible, aggressive, unskilled planner, who miss the opportunities to appropriately question and challenge what interviewees says were deemed poor interviewer.

There is a consensus in the literature that rapport building and keeping a non-coercive and humanitarian style during interviews have crucial positive impacts on the interview outcomes. Studies that questioned interviewees' or interviewers' perceptions regarding the factors that affect the success of the interviews have found that interviewees are more satisfied and more willing to disclose information when they are treated respectfully, empathetically, and in a humanitarian way (Bull, 2013; Kebbell, Alison, Hurren, & Mazerolle, 2010; O'Connor, 2005).

The Home Office of the U.K. developed seven principles of investigative interviewing in 1992 and these principles have been adopted in many countries since then (Schollum, 2005). First of all, the Home Office identified the role of investigative interviewing as eliciting reliable and accurate information from suspects, witnesses, and victims to find the facts about a criminal case. To achieve this goal, the interviewer should have an open mind and pay attention to the relationship between the account given by the interviewee and the previous information in the case file. To act fairly in each interview is another principle that an investigative interviewer should take into consideration. The principles of the Home Office do not consider persistent questioning as unfair. It is clearly explained that the interviewer does not have to accept the first answer given by the interviewee even if he or she prefer to use the right to silence. The only restriction to ask questions

to the interviewee is brought in the interviews with the child victims of sexual or violent abuse. Finally, the interviewer must treat carefully to the vulnerable interviewees (Home Office, 1992). Taking all of these into consideration, interviewers should have a set of aptitudes to overcome the obstacles that might be confronted during this complicated task, reduce the number of mistakes mentioned above, and elicit as much complete and reliable information as possible during the interview (Schollum, 2005).

1.8. The Current Thesis

In the current thesis, the impact of the aptitudes[1] and personality characteristics on the performance in investigative interviewing was examined across three studies. In Study 1, a large general public sample was given a modified version of the PICI to test the validity of the instrument and a four-dimension scale was created using principal components analysis based on the survey results. In Study 2, participants conducted a mock witness interview and the correlation between the interview performance of the participants and their scores in the PICI and FFM scales was analyzed. The research design in this study helped us address the methodological limitations of the PICI study. Specifically, the performance of the interviewer in an interview with the witness of a mock robbery crime was assessed by the researcher and witness instead of the self-evaluation technique used in the study of DeFruyt et al. (2006). Moreover, our study examined the real performances of the interviewers within the context of a designed laboratory paradigm rather than the self-estimated performance based on fictitious vignettes. In Study 3, we used a policing student

[1] The term 'aptitude' is preferred in this study to refer the innate characteristics of the interviewers instead of 'competency' which was used by DeFruyt et al. (2006). Although these two concepts are sometimes used interchangeably in the literature, 'aptitude' is preferred more often to refer to innate abilities and the ability to learn, whereas 'competency' is frequently used in reference to the abilities that can be developed through education or training (Hoekstra & Van Sluijs, 2003; Kurke & Gettys, 1995; Tracey, 2016).

sample to investigate the relationship between training and interview performance and also to analyze how training effect interacts with personality measures when predicting the performance of participants. Through this new research design, five research questions were investigated: (1) Are there certain aptitudes and/or personality characteristics of people that predict their success in investigative interviewing? (2) Are the PICI inventory and its five dimensions still valid and reliable with the new research design used in this study? (3) Is there any relationship between these aptitudes and Five Factor Model personality traits? (4) How does a short training course affect the performance in investigative interviewing? (5) How does the training effect interacts with personality measures when predicting interview performance?

Chapter 2:

Examining the structure of the aptitudes scale (Study 1)

In Study 1, the structure of the aptitudes scale, a modified version of the Police Interviewing Competencies Inventory (PICI), was assessed using a general population sample. By adding ten new items and using a sample that has different cultural and demographic characteristics, the reliability and validity of the PICI was tested. The aim of this study was to examine the structure of the aptitudes scale before using it in Study 2 and Study 3 in which the relationship between the aptitudes and investigative interview performance and how the aptitudes affect interview training efficacy were analyzed.

2.1. Method

Participants

The participants for the online survey were gathered from Canada and the United States through the Qualtrics panels service, which is an online aggregator of research samples for survey-based studies. The final sample consisted of 300 people; 158 (53%) male participants and 142 (47%) female participants. The mean age of the participants was 46 (*SD* = 16.65, *Range* = 18 – 84). In terms of ethnicity, 219 (73%) of participants self-reported as Caucasian, 29 (10%) as Black or African American, 25 (8%) as Asian, 7 (2%) as Latin American, 14 (5%) as Other, and one participant as Middle Eastern.

Materials

A scale was created to examine the interviewing-related aptitudes to use in Study 2. The scale used in this study was an extended version of the PICI instrument created by DeFruyt et al.

(2006), and it consisted of fifty questions. Each question asked the participant to rate himself or herself on a different aptitude. Ten new aptitudes which were derived from the literature on the traits that are necessary to be successful interviewers were added on the forty items in the PICI list and a fifty-item scale was created. These items were added in an attempt to strengthen the existing PICI instrument and to see if there is any change in the variance explained by the instrument with these new items included. The new items added were: being honest, being sincere, being flexible, being confident, being open-minded, being polite, being even-tempered, having a firm tone of voice, being tactful, and being tenacious (Moston, Stephenson, & Williamson, 1992; O'Neill & Milne, 2014; Swanson, Chamelin, & Territo, 2002; Wicklander & Zulawski, 2003). On the first page of the survey, the participants were presented the consent form through which they were informed about the ethical approval given for the study, the types of questions included in the survey, the potential risks of participating for the participant and the ways to deal with these risks, the anonymity and confidentiality of the data collection and storage process, and the rights of participants. To increase the quality of the survey data, the participants were asked on the second page if they commit to providing their thoughtful and honest answers to the questions. The third page of the survey consists of demographics questions (i.e., gender, age, ethnicity, country, income level, and whether English is their first language). In the following pages, the aptitudes questionnaire was presented.

At the beginning of the aptitudes questionnaire, the participants were notified that they were going to be presented with 50 different aptitudes. They were asked to think of their behavior across a variety of situations in their daily lives and rate how characteristic their usual behavior is in each of the aptitudes on the scale. The scale ranged as following: very uncharacteristic, somewhat uncharacteristic, neither characteristic nor uncharacteristic, somewhat characteristic,

very characteristic. The responses for each question were scored from one to five respectively and these self-rated scores indicated to what extent those specific aptitudes characterize the participants.

Procedure

The participants were reached using Qualtrics' panels service. A sample of adults living in Canada or the United States were invited to complete the study via an invitation email from Qualtrics. The email contained a link where the survey was hosted. The participants clicked the link and it brought them to the questionnaire. After they read the consent form at the beginning of the questionnaire and approved to participate by clicking "continue", they started to fill the questionnaire.

Data Analysis

The aim of the Study 1 was to analyze the variety among the participants and examine the underlying structure and dimensions of the aptitudes scale. To this end, a principal component analysis and reliability test were conducted based on the responses of the 300 participants. Principal component analysis helped to group highly correlated aptitudes together and examine which combination of items best measure the components of the aptitudes scale. Orthogonal rotation was applied to create maximum distinctness among the factors as we did not expect any correlation among the factors.

2.2. Results

The result of Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy (.939) and Bartlett's Test of Sphericity (sig = .000) tests allowed us to conduct principal components analysis.

The cut off score for the factor loadings was determined as 0.4 and the items having a value greater than 0.4 were considered as loading on the related factor. Based on the eigenvalues greater than 1.0 and the analysis of the scree plot, the principal components analysis revealed a four-dimensional structure. These dimensions were named as Humane (13 items), Communicative-Insisting (13 items), Self-controlled (9 items), and Careful-Tenacious (10 items; see Table 2). Forty-five aptitudes in the scale explained 52.4 % of the variance among participants. Five items were removed from the scale because they cross-loaded in several dimensions or loaded relatively little. The removed items were Thick-skinned, Quiet, Rigid, Complaisant, and Tongue-tied. Among the 10 new items we added on the PICI list, polite, sincere, open-minded, and honest loaded under Humane dimension; having a firm tone of voice, tenacious, and confident loaded under Communicative-Insisting dimension; even-tempered and flexible loaded under Self-Controlled dimension; and tactful loaded under Careful-Tenacious dimension. The factor loading scores for the items in the scale is shown in Table 4.

Table 2. Four-dimensional structure of the aptitudes scale

Humane	Communicative-Insisting	Self-controlled	Careful-Tenacious
understanding	talkative	calm	aware of my own thoughts
sensitive	assertive	able to keep my	and actions
sympathy	authoritative	head cool	driven
empathetic	persuasive	self-controlled	thorough
tender	a firm tone of voice	even-tempered	work with relentless zeal
able to act gently	speak my mind	patient	careful
good-hearted	offensive	able to handle	able to put things in
polite	take action	pressure	perspective
sincere	tenacious	maintain	pay attention to details
open-minded	confident	concentration	persistent
able to calm others	fluent in social contexts	quick to	perseverance
honest	communicative	understand	tactful
good intuition	able to respond quickly and appropriately	flexible	

To test the internal consistency of the aptitude dimensions, a reliability test was conducted. Coranbach's alpha value for each dimension was higher than 0.87, which meant that all four dimensions were internally consistent (see Table 3). This allowed creating composite scores of these dimensions for each participant.

Table 3. Reliability analysis for the aptitude dimensions

	Number of Items	Cronbach's Alpha
Humane	13	.920
Communicative-Insisting	13	.887
Self-Controlled	9	.871
Careful-Tenacious	10	.879

Table 4. Principal Components Analysis Factor Loadings

Humane dimension	**Factor loading**
understanding	0.803
sensitive	0.789
sympathy	0.766
empathetic	0.682
tender	0.665
able to act gently	0.654
good-hearted	0.638
polite	0.605
sincere	0.596
open-minded	0.579
the ability to calm other people	0.509
honest	0.489
good intuition	0.421
Careful-Tenacious dimension	
aware of my own thoughts an actions	0.621
driven	0.612
thorough	0.605
work with relentless zeal	0.592
careful	0.579
able to put things in perspective	0.577
pay attention to details	0.577
persistent	0.575
perseverance	0.501

tactful	0.429
Communicative-Insisting dimension	
talkative	0.704
assertive	0.697
authoritative	0.695
persuasive	0.658
a firm tone of voice	0.644
speak my mind	0.623
offensive	0.598
take action	0.577
tenacious	0.544
confident	0.488
fluent in social contexts	0.482
communicative	0.457
able to respond quickly and appropriately	0.433
Self-controlled dimension	
calm	0.705
able to keep my head cool	0.699
self-controlled	0.695
even-tempered	0.678
patient	0.659
able to handle pressure	0.551
maintain concentration	0.536
quick to understand	0.492
flexible	0.439

2.3. Discussion

Our analysis revealed a four-dimensional structure for the aptitudes scale. Each dimension refers to a certain number of aptitudes and a specific approach to investigative interviewing. We named these dimensions according to the nature of the items that loaded on them: Humane, Communicative-Insisting, Self-Controlled, and Careful-Tenacious. This structure explained 52.4% of the variance among the participants. These results had some differences with those of DeFruyt et al. (2006). They found a five-dimensional structure for their PICI inventory and their scale explained 44.43% of the variance among the participants. The dimensions were named ‘Careful-tenacious’, ‘Controlled-non-reactive’, ‘Dominant-insisting’, ‘Communicative’ and

‘Benevolent’. Our scale explained higher amount of variance (8%) among participants than the PICI in terms of interviewing-related personality characteristics and suggested a four-dimensional structure rather than a five-dimensional one as the PICI did.

Those who score higher in the Humane dimension are expected to show more empathetic and humanistic behaviors during interviews and build rapport with interviewees more easily. They are also open-minded which might help them to have an unbiased attitude and elicit an objective account from interviewees. Being able to calm others is another advantage of this type of interviewers as they will be able to deescalate heated debates or situations during interviews and create a calm atmosphere that can help interviewees to remember more details about the criminal event. The Humane dimension of the aptitudes scale have similar items with the Benevolent dimension of DeFruyt et al. (2006)’s PICI which refers to eight competences. The additional aptitudes that the Humane dimension has are sensitive, sincere, polite, open-minded, and honest.

The Communicative-insisting dimension of the aptitudes scale has both positive and negative aspects in terms of a successful interview. Some aptitudes in this dimension such as being fluent in social contexts, confident, and communicative might help interviewers to engage interviewees more easily and keep rapport during the interview; while others such as being authoritative, offensive, talkative, and speaking one’s own mind might lead interviewees to keep quiet and to give less details about the incident as the interviewer talks more than him or her, or tend to confirm the interviewer’s version of the story. In the PICI study, however, similar characteristics loaded under two separate dimensions, Dominant-insisting and Communicative. The latter group of aptitudes in the Communicative-insisting dimension, which has negative effects on interviewing performance, had much more factor loading than the former group. Therefore, high scorers in the Communicative-insisting dimension might have a coercive interviewing style

as the high scorers in the Dominant-insisting dimension of the PICI. They might act dominantly during the interview, hold a tense approach, use direct-confrontation, and interrupt the interviewee more frequently than low scorers. In contrast, low scorers might enable interviewees to give their account freely and do most of the talking, interrupt them less, and hold a non-coercive approach.

The Self-controlled dimension of the aptitudes scale refers to a set of behaviors which help interviewers to control their reactions and emotions throughout interviews even if they are provoked verbally or physically by interviewees. They can cope with stressful interview atmospheres and maintain their patience during long interviews. This dimension has almost the same items with the Controlled-non-reactive dimension of the PICI. The fourth dimension of the aptitudes scale was Careful-Tenacious which involves the same items with the first dimension of the PICI that holds the same name. Being able to put things in perspective was under the Controlled-non-reactive factor of the PICI while it appeared under Careful-Tenacious dimension of our aptitudes scale, which makes more sense when we examine the other items in this dimension. Overall, Careful-Tenacious interviewers are expected to have a systematic and thorough approach when they are interviewing, and they examine and scan all details step-by-step by focusing on details (DeFruyt et al., 2006).

At the end of Study 1, we found a similar structure with DeFruyt et al. (2006), except that we found a four-dimensional model instead of five as two of the original factors in the PICI (i.e. Dominant-Insisting and Communicative) were collapsed into single dimension (i.e. Communicative-Insisting) in our analysis. Two items that loaded under the Dominant-Insisting dimension of the PICI, being rigid and not being tongue-tied, were removed from our list due to their low factor loading. Also, being offensive was the highest loading item in Dominant-Insisting dimension of the PICI, However, in our in our analysis being offensive came with lower loading.

Taken together, items like offensive, rigid, and not being tongue-tied with high factor loadings might have played a determining role in identifying the “Dominant” part in the name of this dimension of the PICI.

The difference in our findings may be attributed to the differences between the samples used in two studies. In DeFruyt et al’s (2006) study and the replication of Smets (2009), police officers completed the PICI scale whereas our aptitudes scale was completed by a general sample. Thomas (2011) argue that some personality traits observed in police officers including authoritarianism, cynicism, suspiciousness, prejudice are not innate but developed during the job as a result of constant exposure to calls for service, dealing with the public in worst case situations and other job-related stressors. This may explain why our sample identified themselves less dominant and the aptitudes related to that dimension loaded lower in our factor analysis. Another difference between the samples might be related to the cultural and ethnic variety. DeFruyt et al’s (2006) and Smets’ (2009) studies were conducted in Belgium while our data was collected from the participants living in Canada or the United States. Finally, we collected our data through an online survey platform, however, the PICI inventory was filled out by police officers during a training, which might have affected how carefully the surveys were completed by the participants.

Chapter 3:

The relationship between personality and interview performance (Study 2)

In Study 1, the underlying structure of the aptitudes scale was found and the 50-item scale was reduced to four aptitude dimensions to use in the next step. The purpose of Study 2 was to examine the relationship between the aptitude dimensions found in Study 1 and actual interview performance. In addition, the relationship between the Big Five personality factors and interview performance was analyzed in Study 2. This allowed us to see which scale is a better predictor of interview performance. The association between the aptitude dimension and the Big Five factors were also examined. We hypothesized that high scorers in the four aptitude dimensions (i.e. Humane, Communicative-Insisting, Self-Controlled, and Careful-Tenacious) would perform better than low scorers in interviews. More specifically, high scorers in these dimensions would be rated more positively by the interviewees and researchers than low scorers, and they would elicit more details during interviews than low scores do. One exception might be the questioning style of the Communicative-Insisting participants as they might ask more leading questions and talk more than the interviewee due to their insisting characteristics, which is considered inappropriate in investigative interviewing. We also hypothesized that those who score higher in Conscientiousness, Openness, and Agreeableness dimensions of the FFM would receive higher scores in all success measures, Extraverts and Agreeable interviewers would be rated more positively by the interviewees and researchers, while high scorers in Neuroticism would perform worse than low scorers. We expected correlations between aptitudes dimensions and FFM dimensions (i.e. Humane and Agreeableness, Careful-Tenacious and Conscientiousness, Communicative-Insisting and Extraversion, Self-Controlled and Neuroticism [negatively]).

3.1. Method

Participants

The sample consisted of 308 students enrolled in an introductory psychology course at a Southeastern Ontario University who were compensated with course credit upon their participation. The participants were randomly assigned either as an interviewer or as a witness. Among those who were assigned as interviewer, there were 70 (45.5%) male participants and 84 (54.5%) female participants. The mean age of the interviewers was 20 (*SD* = 2.93, *Range* = 17 – 42). In terms of ethnicity, 52 (33.8%) of the interviewers self-reported as Caucasian, 34 (22.1%) as Asian, 38 (24.7%) as from other ethnicities, 13 (8.4%) as Black, 13 (8.4) as Middle Eastern, 3 (1.9%) as Latin American, 1 (0.6) as Indigenous. The majority of the interviewers (74%) identified English as their first language.

The demographics of witness participants were similar to those of interviewers. Their mean age was 20 (*SD* = 2.93 *Range* = 18 – 38). Ninety-seven (63%) of the witnesses were female, while 57 (37%) of them were male participants. The ethnicity of the witnesses had a similar variation with those reported by interviewers: White – 48 (31.2%), Asian – 42 (27.3%), Other – 29 (18.8%), Black – 20 (13%), Middle Eastern – 9 (5.8%), Latin American – 5 (3.2%), Indigenous – 1 (0.6%). Four participants failed to follow directions and therefore were excluded from the study analyses.

Materials

Mock crime video

The target video used in this study consisted of a mock robbery crime which took place in a bowling alley. The video was 2 minutes in length, and showed the crime from the perspective of security cameras in the bowling alley. The crime is shown in the video from three different

angles—each angle showing different parts of the incident. During the two minutes, the robber who is a male in mid-twenties enters a bowling alley (the first camera angle), hangs around at the hall while pretending to look at the pamphlets (the second camera angle), approaches to the cashier, bangs on the table and shouts at the cashier to demand the money in the till, receives the money from the cashier (the third camera angle), and flees out of the bowling alley (the first camera angle again). While he is running through the hallway, one of the customers tries to stop him but fails to do so (the second camera angle). There are some other details shown in the video such as two customers entering the bowling alley at the beginning and starting to bowl, an elderly woman walking through the hall way before the robber comes in, and a few elderly people sitting on a table besides the cashier who were scared of the robber.

Demographics Questionnaire

In this questionnaire, the participants were asked about their age, gender, education, and ethnicity, the program they study, the year of the program they were in, whether English is their first language (if not, they were also asked about their level of proficiency in English) and, if they previously received any training on interviewing (see Appendix D).

Aptitudes Questionnaire

The fifty-item aptitudes scale, which is an extended version of the PICI, was completed by the participants (see Appendix C). They were asked to think of their behavior across a variety of situations in their daily lives and rate how characteristic their usual behavior is in each of the aptitudes on the scale. The scale ranged as following: very uncharacteristic, somewhat uncharacteristic, neither characteristic nor uncharacteristic, somewhat characteristic, very characteristic. The responses for each question were scored from one to five respectively and these self-rated scores indicated to what extent those specific aptitudes characterize the participants.

Five Factor Model Questionnaire

The FFM inventory was originally created by Costa and McCrae (1992). The version of the FFM that was used in this study was adopted from the Big Five Aspects Scale (BFAS) of DeYoung, Quilty, and Peterson (2007). BFAS consists of one hundred questions asking about some statements which were considered as the indicators of certain personality traits (e.g., I seldom feel blue, I get angry easily). The choices range were as following: strongly disagree, disagree, neither agree nor disagree, agree, strongly agree. As agreed upon in the literature, there are five dimensions of this scale: Neuroticism, Agreeableness, Agreeableness, Extraversion, Openness/Intellect.

Witness Perceptions Questionnaire

A questionnaire was created to assess the interviewees' perception of the performance of the interviewer (see Appendix A). Consisting of 44 questions, the questionnaire was created based on Duke's (2013) rapport scale for investigative interviews and the scale used in Villalba's (2014) study to measure the effect of rapport building in police interrogations (see Appendix A). Common questions existing in both scales were removed and four new questions were added that ask about the performance of the interviewer (e.g. the interviewer was successful in accessing all the information about the event that was in my memory). The responses were provided using a 5-point scale: strongly disagree, somewhat disagree, neither agree nor disagree, somewhat agree, and strongly agree.

Procedure

Participants were invited to the lab as pairs and they were assigned randomly to one of the two roles either as the interviewer or the witness. In each session, the witness was invited to watch the short video of a mock robbery crime on a wide screen with a headphone while the interviewer was waiting in a separate room. Before the interview, the participants in the interviewer role were informed that they would conduct an interview with the other participant about a robbery crime which he or she recently witnessed. They were also told that the main objective was to elicit as much information as possible from the witness during the interview. The key reminder to the interviewers was that there was not a certain procedure that they had to follow during the interview and they just needed to do the task – elicit information about the crime from the witness – as effectively as possible. After the instructions, both participants were asked to complete the questionnaires which are described in the materials section in detail above. After completing the questionnaires, participants were invited to the interview room and the interviews were videotaped. The videotapes were transcribed to use when coding and rating the data.

3.2. Data Coding and Reliability

Measures of Success

The success of the interviewer was assessed based on five different measures:

a) The perception of the interviewee: Previous research has shown that successful interviewers are rated highly by interviewees especially when they are shown empathy, understanding, and respect, and communicated in a non-threatening and non-judgmental way (Bull, 2013; Kebbell et al., 2006; O'Connor and Carson, 2005). The perception of the interviewee was measured in the current study through the interviewee perception questionnaire explained

above. The responses for each question were scored from one to five respectively and a total score on the forty-four questions was computed.

b) The amount of information elicited (number of details): The main purpose of investigative interviewing is to elicit as much information as possible from the interviewee regarding the crime being investigated (Schollum, 2005; Clarke et al., 2011). The amount of accurate information that the interviewer could elicit during the interview was counted based on a coding scheme and the total amount of accurate details was recorded for each interviewer. The maximum number of details that could be elicited was 288.

c) The number of details per question: An additional success measure was created by dividing the number of details gathered by interviewers by the number of questions asked. This gave us a measure of details elicited per question. In the evidence-based interviewing techniques, interviewers are encouraged to talk less than the interviewee, allow interviewees to talk without any interruption, ask fewer but open-ended questions, and avoid leading questions. Therefore, this measure showed us how successfully the participants utilized those techniques.

d) Researcher ratings: The score that was given by the researchers was calculated according to a scale that was created by Clarke et al. (2011) (see Appendix B). This scale includes certain tasks which the interviewer should accomplish to elicit the desired amount of information from the interviewees. Specifically, the scale inquires the rapport building skills, communication skills, types of questions asked (i.e. open-ended, probing, closed-ended), whether a free narrative was requested etc. Based on these measures, a total score for each interviewer was given.

e) Appropriateness of questioning: Whether the personality characteristics were associated with the appropriateness of questioning was also analyzed. Open-ended and closed questions were considered appropriate questions, and leading, multiple, long and complex

questions were considered as inappropriate questions. The appropriateness of the questioning was measured with a simple equation. The number of appropriate questions was divided by the total number of questions asked by the interviewer (appropriateness of the questioning = appropriate questions / [appropriate questions + inappropriate questions]).

Reliability and Inter-rater Agreement

Reliability and inter-rater agreement analyses were conducted for the success measures. The reliability analysis conducted for the 44 items of the interviewee perception questionnaire based on 154 interviewee responses showed that the scale is internally consistent ($\alpha = .83$). For the number of details, an independent research assistant was hired to act as the second coder. Each detail given by the interviewees were coded by two coders (doctoral researcher and the research assistant) based on the coding scheme. Before the coding, the research assistant was provided a one-hour training about the coding technique and how to fill the coding scheme. The inter-coder agreement analysis revealed a high agreement between the coders (Kappa= 0.73). Another independent research assistant was hired to assess the rating agreement of the items in the researcher rating scale. The research assistant rated the 20% of the sample, which was selected randomly among 154 interviews. Before starting to rate the interviews, the rater was provided a one-hour training session through which the instructions and materials that will be used in rating were described. To ensure that the second rater understood the rating procedure, a trial session was held to practice on five interviews and eliminate the potential confusions and misunderstandings. The inter-rater reliability for the items in the scale was analyzed through Weighted Kappa, which ranged from 0.35 to 1.00 for the items in the scale (see Table 5 for individual Kappa weightings for all items in the scale). Weighted Kappa is preferred when the rated items are ordinal variables as it allows assigning different weights to the different levels (Mandrekar, 2011). The overall

Weighted Kappa value was 0.64 (95% CI, .595 to .675), $p < .001$, which means a substantial level of agreement between the two raters was achieved (Landis & Koch, 1977). For the appropriateness of questioning index, there was also a substantial agreement between the two coders, Weighted Kappa = .63 (95% CI, .573 to .683), $p < .001$.

Table 5. Inter-rater reliability for researcher ratings

Item	Weighted Kappa
Introduce self	0.784
Explains interview purpose	0.517
Provides details of routine and route map	0.528
Evidence of Rapport Building	0.474
Encourages interviewee to give an uninterrupted account	0.399
Development of topics	0.629
Did the interviewer encounter with resistance	0.859
Appropriate Structure/ witness led sequencing	0.35
Appropriate use of questions	0.455
Keeps interview to relevant topics	0.573
Summaries and links	0.399
Covers points to prove	0.575
Were there any points to clarify in the witness's account?	0.745
Exploration of information	0.468
Use of pauses and silence	0.444
Self confidence	0.504
Open mind	0.304
Flexibility	0.471
Communication skills	0.411
Active listening	0.624
Invites interviewee to Add/Alter or Correct	0.332
Overall interview outcome	0.602

3.3. Data Analysis

To examine the relationship between the personality characteristics and the interview success, a series of Pearson's correlation tests were conducted. Before the analysis, the data were screened to test the outliers, normality, linearity, and the homogeneity of variance-covariance. The

data of six participants were removed as they were found as multivariate outliers based on the Mahalanobis distance calculation. Normality, linearity, and the homogeneity of variance-covariance of the data were assumed based on scatterplot and standardized residual tests.

Table 6. Descriptive statistics for dependent variables

	Minimum	Maximum	Mean	Std. Deviation
Witness Perception Score	3	77	46.74	17.898
Total number of details	0	99	31.19	15.14
Details per question	0	10	2.00	1.53
Researcher Rating Score	24	87	50.84	12.329
Appropriateness of Questioning	0.27	1	0.68	0.17

Table 7. Correlations between success measures

	Researcher rating	Witness perception	Total number of details	Details per question	Appropriateness of Questioning
Researcher rating	1	.190*	.609**	.049	-.165*
Witness Perception	.190*	1	0.131	-.054	-.183*
Number of details	.609**	0.131	1	.420**	-0.074
Details per question	.049	-.054	.420**	1	.228**
Appropriateness of Questioning	-.165*	-.183*	-0.074	.228**	1

*p<0.05

**p<0.01

After the data screening, the correlation between the aptitude dimensions and the success measures (i.e. number of details elicited, number of details per question, witness perception score, researcher rating score, and the appropriateness of questioning), and the correlations between five factor model personality traits and the success measures were analyzed. To test whether the FFM

and aptitudes scales predicted each of the dependent variables, multiple regressions analyses were conducted separately. Finally, the correlations between the aptitudes dimensions and five factor model dimensions were analyzed through Pearson's correlation test.

3.4. Results

The analysis revealed that two of the individual success measures showed significant correlation with only one of the aptitude dimensions. Those who score higher in Communicative-Insisting dimension were rated by the researchers significantly higher (r = .23) than low scorers. The Communicative-Insisting participants also scored lower on the appropriateness index than the others did (r = -.19) and elicited less details from the interviewee per question that they asked (r =-.171; see Table 8).

Table 8. Correlations between personality measures and success measures

	Witness Perception	Number of Details	Details per question	Researcher Ratings	Appropriateness of Questioning
Big Five Factors					
Neuroticism	.009	.006	.081	.016	-.055
Agreeableness	.200*	.067	.011	.021	.166*
Conscientiousness	.030	-.033	-.016	-.026	.003
Extraversion	.064	-.083	-.148	.211*	-.185*
Openness/Intellect	.051	.130	-.049	.188*	.032
Aptitudes					
Humane	.115	-.003	-.029	.018	.061
Communicative-Insisting	.009	-.113	-.171*	.229**	-.196*
Self-controlled	.109	-.059	-.159	.043	-.035
Careful-Tenacious	.079	.004	-.091	.133	-.060

*p<0.05
**p<0.01

The correlation analysis between the dimensions of FFM (Big Five) and interview success showed that the participants who scored higher on Agreeableness dimension were rated significantly higher (r = .20) by the witnesses (see Table 8). Those who scored higher on Extraversion (r = .21) and Openness (r = .19) were rated more positively by the researchers. The Agreeableness dimension was positively correlated with the appropriateness of questioning (r = .17), while extraverts scored significantly lower than the other participants (r = -.18) in the appropriateness index.

Each aptitude dimension was significantly related to at least one of the FFM dimensions (see Table 9) in parallel with DeFruyt et al. (2006). In accordance with our hypothesis, the highest correlations were found between Humane and Agreeableness dimensions, Communicative-Insisting and Extraversion dimensions, Self-Controlled and Neuroticism dimensions (negatively), and Careful-Tenacious and Conscientiousness dimensions.

Table 9. Correlations (Aptitude Dimensions & FFM Factors)

	N	A	C	E	O
Humane	-0.016	.699**	.303**	.352**	.350**
Communicative-Insisting	-.218**	-0.009	.444**	.752**	.327**
Self-controlled	-.508**	.301**	.420**	.328**	.452**
Careful-Tenacious	-.216**	.339**	.596**	.510**	.593**

N: neuroticism, E: extraversion; O: openness to experience, A: agreeableness, C: conscientiousness.
**p< 0.01

Multiple regression analyses were conducted to test whether the FFM and aptitudes scales predict each of the dependent variables. The only significant model found was when the Big Five dimensions were regressed on the appropriateness of questioning measure (see Table 10). The five

dimensions of the FFM explained 11.1% of the variance (R^2 = .111, F(5, 146) = 3.64, p<.01). It was found that Agreeableness (β = .23, p < .01) and Extraversion (β = -.35, p < .01) significantly predicted appropriate questioning.

Table 10. Multiple regression models to predict interview success measures

	Witness Perception		**Number of Details**		**Researcher Ratings**		**Appropriateness of Questioning**	
	β	**Sig.**	**β**	**Sig.**	**β**	**Sig.**	**β**	**Sig.**
Big-Five Dimensions								
Neuroticism	0.029	0.742	-0.005	0.953	0.070	0.421	-0.078	0.360
Agreeableness	0.213	0.016	0.063	0.468	-0.008	0.925	0.226	0.009
Conscientiousness	0.000	0.998	-0.063	0.523	-0.139	0.157	0.075	0.437
Extraversion	-0.003	0.975	-0.115	0.248	0.174	0.077	-0.354	0.000
Openness/Intellect	-0.011	0.905	0.204	0.035	0.168	0.080	0.053	0.567
R^2	0.044		0.041		0.059		0.111	
Adjusted R^2	0.011		0.009		0.028		0.08	
Aptitude Dimensions								
Humane	0.081	0.466	-0.014	0.898	-0.072	0.511	0.112	0.297
Communicative-Insisting	-0.063	0.547	-0.151	0.148	0.095	0.360	-0.250	0.014
Self-controlled	0.064	0.589	-0.074	0.531	-0.028	0.816	-0.020	0.859
Careful-Tenacious	-0.012	0.930	0.177	0.206	0.109	0.435	0.024	0.861
R^2	0.013		0.018		0.022		0.059	
Adjusted R^2	-0.013		-0.008		-0.004		0.033	

In further analyses, age, gender, and ethnicity of the interviewers were added in both models as control variables in a hierarchical multiple regression method. The new models did not significantly predict any of the success measures and control variables did not make a meaningful contribution to the predictive value of the models.

3.5. Discussion

In Study 2, we experimentally assessed the association between personality characteristics and interview performance by using a new experimental research design. The findings suggested that Big Five personality traits are better predictors of interview success than the aptitudes dimensions. The aptitudes scale used in this study was an extended version of the PICI scale created by DeFruyt et al (2006). Our findings suggested that the four-dimensional aptitudes scale, which was an extended version of the DeFruyt et al.'s (2006) PICI scale, explained a very small portion of the success in investigative interviewing. Specifically, only the Communicative-Insisting dimension of the scale was significantly correlated with three out of five success measures, which were researcher ratings, appropriateness of questioning (negatively), and details per question (negatively). The communicative characteristics of the high scorers in this dimension might have led the researchers rate them more positively, while their insisting sides might have diminished the appropriateness of their questioning style as we hypothesized and caused them elicit less details per question that they asked. On the other hand, Five Factor Model personality traits performed better than the aptitudes scale in predicting interview performance. Agreeableness, Extraversion, and Openness were associated with one or two of the success measures, while Neuroticism and Conscientiousness did not predict any of them. Despite this difference between the aptitude dimensions and Big Five dimensions, each aptitude dimension was significantly correlated with at least on of the Big Five dimensions. Interestingly, the number of details elicited was not associated with any of the aptitude dimensions or Big Five Dimensions. The multiple regression analyses indicated that the Big Five dimensions as a model accounted for more of the variance in all success measures as compared to the aptitudes dimensions. Adding the

age, gender, and ethnicity of the interviewers as control variables did not lead to a meaningful increase in the amount of variance accounted for by each of the personality scales.

In contrast to our findings, DeFruyt et al. (2006) found that some of the PICI dimensions are important to be successful in certain interview types based on the self-rated imaginary performance of interviewers. For example, those who scored higher in Careful-Tenacious considered themselves as more competent in suspect and witness interviews, while Controlled-Nonreactive interviewers viewed themselves more successful in victim interviews and Communicative participants scored themselves higher in all types of interviews. In addition, DeFruyt et al. (2006) found that being Dominant-insisting was important to be successful at interrogations and having a Benevolent approach was related to self-estimated success for interviewing victims and witnesses.

The difference between the findings of the current study and DeFruyt et al.'s (2006) may be a result of differences in research designs and the samples. First, the interview success was measured in the current study through four different objective tools (i.e. witness perception, researcher rating, number of details elicited in interview, appropriateness of questioning) instead of self-estimated performance which was used by DeFruyt et al. (2006). Secondly, in the current study, the interviewers conducted actual interviews with the witnesses of a mock robbery crime as opposed to the imaginary performance based on vignettes read by the participants in DeFruyt et al.'s (2006) research. It can be argued that watching a crime video cannot replace witnessing an actual crime scene. However, when compared with fictitious vignettes that were used in DeFruyt et al.'s (2006) study, a video footage of a crime can be considered more realistic and more likely to enable external validity. There is an agreement between this study and DeFruyt et al.'s (2006) in terms of high correlations between aptitude dimensions and Big Five traits. More research is

needed to conclude the discussion on whether the aptitudes scale or PICI inventory are valid predictors of interview performance.

Third, the irrelevance of the aptitude dimensions for predicting interview performance, which is contrary to DeFruyt et al.'s (2006) findings, might be a result of the lack of external validity of the sample used in our study. While DeFruyt et al. (2006) used a police sample, our study was conducted with undergraduate students from various programs. Future research might replicate our study with a police sample or analyze real-life interview videos of police officers to examine the impact of the aptitude dimensions on the interview performance. Fourth, the students may have had difficulty in understanding the wording of some of the aptitudes such as relentless zeal, perseverance, thick-skinned, rigid, complaisant, and tongue-tied. The culture and language difference between the samples in the two studies might also be a reason for inconsistency. That is to say, Canadian and Belgian participant might have understood some adjectives used in the scale in different ways. Therefore, more research is needed to conclude the discussion on whether the aptitudes scale or PICI inventory are valid predictors of interview performance.

An important finding of the current study was the correlations between three of the FFM dimensions and the interview success measures. High scorers in Agreeableness were found more successful by the interviewees and they performed better in appropriate questioning. Extraverts and high scorers in Openness were rated more positively by the researchers. Agreeableness, Openness/Intellect, and Extraversion personality dimensions might explain various requirements for being successful at investigative interviewing. In a study examining the impact of listening and verbal expression skills on interpersonal influence, Ames, Maisse, and Brockner (2012) found that Agreeableness and Openness mediates the relationship between listening skills and influence. They also found that Extraversion is a mediator in the association between verbal expression and

interpersonal influence. Listening skills and verbal expression are two important skills required for effective investigative interviewing (Bull, 2013; Russano & Narchet, 2013).

There are also other skills necessary to conduct effective interviews which might be related to Agreeableness, Openness, or Extraversion. Rapport building, for instance, is a crucial skill to perform better at interviewing (Duke, 2013; Milne & Bull, 1999; Villalba, 2014). Those who score higher in Agreeableness might have built better rapport with the witnesses in the current study because they were rated more positively by interviewees than low scorers in Agreeableness. As agreeable individuals are more courteous, flexible, trusting, good-natured, and tolerant (Ono, Sachau, Deal, Englert, & Taylor, 2011), they can engage with the interviewee more easily, establish a good rapport, and adopt an empathetic attitude towards the interviewee, which result in a more successful interview (Smets, 2009).

Openness is relevant to the performance in the tasks in which creative achievement and originality is needed (Judge et al., 2013). In an investigative interview, interviewees might not be willing to share information for different reasons and this brings an additional challenge to the interviewer in the efforts to reach the needed information (Yeschke, 2003). Furthermore, interview is a dynamic process of interaction between the interviewer and the interviewee, thus, it is almost impossible to predict before the interview what will happen in the interview room. During this dynamic process, the interviewer should consider all possibilities and find instant solutions for unpredictable incidents (Smets, 2009). High scorers in the Openness factor are known to be imaginative, cultured, curious, broadminded, and artistically sensitive (Barrick & Mount, 1991). Those who score higher in Openness may perform better in investigative interviewing by using these skills.

One surprising result was the irrelevance of Conscientiousness with any of the performance measures although it has been shown as the best predictor of job performance among the Big Five traits in general (Judge, Rodell, Klinger, Simon, & Crawford, 2013) including law enforcement professionals (Barrick, Mount, & Judge, 2003; Detrick & Chibnall, 2006; Ono et al., 2011). According to Costa, McCrae, and Dye (1991), Conscientiousness have both "proactive and inhibitive aspects" (p. 887). The proactive aspect includes traits like achievement striving, competence, and self-discipline (DeYoung et al., 2007), which might make high scorers on Conscientiousness successful in most professional tasks. The inhibitive aspect, on the other hand, includes traits like perfectionism and cautiousness. These traits might prevent interviewers from engaging with interviewees, building rapport, and asking creative questions during interviews. Therefore, these traits might have played a dominant role in some of the high scorers on Conscientiousness in our sample and led them score lower in the performance measures.

To our knowledge, the only attempt in the literature that assessed the relationship between FFM dimensions and interview performance before the current study found no significant relationship between personality factors and the interviewing scores based on the thirteen interview-related skills (Sear & Stephenson, 1997). In contrary to their finding, we found moderate correlation between three FFM dimensions and interview performance. The difference in the findings can be explained through the homogenous characteristics and the small size of the police officer sample used by Sear and Stephenson (1997) as they mentioned this aspect as a limitation of their study. Our study, however, used a demographically more heterogeneous and significantly larger sample than the one used by Sear and Stephenson (1997).

Overall, Study 2 aimed to examine if there are certain personality characteristics that affect the success in investigative interviewing by using a new experimental research design. Those who

scored higher in the Agreeableness, Openness/Intellect, and Extraversion dimensions of the FFM scale and high scorers in the Communicative-Insisting dimension of the aptitudes scale were found more successful in the interview task in our research design than the low scorers in these dimensions.

There are some limitations of our study that can be addressed in the future research. First, the performance of the participants in a laboratory-based interview might not have represented the whole picture of their skills and personality. Therefore, an analysis of the real-life interview videotapes might have more external validity. Second, using self-rated personality traits and aptitudes might suffer from self-serving bias. However, by using a large enough sample size in this study, such a self-serving bias was diminished (Cucina, Vasilopoulos, & Sehgal, 2005). Third, the variation in the ability to recall across 154 witness participants might have affected the interview outcome in this study. However, having the same witness participate in all session would have been meaningless as their memory would change in every session that they are interviewed. Fourth, the second rater did not rate all interviews due to time and resource restrictions but rated 20% of them which were randomly selected. Fifth, the fact that the personality questionnaires were filled by the participants jest before they conducted interviews might have affected the demand characteristics of the study. The participants might have guessed the research questions that we were investigating and performed accordingly. Finally, the participants were not actual police officers or investigative interviewers. The usage of student participants as investigative interviewers might be considered as a concern in terms of external validity of the research design, as they do not have any experience and they had not received any training on investigative interviewing. However, the primary goal of this study was to understand the specific unique impact of personality characteristics on investigative behaviours and outcomes. Therefore, using non-

expert participants allowed us to evaluate the pure impact of personality on interviewing behaviours independent of training and real-world experience. In addition, the need for a larger sample in research studies that include personality-related measures (Tett et al., 2009) made our sample preferable and convenient to gather in a limited time. The same research design can be replicated with an externally valid sample in future research.

The key contribution of Study 2 to the literature on investigative interviewing is a better understanding of the impact of personality on the interviewing performance in crime investigations through an objective, internally valid, and realistic research design. Police departments can consider the findings of this study when they are assigning interviewers in criminal investigations. Specifically, FFM inventory can be used when assigning interviewers. The officers who score high in Openness, Agreeableness, and Extraversion can be selected as investigative interviewers. More research is needed to understand the nature of the correlation between personality and interview performance, especially through the analysis of real-life police interviews.

Chapter 4:

Examining the Training Effect and the Interaction between Personality and Training (Study 3)

In Study 3, we investigated how a short training course on the PEACE model of investigative interviewing[2] affected interview performance and how this training effect interacted with personality measures when predicting interview performance.

4.1. Training Effect on Investigative Interview Performance

Considering the poor interviewing practices observed (Baldwin, 1993; Fisher et al., 2010; Milne & Bull 1999; Snook et al., 2010), training efforts have focused on increasing the interviewing skills of police officers and enabling them to obtain truthful information through the use of non-suggestive and non-coercive interview methods (Milne & Bull, 1999; Schollum, 2005; Smets, 2009). The basic assumption behind these efforts was that most officers can learn to be effective interviewers with right instruction (Baldwin, 1993; Memon, Holley, Milne, Koehnken, & Bull, 1994; Williamson, 1994). Laboratory research (Köhnken et al. 1999; Memon et al. 2010) and real-life police training studies (MacDonald, Snook, & Milne, 2017) have shown that people can be trained to apply the evidence-based interviewing techniques such as Cognitive Interview and PEACE model. However, studies on the effectiveness of the training courses on these interview techniques have found mixed results. MacDonald, Snook, and Milne (2017), for instance, analyzed the impact of PEACE interview training on the performance of interviewers by examining 80 interviews conducted by Canadian police officers based on 38 interview practices.

[2] PEACE is an evidence-based interview model developed in the UK in early 1990s. The components of the acronym PEACE stands for Planning & Preparation, Engage & Explain, Account, Closure, and Evaluation.

They found that trained officers performed better than their untrained colleagues. Specifically, trained officers asked more open-ended and fewer leading questions and they used the relatively simple skills they learned on the three phases of the PEACE model (i.e. engage and explain, account, and closure) more frequently. However, complex skills such as memory enhancing techniques were not applied by the trainees at the expected level, and despite the improvements in the number of open-ended questions, the percentage of these questions was found to be lower than the findings in other similar studies. (Myklebust & Alison 2000; Snook & Keating 2010; Wright & Alison 2004).

Some studies found that the impact of interview training on skills such as effective communication and building rapport has been limited (Aldridge & Cameron, 1999; Clarke, Milne & Bull, 2011). For example, Aldridge and Cameron (1999) found that training had little effect on the questioning style and rapport building skills of police officers. When they compared the performance of trained and untrained police officers in certain aspects of child witness interviews such as requesting free accounts and the number of open, specific, leading, and non-leading questions used in the interview, they found no differences between the two groups and both groups were rated poorly in these skills. Similarly, Clarke, Milne and Bull (2011) analyzed the performance of interviewers from six police forces in England and Wales in 174 real-life suspect interviews and found no difference between trained and untrained officers based on a rating scale constructed to evaluate the behaviours underpinning the PEACE model.

Studies on the effectiveness of Cognitive Interview training also showed mixed results. Memon et al. (1994) examined the effects of a four-hour Cognitive Interview training on the witness interview performances of experienced police officers in a staged armed robbery case. They found that the amount of correctly recalled details did not significantly increase after the

training. Moreover, they found that the trained officers could not successfully apply the techniques taught in the training and continued to use leading questions. More recently, Dando, Wilcock, and Milne (2008) asked 221 young non-specialist police officers from five UK police forces complete a questionnaire about their use of the Cognitive Interview techniques and their practical experiences. For the question on how well their training had equipped them with the interviewing skills necessary to interview the witnesses, 71% of the respondents reported that they felt not at all equipped or not very well equipped although they received initial recruit training.

There are also studies that found positive outcomes of Cognitive Interview training. For example, Clifford and George (1996) analyzed the impact of a four-day training on the interview performances of twenty-eight experienced police officers. The officers were split into four groups and the first three groups were trained either on one of the Cognitive Interview (CI) and Conversation Management (CM) techniques or both of them. The last group did not receive any training as they served as the control group. The CI-only training led to the greatest shift after training specifically in terms of the amount of information gathered and appropriate questioning. In fact, the only meaningful increase in the amount of information was observed only after the CI training. The comparison of pre and post training real life interview performances showed that the five techniques taught in the training (i.e. context reinstatement, change perspective, change of temporal order, refrain from editing anything out, and working hard to recall the details) were utilized in less than 10% of interviews after training. They also found that none of the interviewers applied all four Cognitive Interview techniques. The only significant increase in the usage of CI techniques was observed in the context reinstatement.

Griffiths and Milne (2006) analyzed the audio tapes of sixty interviews conducted by fifteen experienced interviewers (four interview per interviewer) after an advanced level three-

week course on suspect interviewing. The evaluation of the interviews based on a scale created by Clarke and Milne (2001) revealed that although the five-tier training helped develop the interview skills of police officers, their performance significantly declined as time passed. They also compared the performance of these interviewers with the untrained officers and found that the trained officers used some practices that they learned (e.g. delivery of legal requirements) more than untrained officers. However, there was no significant difference in the usage of some evidence-based interviewing practices such as appropriate questioning, sequence of questioning, and topic structure. Also, while simple skills such as delivering legal rights to suspects were applied by the trainees, they were less likely to use more complex skills that they learned such as structuring the areas of the interview and explanation of the legal requirements.

Warren et al. (1999) evaluated a ten-day interview training program designed to improve the knowledge and skills of child-abuse investigators. The participants (n = 27) completed pre- and post-training questionnaires that assessed their knowledge of the scientific evidence regarding memory, suggestibility, and other key elements that affect recall, and conducted pre-and post-training interviews with preschool children about two experienced events. Participants' knowledge of the scientific basis of the various interview protocols improved post-training, but the training did not have a significant effect on participants' questioning ability or the amount of information obtained from the child interviewees.

According to St-Yves et al. (2014), the difference in the training outcomes in various studies stem from the different designs, the purpose of the interview (e.g. suspect, witness, victim interview etc.), and the level of motivation and capability of the interviewers attended to the research studies evaluating these training courses. For instance, St-Yves et al. (2014) argued, the initial assessments of the PEACE model training programmes were accomplished through the

participation of highly motivated and capable ranked officers while later studies were conducted during the mandatory courses for low-ranked officers. Therefore, the profile and motivation levels of the participants might have affected the outcomes.

One potential explanation for the relative lack of training effectiveness found in some studies outlined above is that there might be some innate differences among individuals which make some of them more open to learn and have better interviewing skills than the others (St-Yves, 2009). If that is the case, the impact of training courses might vary among police officers based on their personalities. Smets (2009) suggests, for example, a non-empathetic police officer is less open to learn to build rapport with interviewees than an empathetic police officer. Therefore, it is important to examine the relationship between personality measures and training effect in order to identify which individuals are more likely to be trained on best practices in investigative interviewing.

To our knowledge, there is no such study examining the impact of personality on the effectiveness of interview training. However, there is some evidence showing the interaction between personality and training performance of police officers in a broader sense (Forero et al., 2009) and other professionals such as marine corps recruits (Dean, Conte, & Blankenhorn, 2006) and military pilots (Carretta et al., 2014). In a longitudinal research design, Forero et al. (2009) examined the impact of personality and training together on the actual performance of 2,010 police candidates. The performance of the subjects were evaluated at three different steps: after academy training, at the end of a 1-year of real-life practice, and at the end of their sixth year in their job through the same evaluations scale. Their performances were also evaluated through the number of disciplinary sanctions, disciplinary proceedings, probations, commendations for excellent performance, and medals that they received during the 6-year period. The personality of the

subjects were measured through Law Enforcement Assessment and Development Report, which includes a set of personality traits scores based on the 16 Personality Factor Questionnaire (16PF-5; Cattell, Cattell, & Cattell, 1993) and the Clinical Analysis Questionnaire (CAQ; Krug, 1980). As a result of a Structural Equation Model analysis, Forero et al. (2009) found that the job performance of police officers is influenced by both personal dispositions and training. However, the impact of personality alone moderately predicted the variance in the performance in both first year and sixth year evaluations. When the influence of personality is mediated by training, the new model explained the 60% of performance variance. Based on these findings Forero et al. (2009) concluded that training is a mediator of personal dispositions when predicting the job performance, thus, police selection should be accomplished by considering the mediating impact of academy training program on the personality measures.

4.2. Study 3

As mentioned, previous research has not reached a conclusion on whether training improves the performance in investigative interviewing as they found mixed results. There is an obvious need for understanding the reasons for the ineffectiveness of training in this domain. The impact of individual differences might be one explanation and, to our knowledge, there is no such study examining the relationship between personality characteristics and efficacy of training on investigative interviewing. Therefore, in Study 3, we attempted to answer the following research questions: (1) Does a short training on PEACE investigative interviewing improve the performance of participants? (2) Are there certain personality traits and aptitudes that predict the effectiveness of interview training?

We hypothesized that the training would increase the performance of the participants in all five success measures that will be explained below. We also hypothesized that high scorers in the four aptitude dimensions (i.e. Humane, Communicative-Insisting, Self-Controlled, and Careful-Tenacious) would improve their interview performance after the training more than the low scorers and will perform better than low scorers in post-interviews. Similar to our hypothesis in Study 2, high scorers in the Communicative-Insisting might perform poorly in appropriate questioning as they might ask more leading questions and talk more than the interviewee talks. Finally, we hypothesized that the high scorers in Conscientiousness and Openness dimensions of the Big Five would have a better training performance than low scorers, while high scorers in Neuroticism would perform worse than low scorers. Similar to the Communicative-Insisting dimension of the aptitudes scale, Extraversion might be related to lower performance in appropriate questioning.

4.3. Method

Participants

The sample consisted of 45 policing students enrolled in a Southeastern Ontario College and 83 undergraduate students enrolled in an introductory psychology course at a Southeastern Ontario University. The policing students were assigned as interviewers and compensated with money and a certificate for the training course. The undergraduate students were assigned as interviewees and compensated with extra credit upon their participation. Thirty-eight of the policing students completed all three steps (i.e. pre-interview, training course, and post-interview). Among the policing students who completed all steps (interviewers), there were 27 (71%) male participants and 11 (29%) female participants. The mean age of the interviewers was 23 (*SD* = 6.29, *Range* = 17 – 48). In terms of ethnicity, 22 (57.9%) of the interviewers self-reported as

Caucasian, 7 (18.4%) as Asian, 4 (10.5%) as from other ethnicities, 3 (7.9%) as Black, 1 (2.6%) as Middle Eastern, 1 (2.6%) as Latin American. The majority of the interviewers (89.5%) identified English as their first language.

The demographics of interviewee participants were similar to those of interviewers. Their mean age was 21 (*SD* = 5.12 *Range* = 18 – 50). Fifty-seven (67.9%) of the witnesses were female, while 27 (32.1%) of them were male participants. The ethnicity of the witnesses had a slightly different variation from those reported by interviewers: White – 27 (32.1%), Asian – 23 (27.4%), Other – 13 (15.5%), Middle Eastern – 10 (11.9%), Black – 7 (8.3%), Latin American – 4 (4.8%). Fifty-two of the interviewees (62.6%) identified English as their first language.

Materials

Mock crime videos

Two different mock crime videos were used for pre and post interview sessions. The target videos used in this study consisted of mock theft crimes and both of the crime scenarios were staged in an office setting. Both videos were recorded by a person wearing a body-worn camera, thus, they were showing the crime from a first person perspective, each video was approximately 3 minutes in length. The following statement was shown at the beginning of the videos:

> In a moment, you will see a short video of an event. The event happened yesterday at 3 PM in an office on the 6th floor of the X Bank building in downtown (In Video 2: yesterday at 11 AM in an office on the 3rd floor of the Y Hall building in downtown). Please note that the video was made from a first-person perspective by someone wearing a body-worn camera. Please imagine, while you are watching, that you are actually experiencing the event from this perspective as it unfolds in the video.

In Video 1, a male entered in a lab and stole the laptop of a female graduate student when she went out of the office to get some water. In Video 2, another male stole the wallet and phone of a graduate student while he was in the next room with his colleague to discuss on a stats issue. There was similar amount of details in the two videos such as the office supplies, discussions between graduate students, clothes of the actors etc.

Procedure

The study was conducted in three steps: pre-interview, training, and post-interview. In the pre-interview step, the policing students and undergraduate students were invited to the lab in pairs. In each session, the policing students were assigned as interviewer and the others as the witness. After both participants signed consent forms, the witness was invited to watch the short video of a mock theft crime on a wide screen with a headphone while the interviewer was waiting in a separate room. Before the interview, the interviewer was informed that they would conduct an interview with the other participant about a theft crime which he or she recently witnessed. They were also told that the main objective was to elicit as much information as possible from the witness during the interview. The key reminder to the interviewers was that there was not a certain procedure that they had to follow during the interview and they just needed to do the task – elicit information about the crime from the witness – as effectively as possible. After the instructions, both participants were asked to complete some questionnaires. The interviewers completed the five-factor model, aptitudes, scale, and demographics questionnaires. The witnesses completed the demographics and five factor model questionnaires before the interview and the witness perception questionnaire after the interview. These questionnaires were described above in the methodology part of Study 2. After completing the questionnaires, participants were invited to the interview

room and the interviews were videotaped. The videotapes were transcribed to use when coding and rating the data.

The second step was a half-day training course on PEACE model of investigative interviewing. In five different sessions, the participants who conducted the pre-interview were trained by Dr. Joseph Eastwood on the basics of PEACE model of investigative interview and Cognitive Interview techniques. The following topics were covered in the training: (1) The importance of interviewing in criminal investigations, (2) planning and preparation phase before the interview, (3) rapport building, (4) ways to engage interviewees in the process, (5) explaining the purpose of interview, (6) the basics if how human memory works and its implications for interviewing, (7) Cognitive Interview techniques, (8) appropriate and inappropriate questions types, (9) proper interview behaviors, and (10) note-taking. The videotape of a portion of a real-life interview in which some Cognitive Interview techniques were used (e.g. eye-closure, focused concentration, context reinstatement, sketching) was shown to the trainees during the training. Also, one volunteer trainee was interviewed in class about a traffic accident that he or she experienced using several of the Cognitive Interview techniques in order to further show the trainees how those techniques are applied.

The third step was a post-interview session. Within the first two weeks after the training, the trainees were invited to the lab again and conducted a second interview with a new undergraduate student serving as the witnesses. After signing the consent form, the witnesses watched the second video which had a similar scenario and number of details with the first video. Similar to the pre-interview step, the interviews were videotaped and transcribed. After the interview, the witnesses evaluated the performance and attitude of the interviewers through the same questionnaire used in the pre-interview session.

4.4. Data Coding

Similar to Study 2, the success of the interviewer was assessed in both pre and post interview tasks based on five different measures: (1) perception of the interviewee, (2) amount of information elicited, (3) the number of details per question, (4) researcher ratings, (5) appropriateness of questioning.

a) The witness perception: The witness perception was measured through the witness perception questionnaire which is explained in Study 2 (see Appendix A). The responses for each question were scored from one to five respectively and total score based on the forty-four questions was computed.

b) The amount of information elicited (The number of details): The amount of accurate information that the interviewer could elicit during the pre and post interviews was counted based on the coding scheme created for both crime videos. In video 1, the maximum number of details that could be elicited was 270, and that number in video 2 was 238. Two research assistants coded each detail given by the interviewee and the total amount of accurate details was recorded for each interviewer. The coders were provided a 2-hour training on the methodology that they need to follow when coding the details in the videotapes. After the research assistants coded the first five videotape another 1-hour training session was held. In this session, the differences between the decisions in coding were discussed and an agreement was reached on those points.

c) The number of details per question: The number of details elicited per question was measured by dividing the number of details gathered by interviewers by the number of question asked in both pre and post interviews. In the evidence-based interview training that the participants received, they were taught that the interviewer should talk less than the interviewee, allow them

talk more without any interruption, ask fewer but open-ended questions, and avoid leading questions. Therefore, the trainees were expected to elicit more details with less questions in the post-interview than they did in the pre-interview.

d) Researcher ratings: The performance of the interviewers were rated by two research assistants based on a scale that was created by Clarke et al. (2011), which was explained in the Methodology part of Study 2 (see Appendix B). Before starting to rate the interviews, the raters were provided a 2-hour training session through which the instructions and materials that will be used in rating were described. To ensure that the second raters understood the rating procedure, a 2-hour trial session was held to practice on five interviews and eliminate the potential confusions and misunderstandings.

e) Appropriateness of questioning: Similar to Study 2, the appropriateness of the questioning was measured through division of the number of appropriate questions by the total number of questions asked by the interviewer (appropriateness of the questioning = appropriate questions / [appropriate questions + inappropriate questions]). Two coders counted the number of questions per each type (open-ended, closed, multiple, long/complex, leading) for each participant.

Reliability and Inter-rater Agreement

Reliability and inter-rater agreement analyses were conducted for the success measures. The reliability analysis conducted for the 44 items of the interviewee perception questionnaire based on 83 interviewee responses given by the pre and post interview witnesses showed that the scale is internally consistent ($\alpha = .85$). The number of details were coded by two coders and the inter-coder agreement analysis revealed a high agreement between the coders (Kappa= 0.89). The inter-rater agreements for the researcher rating scale and the appropriateness of questioning were analyzed through Weighted Kappa. The overall Weighted Kappa value for the researcher rating

scale was 0.63 (95% CI, .610 to .656), $p < .001$, which means a substantial level of agreement between the two raters was achieved (Landis & Koch, 1977). The inter-rater reliability for the items in the appropriateness index indicated a moderate level of agreement between the two coders, Weighted Kappa = .59 (95% CI, .526 to .651), $p < .001$.

4.5. Data Analysis

Before the analysis, the data were screened to test the outliers, normality, linearity, and the homogeneity of variance-covariance. There were no outlier case based on the Mahalanobis distance calculation. Normality, linearity, and the homogeneity of variance-covariance of the data were assumed based on scatterplot and standardized residual tests. After the data screening, the impact of the training on the success measures and the relationship between personality measures and training effect were analyzed. To examine the training effect by comparing the mean success measure scores in pre and post interviews, a paired samples t-test analysis was conducted for each success measure (i.e. number of details elicited, witness perception score, researcher rating score, and the appropriateness of questioning). Before the t-test analyses, the normality of each success measure in both pre and post-interview conditions was tested. Based on the skewness and kurtosis measures, normality was assumed for all success measures except for the detail per question variable. The skewness and kurtosis levels for the detail per question variable in post-interview condition were estimated at 3.34 and 14.54 respectively, which is more than the maximum allowable values for a t-test (i.e. skewness < 2.0 and kurtosis < 9.0; Posten, 1984). A log10 transformation was applied to the detail per question variable and a new variable was created which has a normal distribution (skewness = .43 and kurtosis = -.04).

The relationship between the personality measures and interview success measures for both pre and post interview tasks were analyzed through Pearson's correlation tests. To examine the relationship between the personality characteristics (i.e. aptitude dimensions and Big Five dimensions) and the training effect, a series of Pearson's correlation tests were conducted. Before the analysis, the changes in each success measure were calculated as percentage and new dependent variables created for each percentage change. Table 11 shows the descriptive statistics for the dependent variables that indicate the percentage changes for each success measure after the training. The change in the interview length and number of questions asked are also given in Table 11. The correlations between the dependent variables are given in Table 12.

Table 11. Descriptive Statistics for the Dependent Variables

	Mean	Std. Dev.	Minimum	Maximum
Change-WitnessScore (%)	28.88	72.03	-51.47	266.67
Change- N. of Details (%)	57.50	142.97	-59.32	739.50
Change- Researcher Rating (%)	13.52	20.69	-19.40	100.00
Change-Appropriateness (%)	5.53	17.35	-32.77	55.56
Change-Details per question (%)	504.62	848.28	-36.09	3772.00
Change-N. of questions (%)	26.31	214.21	-81.58	1250.00
Change-Time (%)	123.07	175.94	-56.01	818.54

Table 12. Correlations Between the Dependent Variables

	Change-Witscore	Change-Details	Change-Researcher Rating	Change-Appropriateness	Change-Details per Q.
Change-WitnessScore	1	-0.082	0.018	0.009	-0.166
Change-N. of Details	-0.082	1	0.25	-.378*	.420**
Change-Researcher Rating	0.018	0.25	1	-0.151	-0.201
Change-Appropriateness	0.009	-.378*	-0.151	1	-.315*
Change-Details per Q.	-0.166	.420**	-0.201	-.315*	1

**p < 0.01 (1-tailed).
*p < 0.05 (1-tailed).

4.6. Results

The paired-samples t-test analyses showed that attending the training course had a significant positive effect on all success measures except for the appropriateness of questioning (see Table 13). After the training, there was a significant increase in the witness perception scores ($t(37) = -2.11$, $p < .05$), number of details ($t(37) = -2.062$, $p < .05$), researcher ratings ($t(37) = -4.525$, $p < .001$), and detail per question ($t(36) = -6.608$, $p < .05$). We also analyzed the difference in the length of interview time and number of questions asked. There was a decrease in the mean number of questions asked but this change was not statistically significant. We found that the mean length of interview time almost doubled after the participants received the training. To test whether the longer time period led to a rise in the amount of information elicited in the post-interviews, a separate correlation analysis was conducted. A strong correlation was found between the time spent and number of details elicited ($r = .62$, $p < .001$).

Table 13. Difference in the Success Measure Scores (Pre-Post Interview)

		Mean	Std. Deviati	t	df	Sig. (2-tailed)
Witness Perception Score	Pre	50.21	17.344	-2.115*	37	0.041
	Post	57.58	15.613			
Number of Details	Pre	12.98	6.414	-2.062*	37	0.046
	Post	16.16	7.151			
Researcher Rating Score	Pre	61.55	8.858	-4.525**	37	0.000
	Post	68.74	9.366			
Appropriateness Index Score	Pre	0.88	0.112	-1.715	37	0.095
	Post	0.92	0.11			
Detail per question	Pre	0.91	0.261	-6.608**	37	0.000
	Post	3.64	0.347			
Interview Length (sec)	Pre	337.13	226.324	-4.289**	37	0.000
	Post	614.13	419.599			
Number of Questions	Pre	23.37	18.201	1.166	37	0.251
	Post	19.76	16.935			

* $p < .05$
** $p < .01$

After comparing the overall success measure scores, we analyzed the differences in specific items in the rating scale by conducting a paired samples t-test at item level. The analysis showed that the interviewers were rated significantly higher after the training on 10 of the 23 items: introducing self, explaining interview purpose, providing details of routine and route map, rapport building, encouraging interviewee to give an uninterrupted account, appropriate structure, using pauses and silence, self-confidence, summarizing interview, inviting interviewees to add, alter or correct their account. There was no statistically significant change on the remaining 13 items in the scale.

The analysis of the correlations between personality measures and the interview success measures in the pre and post interview tasks showed some significant outcomes. According to the pre-interview outcomes, none of the Big Five Dimensions was correlated with the success measures. In post-interviews, however, there were several significant correlations (see Table 14).

High scorers on the Openness/Intellect dimension performed better than low scorers both in researcher rating score ($r = .49$, $p < .01$) and witness perception score ($r = .35$, $p < .01$). They also asked more questions than low scorers ($r = .36$, $p < .05$) and their interviews took longer than those of low scorers ($r = .44$, $p < .01$). Extraverts scored higher in researcher ratings ($r = .40$, $p < .05$) and held longer post-interviews ($r = .34$, $p < .05$). In addition, Extraverts asked more questions in both pre-interview ($r = .32$, $p < .05$) and post-interview tasks ($r = .34$, $p < .05$). Researchers rating scores in post-interviews were positively correlated with Agreeableness ($r = .37$, $p < .05$) and negatively correlated with Neuroticism ($r = -.37$, $p < .05$).

Table 14. Correlations between Big Five Dimensions and Success Measures (pre and post)

		N	A	C	E	O
Witness Perception Score	Pre	0.002	0.115	0.078	0.291	-0.017
	Post	-0.092	.369*	0.116	0.206	.352*
Number of Details	Pre	0.011	-0.176	-0.165	-0.043	-0.071
	Post	-0.1	-0.167	0.114	0.141	0.197
Researcher Rating Score	Pre	-0.226	0.097	0.037	0.216	0.140
	Post	-.370*	0.282	0.22	.399*	.491**
Appropriateness Index Score	Pre	0.006	-0.29	-0.046	-0.057	0.108
	Post	-0.106	-0.092	0.041	0.105	0.053
Detail per question	Pre	0.011	0.052	-0.051	-0.053	-0.142
	Post	0.034	-.454**	0.04	-0.137	-0.239
Interview Length (sec)	Pre	-0.031	0.01	0.02	0.261	0.057
	Post	-0.153	0.083	0.132	.345*	.440**
Number of Questions	Pre	-0.044	0.064	0.05	.320*	0.095
	Post	0.12	-0.042	0.161	.338*	.363*

N: neuroticism, E: extraversion; O: openness to experience, A: agreeableness, C: conscientiousness.
*p<0.05
**p<0.01

In the analysis of the correlation between the aptitude dimensions and success measures (see Table 15), we found that researcher rating scores in post-interviews were significantly correlated with Communicative-Insisting ($r = .40$, $p < .05$) and Careful-Tenacious dimensions ($r = .46$, $p < .01$). High scorers in the Humane dimension scored lower in the appropriateness of questioning in the pre-interview task ($r = -.36$, $p < .05$) and elicited less details per question in post-interviews than low scorers ($r = -.39$, $p < .05$). High scorers in Careful-Tenacious dimension held longer interviews in both pre-interview ($r = .34$, $p < .05$) and post-interview ($r = .36$, $p < .05$) and asked more questions in pre-interview ($r = .36$, $p < .05$). The post-interviews conducted by Communicative-Insisting participants took longer ($r = .35$, $p < .05$) and more questions asked in the interviews ($r = .34$, $p < .05$).

Table 15. Correlations between Aptitudes and Success Measures (pre and post)

		Humane	Communicative-Insisting	Self-Controlled	Careful-Tenacious
Witness Perception Score	Pre	0.285	0.109	0.075	0.13
	Post	0.304	0.182	-0.016	0.16
Number of Details	Pre	-0.208	-0.003	-0.125	0.202
	Post	-0.153	0.002	-0.062	-0.015
Researcher Rating Score	Pre	-0.136	0.128	0.167	0.25
	Post	0.167	.404*	0.314	.459**
Appropriateness Index Score	Pre	-.364*	0.097	-0.01	0.163
	Post	-0.021	0.226	-0.058	0.197
Detail per question	Pre	0.01	0.04	-0.02	0.166
	Post	-.393*	-0.096	-0.208	-0.186
Interview Length (sec)	Pre	0.07	0.182	-0.012	.343*
	Post	0.238	.347*	0.181	.361*
Number of Questions	Pre	0.098	0.162	0.041	.365*
	Post	0.199	.342*	0.129	0.295

*p<0.05
**p<0.01

The change in the performance scores was partially explained through personality measures. The Pearson's correlations test indicated that only the Humane dimension of the aptitudes scale was significantly correlated with the change in two of the success measures: researcher rating ($r = .30$, $p < .05$) and appropriateness of questioning ($r = .29$, $p < .05$) (see Table 14). It was also found that Communicative Insisting dimension was significantly correlated with the change in the length of interview ($r = .33$, $p <.05$) time and number of questions asked ($r = .27$, $p < .05$) (see Table 16).

Table 16. Correlations between Aptitude Dimensions and Success Measures

	Humane	Communicative-Insisting	Self-controlled	Careful-Tenacious
Change-WitnessScore	-0.092	0.109	-0.102	0.075
Change- Number of Details	0.026	-0.263	0.104	-0.120
Change-Researcher Rating	0.304*	0.262	0.135	0.209
Change- Appropriateness of Q.	0.296*	0.106	-0.046	0.010
Change-Detail per question	-0.025	0.064	-0.086	0.013
Change-Time	0.218	0.333*	0.153	0.139
Change-Number of Questions	0.082	0.272*	0.017	0.035

**p < 0.01 (1-tailed).
*p < 0.05 (1-tailed).

The analysis of Big Five dimensions in relation to the training outcome showed that the Openness/Intellect dimension was significantly correlated with the change in the researcher rating scores (see Table 17). Another interesting finding about the Openness/Intellect dimension was that high scorers in this dimension spent significantly much higher time in their post-interviews than the low scorers did ($r = .29$, $p < .05$), while the other Big Five dimensions have no effect on the change in the time spent in interview.

Table 17. Correlations between Big Five Dimensions and Success Measures

	N	A	C	E	O
Change-WitnessScore	-0.077	0.145	0.107	0.015	0.126
Change-Details	0.086	0.145	0.008	-0.152	-0.224
Change-Researcher Rating	-0.122	0.173	0.152	0.168	0.275*
Change-Appropriateness	-0.050	0.177	0.039	0.145	-0.091
Change-Detail per question	0.201	0.068	-0.226	0.005	0.146
Change-Time	-0.001	0.037	0.011	0.174	.294*
Change-Number of Q.	0.264	-0.188	0.036	0.086	0.217

N: neuroticism, E: extraversion; O: openness to experience, A: agreeableness, C: conscientiousness.

**p < 0.01 (1-tailed).
*p < 0.05 (1-tailed).

To better understand the nature of the correlations between the Big Five dimensions and the change in the success measure scores, the two facets of each dimension were also analyzed. The version of the Five Factor Model that was used in this study was adopted from the DeYoung et al.'s (2007) Big Five Aspects Scale (BFAS) as explained in Study 2. BFAS has ten aspects and each of the Big Five dimensions consists of two aspects (Neuroticism: Withdrawal and Volatility; Agreeableness: Compassion and Politeness; Conscientiousness: Industriousness and Orderliness; Extraversion: Enthusiasm and Assertiveness; Openness/Intellect: Intellect and Openness). As shown in Table 18, the Intellect facet of the Openness/Intellect dimension was significantly correlated with the change in the researcher rating scores ($r = .28$, $p < .05$), whereas the Openness facet of the same dimension was not correlated. It is also important to note that the change in the length of time was negatively correlated with the Politeness facet of the Agreeableness dimension. Finally, the Withdrawal facet of Neuroticism was positively correlated with the number of questions asked.

Table 18. Correlations between 10 Big Five Aspects and Success Measures

Big Five	Facets	Change-Witness Score	Change-Details	Change -Rating Score	Change-Appropr.	Change-Time	Change-Number of Q.
N	Withdrawal	-0.065	0.159	-0.121	0.022	0.029	.273*
	Volatility	-0.070	-0.005	-0.093	-0.108	-0.030	0.191
A	Compassion	0.245	0.008	0.183	0.233	0.177	0.054
	Politeness	-0.016	0.268	0.110	0.058	-0.138	-.418**
C	Industriousness	-0.014	-0.049	0.152	-0.104	-0.006	0.022
	Orderliness	0.213	0.076	0.097	0.205	0.029	0.040
E	Enthusiasm	0.060	0.027	0.202	0.141	0.069	-0.001
	Assertiveness	-0.031	-.278*	0.092	0.111	0.226	0.145
O	Intellect	0.200	-0.151	.278*	-0.106	.322*	0.211
	Openness	-0.020	-0.228	0.164	-0.038	.294*	0.217

**p < 0.01 (1-tailed).

*p < 0.05 (1-tailed).

The facet level analysis of the Big Five dimensions and post-interview performance also helped better unpack the nature of the relationship between personality and training effect (see Table 19). The researcher rating scores for the post-interview performances of the participants were significantly correlated with the Withdrawal facet of Neuroticism ($r = .36$, $p < .05$), the Compassion facet of Agreeableness ($r = .44$, $p < .01$), the Enthusiasm ($r = .36$, $p < .05$) and Assertiveness ($r = .34$, $p < .05$) facets of Extraversion, and the Intellect facet of Openness/Intellect ($r = .59$, $p < .01$). The witness perception scores were correlated with the Compassion facet of Agreeableness ($r = .53$, $p < .01$) and the Intellect facet of Openness/Intellect ($r = .33$, $p < .05$).

Table 19. Correlations between 10 Big Five Aspects and Post-Interview Success

Big Five	Facets	Witness Score	Details	Rating Score	Appropr.	Time	Number of Q.
N	Withdrawal	-0.01	-0.233	-.357*	-0.144	-0.205	0.026
	Volatility	-0.15	0.055	-0.292	-0.043	-0.065	0.182
A	Compassion	.519**	-0.117	.435**	0.035	0.257	0.147
	Politeness	0.084	-0.177	0.02	-0.214	-0.145	-0.252
C	Industriousness	0.02	0.177	0.216	0.034	0.121	0.073
	Orderliness	0.196	-0.006	0.147	0.036	0.099	0.215
E	Enthusiasm	0.209	0.092	.356*	0.054	0.237	0.29
	Assertiveness	0.15	0.15	.335*	0.126	.356*	0.294
O	Intellect	.334*	0.212	.586**	0.004	.377*	0.317
	Openness	0.237	0.101	0.181	0.094	.346*	0.279

**$p < 0.01$ (1-tailed).
*$p < 0.05$ (1-tailed).

4.7. Discussion

The purpose of Study 3 was to measure the impact of a short training on the investigative interview performance and how that training effect interacted with personality measures when predicting the interview performance within a policing student sample. The findings on the training effect and personality-training relationship will be discussed separately.

4.7.1. Training effect

Findings indicated that the training improved the performance of the participants in most of the success measures. The highest increase was observed in the number of details per question (57% on average), which means that the interviewers elicited more information with less questions after they received the training. The mean percentage increase in the witness perception scores was 29%. This shows that participants built better rapport and adopted better attitudes towards interviewees in post-interview session than they did in the pre-interview. There was a very high

increase in the time spent in the interviews on average (123%), while the mean number of questions asked slightly dropped (-26%). In addition, there was a strong correlation between the length of time and the number of details elicited in post interviews. Taken together, it can be concluded that the training improved the participants' ability to elicit more information and build better rapport while asking less questions.

The mean increase in the researcher rating scores was 13%, which is a lower effect than that was observed in the witness perception score. The reason for this relatively limited effect was the fact that the participants rating score did not increase in more than half of the items although there is an increase in the total researcher rating score. Thirteen out of 23 items in the researcher rating scale did not significantly increase in the post-interview. The 10 items that the interviewers performed better after the training were more about the Engage and Explain phase of the PEACE model such as introducing self, explaining interview purpose, providing details of routine and route map, rapport building. On the other hand, the items of the rating scale that we could not observe any significant increase were more about the Account phase and the questioning style of the interviewer such as development of topics, keeping interview to relevant topics, covering points to prove, and exploration of information. In addition, four out of five interviewer characteristics did not significantly improve after the training according to the research rating scores. The interviewers received higher score only in the self-confidence item in that portion of the scale, while there was no significant change in the scores for open-mindedness, flexibility, communication skills, and active listening items.

This differential finding was very similar to that of Griffiths and Milne (2006) in their analysis of the impact of an advanced level three-week course on suspect interviewing performance of experienced police officers. They found that the training improved the basic skills

of the interviewers such as delivery of legal requirements but not the usage of some complex practices such as appropriate questioning, sequence of questioning, and structuring the areas of the interview. Griffiths and Milne's (2006) concluded that the PEACE training which the experienced officers received was "less effective at improving the ability of officers to obtain and probe accounts" (p.187). Similar to that comment, we can conclude that the training in this study was more effective in improving the participants' attitudes toward interviewees, their rapport building and other skills related to the Engage and Explain Phase, and their self-confidence than improving the questioning and probing skills of the participants. These improvements led to a substantial increase in the witness perception scores and the number of details elicited in the interview.

4.7.2. Personality-Training Effect Relationship

The impact of personality on the training effect was analyzed by using two different personality measures: (1) The aptitudes scale created in Study 1, which is an extended version of DeFruyt et al.'s (2006) scale, (2) Five Factor Model (Big Five) personality scale. The training effect was measured in two different ways: (1) By measuring the percentage changes that occurred in each success measures after the training and (2) By comparing the personality effects on pre and post interview performances.

4.7.2.1. Aptitudes and Training Efficacy

Out of four dimensions of the aptitudes scale, two of them (i.e. Communicative-Insisting and Careful Tenacious) predicted high researcher ratings in post-interview. The performance of high scorers in the Careful-Tenacious dimension was rated by researchers significantly higher than the low scorers in post-interview ($r = .46$, $p < .01$), which shows their adherence to training based on the items in the rating scale. Careful-Tenacious dimension includes aptitudes like being aware

of one's own thoughts and actions, thorough, careful, able to put things in perspective, and paying attention to details. The PICI scale of DeFruyt et al. (2006) has also Careful-Tenacious dimension with some differences in its components. According to DeFruyt et al. (2006), "high scorers on this dimension systematically and thoroughly examine, and scan all elements using a step-by-step approach with much attention to detail and sustained focus" (p.574). Thanks to these characteristics, high scorers might have benefited from training more than low scorers did and apply what they learned carefully as much as they could.

The interviewers who scored higher on Communicative-Insisting dimension were also rated more positively than low scorers in post-interview ($r = .40$, $p < .05$). Also, they spent more time and asked more questions in their post-interviews than they did in the pre-interviews. Communicative-Insisting dimension is characterized with items like being talkative, assertive, fluent in social contexts, able to respond quickly and appropriately, taking action, and speaking one's own mind. After the training, high scorers on this dimension might have become more eager to talk and elicit more information, and this motivation might have led them to spend more time and ask more question in the post-interview than they did in the pre-interview.

Only the Humane dimension was correlated with the percentage change in the researcher rating score ($r = .30$, $p < .05$) and the change in the appropriateness of questioning ($r = .29$, $p < .05$). This means that the performance of the participants who scored higher in the Humane dimension improved more than the performance increase of the low scorers after the training. As mentioned above, much of the performance increase based on the researcher rating scores stemmed from the skills in the Engage and Explain phase of the interview which includes the items like rapport building, introducing self, explaining interview purpose, and providing details of routine and route map. As we found in Study 1, the Humane dimension includes the aptitudes such as

being understanding, sensitive, sympathy, empathetic, sincere, open-minded etc. The relationship between the Humane dimension and training effect suggests that the people who have Humane characteristics can be trained better than others on the skills such as rapport building, engaging with the interviewee, and explaining the interview purpose. This finding was supported by Smets' (2009) argument that a non-empathetic police officer might be less open to learn to build rapport with interviewees than an empathetic police officer was.

Along with the increase in researcher rating scores, high scorers in the Humane dimension also developed their appropriateness of questioning after the training more than the low scorers. In the post-interview, they asked less inappropriate questions (i.e. leading, complex, long and multiple questions) and more appropriate questions (open ended and closed questions). This shows the adherence of the high scorers in the Humane dimension to the training in terms of appropriate questioning. Humane dimension of the aptitudes scale includes items like open-minded, understanding, polite, and able to act gently. These aptitudes might have prevented the participants from asking inappropriate questions after they learned in the training that these types of questions would be harmful for the interviewee and their memory. Especially their open-minded characteristic might have helped them to make fewer assumptions, interrupt the interviewee less, and keep an open-mind during the interview by asking more open–ended questions.

4.7.2.2. Big Five Traits and Training Efficacy

The performance of the participants in the pre-interview task was not correlated with any of the Big Five dimensions. However, in post-interviews, high scorers on Agreeableness, Extraversion and Openness performed better than low scorers, and high scorers on Neuroticism performed worse than low scorers. In post-interviews, researcher rating scores were positively

correlated with Extraversion ($r = .40$, $p < .05$) and Openness/Intellect ($r = .49$, $p < .01$), and Agreeableness was positively correlated with witness perception score ($r = .37$, $p < .05$). In Study 2, similar correlations between Big Five dimensions and success measures were found with a larger student sample. What we found differently in the post-interview task in Study 3 were the correlation between the Openness/Intellect dimension and witness perceptions score ($r = .35$, $p < .05$) and that between the Neuroticism dimension and researcher rating score ($r = -.37$, $p < .05$).

The increase in the success measures was explained through only one of the Big Five personality dimensions. Although we found that four of the Big Five dimensions were significantly correlated with the post-interview performance, only the Openness/Intellect dimension was found to be significantly correlated with the performance increase ($r = .28$, $p < .05$). High scorers in the Openness/Intellect dimension were rated more positively for their post-interview performance than their pre-interview performance. Among the Big Five factors, Openness/Intellect dimension was the best predictor of performance increase and adherence to training as it was correlated with more success measures with higher correlation values than the other dimensions were. This finding is in line with the previous research on the impact of the Openness/Intellect traits on learning in professional settings. In a recent study, Lafontaine and Cyr (2016) assessed the relationship between personal attributes of twenty-four Canadian police officers and their adherence to training on interviewing with child victims of sexual abuse. Through mock interviews conducted by the officers who received a one-week training on an interview protocol, Lafontaine and Cyr (2016) assessed the performance of the police officers. They found that the interview performance of the officers was positively predicted by cognitive abilities ($\beta = .41$, $p < .05$), female gender ($\beta = .48$, $p < .05$), and the personality trait of Openness to Experience ($\beta = .35$, $p < .05$), and negatively predicted by stress management ($\beta = -.38$, $p < .05$) and level of experience ($\beta = -.24$, $p < .05$).

Barrick and Mount's (1991) meta-analysis of 117 studies on the relationship between personality traits and job performance measures including training proficiency showed that the training proficiency was predicted by the Openness to Experience (p = .25) dimensions of the Big Five. Similarly, Salgado's (1997) meta-analysis of 36 studies conducted in European countries showed that training performance is significantly correlated with Openness (p = .26) and Agreeableness (p = .31) dimensions. Based on their findings on the Openness dimension, Barrick and Mount (1991) concluded that the reason for the relatively better training success of the high scorers in this dimension is their positive attitudes toward learning experiences in general; therefore, the Openness dimension can be an actual measure of learning and motivation to learn. High scorers on the Openness/Intellect dimension also spent longer time in the post-interviews than they did in the pre-interviews. This might be an indicator of their relative eagerness to apply the techniques that they learned in the training when compared with low scorers. As discussed in Study 2, Openness has been shown as a predictor of success in the tasks in which creative achievement and originality is needed (Judge et al., 2013). Creativity and originality might help overcome the obstacles that might occur during an investigative interviewing such as unwillingness of the interviewees to share information and unpredictable and dynamic nature of investigative interviews (Smets, 2009; Yeschke, 2003). The participants who scored higher on Openness/Intellect in Study 3 might have become more aware of the necessities of investigative interviewing than low scorers did after they received training on interviewing and they might have presented their creativity and originality in post-interviews.

The analysis of the two-facets of the Openness/Intellect dimensions showed that only the items in the Intellect facet of the dimension contributes to the correlation with higher change in the researcher ratings. In the post-interview task, the Intellect facet also predicted higher witness

perception scores, researcher rating scores, and longer interviews. The Intellect facet includes traits like thinking, learning, and understanding quickly, handling a lot of information, liking to solve complex problems, and formulating ideas quickly (DeYoung et al., 2007). These characteristics of the high scorers in our sample might have helped them to benefit more from the training on investigative interviewing, understand the material quickly and apply them in the post-interview session. DeYoung et al. (2007) explained the distinction between the Openness and Intellect aspects of the Openness/Intellect dimension through their basic markers. The markers of the Openness facet are Fantasy, Aesthetics, and Feelings, while the Intellect facet is known through its association with Ideas which "fluid intelligence or working memory" (p.894). Therefore, high scorers in the Intellect aspect might have better grasped and applied the basic ideas and interview techniques presented in the training than the low scorers did thanks to their fluid intelligence and working memory.

The correlation between the Extraversion dimension and researcher rating scores in post-interviews shows the extent of the adherence of Extraverts to the training. The facet level analysis showed that both the Enthusiasm and Assertiveness aspects of Extraversion are similarly correlated with the researcher rating scores. Extraverts are known with having more positive emotions, friendliness, warmth, being more oriented toward others, active, and sociable (Costa & McCrae, 1992; DeYoung et al., 2007). The effect of these characteristics might have emerged with the contribution of training and led the high scorers on Extraversion to build rapport and engage with the interviewees, and therefore, perform better than low scorers in post-interviews. While the post-interview performance of Extraverts were rated more positively by researchers, high scorers on the Agreeableness dimension were more successful than low scorers from witnesses' perspective similar to the finding of Study 2. Interestingly, the facet-level analysis indicated that

only the Compassion aspect of the Agreeableness contributed to this relationship. Moreover, the Compassion aspect was also correlated with the researcher rating score. The Compassion aspect refers more to the "compassionate emotional affiliation with others" such as warmth, empathy, sympathy, and tenderness (DeYoung et al., 2007, p. 885). Furthermore, high scorers on Agreeableness are known as more courteous, flexible, trusting, good-natured, and tolerant (Ono et al., 2011). With the impact of these traits and the training, high scorers on Agreeableness dimension and specifically Compassion aspect might have built better trust and rapport with interviewees, and thus, were rated more positively by both researchers and witnesses in post-interviews.

Another interesting finding was the negative correlation between the Neuroticism dimension and researcher rating scores. At facet level, the Withdrawal aspect of Neuroticism was the indicator of this correlation. Similar to this finding, Ono et al. (2011) found that the interrogation performance of the law-enforcement agents from the US Air Force Office of Special Investigation was negatively correlated with high scores in Neuroticism ($r = -.40$, $p < .05$). High scorers on Neuroticism are more likely to experience psychological distress, such as anxiety and depression. (Costa & McCrae, 1992). The Withdrawal aspect includes traits like being easily discouraged, embarrassed, and overwhelmed (DeYoung et al., 2007). Although the high scorers in the Neuroticism dimension and Withdrawal aspect received the same training with the other participants, these traits might have made them feel anxious and discouraged when they were asked to perform what they learned in the training in post-interviews.

The correlation between the change in the number of questions asked and two of the Big Five facets (i.e. positively with Withdrawal and negatively with Politeness) is also worth discussing further. The Politeness facet of the Agreeableness dimension is known with the

characteristics like hating to seem pushy, avoiding imposing one's will on others, rarely putting people under pressure (DeYoung et al., 2007). These characteristics of the Politeness facet might have led the high scorers to avoid asking questions in post-interviews even less than they did in their pre-interviews because they learned well the necessity to talk less than interviewee in the training course. On the other hand, the Withdrawal facet of the Neuroticism dimension is characterized by being filled with doubts about things, feel uncomfortable with one's self, worrying about things, and being easily discouraged (DeYoung et al., 2007). Due to these traits, the high scorers in this dimension might have given up the rule of talking less than the interviewee in the post-interview although they learned to do so in the training.

4.8. Concluding thoughts

In Study 3, the impact of a training course on investigative interview performance was evaluated and how the training effect was correlated with individual differences was analyzed. The findings showed that the training improved the interviewers' performance in most of the success measures and individual differences partially explained the increase in the performance and post-interview performance. The item-level analysis of the researcher rating scores showed that not all items in the scale improved after the training. Those that improved were more related to the Engage and Explain phase of the PEACE model of interview. The Humane dimension of the aptitudes scale and the Openness/Intellect dimension of the Five Factor Model were significantly correlated with the percentage change in some of the success measure scores (i.e. Humane with researcher rating score and appropriateness of questioning and Openness/Intellect with researcher rating score). Considering the fact that the increase in the researcher rating scores mostly stem from the Engage and Explain phase, we concluded that high scorers in the Humane dimension could be trained on the skills in this phase more easily than low scorers thanks to their empathetic and

sincere nature. In line with the previous research on training-personality relationship, this study also showed that high scorers in the Openness/Intellect dimension are more likely to benefit from training on investigative interviewing.

The post-interview performance was correlated with four of the Big Five dimensions (Openness/Intellect, Agreeableness, Extraversion, and Neuroticism-negatively) and two of the aptitude dimensions (Careful-Tenacious and Communicative-Insisting). Similar to the findings in Study 2, Five Factor Model was a better predictor of training efficacy than the aptitudes scale. Each of these dimensions explain certain aspects of successful interviewing and a combination of these findings can help identifying the right personnel for this task in police departments and develop training policies accordingly.

There are some limitations of Study 3. First, our sample consisted of 38 policing students. This number of participants might be adequate to test the training effect on the interview performance; however, a larger sample could have given more reliable findings in terms of the personality effect on the training performance. The same research design can be replicated with a larger sample in future research. Secondly, the training given in this study was a half-day course which provided the basics of the PEACE model of investigative interviewing. The real life interview trainings usually takes at least a week and provides extensive information on the interview techniques and gives the opportunity to practice what is learned on the spot during the training. Future research can examine the personality effect after longer and comprehensive real life training courses. Thirdly, we tested training efficacy through the post-interviews conducted within the first two weeks after the training, which enabled us the examine the short run effect of the training. A longitudinal approach might give the opportunity to compare short and long run effects and how personality interacts with both effects. Fourth, we did not use a control group in

this study due to the time-related restrictions and availability of participants. There might have been some factors other than the training that affected the difference between pre and post interview performances of the participants. For example, participants might have been familiarized with the lab atmosphere in pre-interviews, and, coming to the lab as a second time in post-interviews might have eliminated some of their concerns and anxiety, which have contributed to their post-interview performances in addition to the training effect. A control group would have helped us to control such variables. Finally, the interviews conducted in this study was on mock crime scenarios. This helped us to control the amount of information available to elicit across different participants and allowed to compare performances between pre and post interview sessions. However, interviewers might have been more motivated to elicit information if they were to investigate a real crime. Future research can analyze the interview performances of police officers in real life crime investigations before and after they are trained.

Chapter 5: General Discussion

The purpose of the current series of studies was to identify the impact of individual differences on investigative interviewing performance. Testimonial evidence, which consists of the accounts of suspects, victims, and witnesses, plays a crucial role in crime investigations, thus, it is necessary to improve the skill levels and efficacy of investigative interviewers. The poor interviewing performance of police officers and the lack of training effectiveness that were found in previous research indicates the need for examining the reasons for these deficiencies. One potential explanation is the impact of individual differences on the interview performance and training effectiveness. That is to say, there might be some individual differences which make some people more successful in investigative interviewing and enable them to benefit more from the training on investigative interviewing. Whether personality matters in job performance has been discussed and well-researched for the last few decades and substantial evidence has been found to support this idea in a broader sense. As research in personality psychology became more sophisticated, researchers in this field has focused their attention on the performance in specific jobs and tasks rather than considering job performance as a unique outcome. Following this trend, this three-step dissertation attempted to identify which individual differences lead people to be more successful in investigative interviewing and to benefit more from a training course on this task.

Through this three-step research, the following tentative conclusions have been reached: (1) The aptitudes scale created by modifying and extending the Police Interviewing Competences Inventory (DeFruyt et al., 2006) is a valid measure of individual differences and has a four-dimensional structure (i.e. Humane, Communicative-Insisting, Self-controlled, and Careful-Tenacious); (2) Individual differences partially explains the variation in the performance in

investigative interviewing, adherence to training, and training efficacy; and (3) The Five Factor Model is a better predictor of investigative interview performance and training efficacy than the aptitudes scale.

The following correlations between Big Five personality traits and success measures were found: (1) Openness with researcher ratings in both interview performance and training efficacy, and Openness with adherence to training based on witness perception, (2) Agreeableness with witness perception, appropriate questioning, and adherence to training based on witness perception, (3) Extraversion with researcher ratings, inappropriate questioning, and adherence to training based on researcher ratings, and (4) Neuroticism (negatively) with researcher ratings. The correlations between the aptitudes scale and performance measures were as following: (1) Communicative-Insisting with researcher ratings and adherence to training based on researcher ratings, (2) Humane dimension with researcher ratings in training efficacy and the change in appropriate questioning after training, and (3) Careful-Tenacious with adherence to training based on researcher rating. To reach more robust conclusions on the impact of personality and which personality traits are effective, more research is needed. However, some practical implications can be derived from the findings of this dissertation.

One of the purposes of these studies was testing the reliability and validity of the Police Interviewing Competences Inventory, which was created by DeFruyt et al. (2006) as an initial attempt to identify the competences related to investigative interviewing. The research design used both in their study and the replication of their study by Smets (2009) had important limitations that needed to be addressed. As mentioned, they used the self-estimated and imaginary performances of police officers to analyze the personality effect on interview performance. That is, the participants in both studies did not conduct interviews but imagined that they did so and evaluated

their performance. In Study 2, this limitation was addressed through an experimental research design and the actual performances of interviewers were evaluated. In addition, instead of a self-estimation, objective measures of performance were used to evaluate the interview performance. The interviewers were evaluated based on the witness perception scores, the amount of information elicited, and the researcher ratings on the interview-related criteria and question types. Although the four-dimensional structure of the aptitudes scale explained a substantial variance among the participants in Study 1, the findings of Study 2 showed that the scale has a relatively weak predictive validity.

As discussed in the introduction, the literature suggests using narrow traits to assess the performance in specific tasks of a job because broad personality measurement tools such as the Big Five traits may disregard the situational factors that influence the performance in those tasks (Murphy & Dzieweczynski, 2005). It is also suggested to develop specific personality measurement tools for each task based on job-related criteria through examining the necessities of specific tasks (Barrick & Mount, 2005; Schmidt & Hunter, 2004). However, the aptitudes scale that we used in this study was not as efficient as the Big Five scale when predicting the performance in investigative interviewing despite the aptitudes scale included the items being specifically related to investigative interviewing. There are several potential explanations for this finding. First of all, the items in the aptitudes scale were not created through a personality-oriented job analysis as recommended by Tett et al. (1991). Forty of the items in the scale were listed in the PICI scale of DeFruyt et al. (2006) and the rest of the items were gathered from the previous studies and literature on investigative interviewing. The forty items in the PICI were compiled by asking the views of experienced police interviewers on which characteristics make a successful interviewer and reviewing the literature on this topic. A job analysis "is the process of gathering,

analyzing, and structuring information about a job's components, characteristics, and job requirements" (Sanchez & Levine, 2000, p. 809). In personality-oriented job analysis, analytic methods such as interrater reliability agreement on the traits by the subject matter experts (SMEs) are frequently used to test the accuracy of measurement scales (Prien, Prien, & Wooten, 2003). The aptitudes scale used in this study and the PICI were not created through such an analytical method.

A second potential explanation for the difference between the findings in the current research and the literature may be that the fallacy of the assumption that narrow personality measures are superior to broad traits in predicting job performance is not true in this context. Singh (2008) drew the framework of strategic job analysis that addresses the needs for changing workplaces and stresses the importance of using "general and broad traits instead of specific skills and behaviours" (p. 97). He suggests that today's professional organizations change very rapidly, therefore, broader definitions of work-related skills and traits can serve as better measures of "the flexibility and adaptability of employees" and how they can fit the changing environments and strategies of the workplace (Singh, 2008, p. 97). Policing in general and investigative interviewing in particular are among the most rapidly changing and evolving tasks today given the subjectivities and changing dynamics involved. Thus, it can be argued that Big Five personality dimensions can predict the success in investigative interviewing thanks to their ability to foresee the dynamic and subjective nature of this task. Finally, the comprehensiveness of the Big Five scale and its applicability to a variety of contexts that have been well evidenced for decades (Costa & McCrae, 1992; Ozer & Benet-Martínez, 2005) can also explain its superiority to the aptitudes scale in predicting interview performance in this study.

Overall, three of the Big Five dimensions (i.e. Agreeableness, Extraversion, and Openness) were associated with one or two of the measures of interview performance in Study 2, four dimensions (i.e. Openness/Intellect, Agreeableness, Extraversion, and Neuroticism-negatively) predicted adherence to training, and one dimension (i.e. Openness/Intellect) predicted training efficacy in Study 3. Each of these traits might have contributed to different skills needed to be successful in investigative interviewing. The Openness/Intellect dimension was the best predictor of success as it was related with interview performance, training efficacy, and adherence to training. Openness/Intellect is defined as "the breadth, depth, originality, and complexity of an individual's experiential life (John, Naumann, & Soto, 2008, p. 120). As mentioned, this dimension is known with traits like creativity, quickness, intellect, competence, reflection, and imagination (DeYoung et al., 2007; Vartanian et al., 2018). All of these traits are necessary to explore information, keep an open-mind, find creative and flexible solutions for the problems that might occur during the interview. In addition, the previous research consistently found that high scorers in Openness/Intellect are more likely to benefit from job training in various professional groups (Barrick & Mount, 1991; Ziegler et al., 2014), and our finding contributed to this literature from an investigative interviewing point of view.

One of the strengths of the research design used in these series of studies is the usage of multiple variables to measure both interview performance and training efficacy. The four measures used to assess the investigative interview performance were the number of details elicited, researcher ratings, witness perception scores, and appropriateness of questioning. To examine the training efficacy, the percentage changes in these measures were used as dependent variables. Using multiple criteria enabled to minimize the biases and errors that could be faced when single measurement is used and also provided the opportunity to assess different aspects of interview

performance from different perspectives (e.g. rapport building, questioning, listening etc.). According to Gottfredson (1991), job performance is complex and multidimensional, therefore, multiple criteria that can measure different aspects of a job is more useful than single or composite measurement. The benefits of such an approach were seen in both Study 2 and Study 3 because the personality traits that predicted interview performance were not correlated with all measures of performance at the same time, which provided meaningful outcomes. For example, in Study 2, Agreeableness was correlated with witness perception scores and appropriate questioning whereas high scorers on Agreeableness were not found successful by the researchers. On the other hand, Extraverts and high scorers in Openness were found to be relatively successful by the researchers but did not receive high scores from the witnesses. Interestingly, the number of details elicited was not predicted by any of the personality dimensions. The reason for this might be the effect of the variation in the witness memory on the amount of information elicited. The impact of Openness on training efficacy was observed on the change in researcher ratings but not the other success measures. Taken together, this multiple-criteria approach allowed to see the nature of the personality effect more closely and differentiate the specific effects from each other. For instance, the witness perceptions and researcher ratings did not measure the exact same skills. The rapport building skills and attitudes of the interviewer towards the witness can be better understood from witness' perspective, while the questioning style can be evaluated more objectively by the researchers.

One major limitation of the current series of studies is using only witness interview performances of the interviewers as dependent variable. The individual differences that predict the success in investigative interviewing might vary based on the type of interview. Even though DeFruyt et al. (2006) used the self-evaluated imaginary performance of police officers, their

method of performance assessment addressed this limitation. They asked the officers to rate their imaginary performance across different types of interviews (i.e. suspect, witness, and victim interviews). In response to these questionnaires, Benevolent officers, for example, rated themselves as more successful in victim and witness interviews but not in suspect interviews. On the other hand, Dominant-Insisting dimension was found positively correlated with suspect interview performance and negatively correlated with victim interview performance. Future research can use real-life interview videotapes to assess the interviewers' performances in the three different types of interviews that were mentioned and identify which individual differences predict success in which type of interview. An alternative way to assess suspect interview performance in relation to individual differences might be using the cheating paradigm of Russano, Meissner, Narchet, and Kassin (2005). In this paradigm, participants are asked to complete an experimental task with a confederate and they are induced to either cheat or not cheat by the confederate. Then, they are interrogated by researchers to elicit true or false confession depending on the guilt condition (i.e. guilty or innocent). Through this paradigm, participants who cheated can be considered as suspects and interviewed by the other participants. Then, the relationship between the performance of the interviewers and their personality measures could be analyzed.

The findings of this dissertation have significant implications for law enforcement policy and future research on investigative interviewing. Together with other assignment criteria, such as experience and education, high scores on the Openness, Agreeableness, and Extraversion dimensions of the Big Five scale and the Communicative-Insisting dimension of the aptitudes scale can be used as criteria when selecting the right personnel as investigative interviewers. High scorers in the Openness, Agreeableness, and Extraversion dimension of the Big Five scale and the Humane, Careful-Tenacious and Communicative-Insisting dimensions of the aptitudes scale can

be provided with more in-depth training as they are more likely to benefit from interview training. Being equipped with evidence-based interview techniques, those officers can serve as mentors or supervisors to new recruits who are potentially planned as future interviewers in the police force. On the other hand, training efforts should focus more to improve the weaknesses of low scorers in the Openness, Agreeableness, and Extraversion dimensions of the Big Five scale and the Communicative-Insisting, Careful-Tenacious, and Humane dimensions of the aptitudes scale.

Future research

The findings in these series of studies are not conclusive. Future research can address the limitations of these studies such as the limited external validity of the sample (i.e. student participants), usage of mock crime scenarios, the length of the training course, the small sample size used to assess personality-training effect relationship, the lack of different interview types (i.e. suspect and victim interviews), and the assessment of training effect only in the short run (i.e. within the first two weeks).

Future research can address these limitations through the following recommendations:

- Actual police officers can participate in the experiments as interviewers if a lab-based design similar to this dissertation is used.
- Alternatively, the videotapes of real life interviews conducted by police officers can be analyzed. However, such an analysis might be affected by the variation in the criminal events investigated by different officers.
- The analysis of real-life interviews would also allow the analysis of different types of interviews such as victim and suspect interviews to be conducted.

- The sample size in Study 2 was sufficient; however, Study 3 was conducted with a smaller sample. Larger samples can be used to analyze the personality-training effect relationship in future research.
- The effect of training and its relationship with personality can be assessed through a more extensive training program.
- The training effect can be assessed in a longitudinal research design which will allow to compare the long run and short run effects and how these effects are related to individual differences.

Conclusion

This three-step study indicated that individual differences played a moderate role in the performance of investigative interviewers and the efficacy of training on investigative interviewing. Personality traits which are broadly defined, such as Five Factor Model (FFM), predicted the interview performance and training efficacy better than specific aptitudes did. The aptitudes scale which is a modified and extended version of the Police Interviewing Competences Inventory (PICI) consisted of four dimensions which were named as Humane, Communicative-Insisting, Self-controlled, and Careful-Tenacious. The findings of Study 2 indicated that three dimensions of the FFM (i.e., Agreeableness, Extraversion, and Openness) and only one dimension of the aptitudes scale (i.e., Communicative-Insisting) were correlated with the interview success of the student sample. Study 3 showed that the efficacy of a short training on investigative interviewing performance was correlated with the Humane dimension of the aptitudes scale and the Openness/Intellect dimension of the FFM. The post-interview performance of the participants was predicted by Openness/Intellect, Agreeableness, Extraversion, and Neuroticism (negatively)

dimensions of the Big Five and Careful-Tenacious and Communicative-Insisting dimensions of the aptitudes scale.

References

Aldridge, J. & Cameron, S. (1999). Interviewing child witnesses: Questioning techniques and the role of training. *Applied Developmental Science, 3*, 136–147.

Ames, D., Maissen, L. B., & Brockner, J. (2012). The role of listening in interpersonal influence. *Journal of Research in Personality, 46*(3), 345–349.

Baldwin, J. (1993). Police interview techniques; establishing truth or proof? *British Journal of Criminology, 33(3)*, 325-352.

Barrick, M. R., & Mount, M. K. (1991). The Big Five Personality Dimensions And Job Performance: A Meta-Analysis. *Personnel Psychology, 44*(1), 1–26.

Barrick M. R., Mount M. K., & Judge T. A. (2003). Personality and Performance at the Beginning of the New Millennium: What Do We Know and Where Do We Go Next? *International Journal of Selection and Assessment, 9*(1-2), 9–30.

Berresheim, A., & Weber, A. (2003). Structured witness interviewing and its effectiveness. *Kriminalistik, 57*, 757–71.

Black, J. (2000). Personality testing and police selection: Utility of the "Big Five." *New Zealand Journal of Psychology, 35*, 353-374.

Bull, R. (2013). What is 'believed' or actually 'known' about characteristics that may contribute to being a good/effective interviewer?. *Investigative Interviewing: Research and Practice, 5.* 128-143.

Bull, R., & Cherryman, J. (1996). *Helping to identify skills gaps in specialist investigative interviewing: Enhancement of professional skills*. London: Home Office Police Department.

Carretta, T. R., Teachout, M. S., Ree, M. J., Barto, E. L., King, R. E., & Michaels, C. F. (2014). Consistency of the Relations of Cognitive Ability and Personality Traits to Pilot Training Performance. *International Journal of Aviation Psychology*, *24*(4), 247–264. https://doi-org.uproxy.library.dc-uoit.ca/10.1080/10508414.2014.949200

Dean, M. A., Conte, J. M., & Blankenhorn, T. R. (2006). Examination of the predictive validity of Big Five personality dimensions across training performance criteria. *Personality and Individual Differences*, *41*(7), 1229–1239. https://doi.org/10.1016/j.paid.2006.04.020

Cattell, R. B., Cattell, A. K., & Cattell, H. E. P. (1993). *Sixteen Personality Factor Questionnaire* (5th ed.). Champaign, IL: Institute for Personality and Ability Testing, Inc.

Cederborg, A. C., Orbach, Y., Sternberg, K. J. & Lamb, M. E. (2000). Investigative interviews of child witnesses in Sweden. *Child Abuse and Neglect*, *24*, 1355–61.

Cherryman, J., & Bull, R. (2001). Police Officers' Perceptions of Specialist Investigative Interviewing Skills. *International Journal of Police Science & Management*, *3*(3), 199–212. https://doi.org/10.1177/146135570100300302

Clark, W. (2016, July 13). Keep an Open Mind During Interviews & Investigations. Retrieved August 26, 2016, from https://www.linkedin.com/pulse/keep-open-mind-during-interviews-investigations-wesley-clark

Clarke, C., Milne, R., & Bull, R. (2011). Interviewing suspects of crime: the impact of PEACE training, supervision and the presence of a legal advisor. *Journal of Investigative Psychology and Offender Profiling*, *8*(2), 149–162.

Clifford, B.R., & George, R. (1996) A field evaluation of training in three methods of witness/victim investigative interviewing. *Psychol Crime Law, 2*(3), 231–248. doi:10.1080/10683169608409780

Costa Jr, P. T., & McCrae, R. R. (1992). The five-factor model of personality and its relevance to personality disorders. *Journal of Personality Disorders*, 6(4), 343-359.

Costa, P. T., Jr., McCrae, R. R., & Dye, D. A. (1991). Facet scales for Agreeableness and Conscientiousness: A revision of the NEO Personality Inventory. *Personality and Individual Differences, 12,* 887–898.

Cucina, J.M., Vasilopoulos, N.L., & Sehgal, K. (2005). *Journal of Business and Psychology, 20*, 275-290.

Dando, C., Wilcock, R., & Milne, R. (2008). Victims and witnesses of crime: police officers' perceptions of interviewing practices. *Leg Criminol Psychol, 13*, 59–70. doi:10.1348/135532506X162498

DeFruyt, F., Bockstaele, M., Taris, R., & Van Hiel, A. (2006). Police interview competences: assessment and associated traits. *European Journal of Personality*. *20* (7). 567-584

Detrick, P., & Chibnall, J. T. (2006). NEO PI-R personality characteristics of high-performing entry-level police officers. *Psychological Services, 3*(4), 274-285.

Duke, M. C. (2013). *The development of the rapport scales for investigative interviews and interrogations* (Unpublished doctoral dissertation). The University of Texas at El Paso, El Paso, TX.

Fahsing, I. A., & Rachlew, A. (2009). Investigative interviewing in the Nordic region, in T. Williamson, B. Milne and S. P. Savage (eds), *International Developments in Investigative Interviewing*. Cullompton: Willan, pp. 39–65.

Fisher, R.P., Ross, S.J., Cahill, B.S. (2010). Interviewing witnesses and victims. In G. P. Anders (Ed). Forensic psychology in context: Nordic and international approaches (pp. 56-74). Devon, United Kingdom: Willan Publishing, xviii, 334 pp.

Forero, C., Gallardo-Pujol, D., Maydeu-Olivares, A., & Andrés-Pueyo, A. (2009). A Longitudinal Model for Predicting Performance of Police Officers Using Personality and Behavioral Data. *Criminal Justice and Behavior*, *36*(6), 591–606. http://doi.org/10.1177/0093854809333406

Griffiths, A., Milne, B., & Cherryman, J. (2011). A Question of Control? The Formulation of Suspect and Witness Interview Question Strategies by Advanced Interviewers. *International Journal of Police Science & Management*, *13*(3), 255–267. https://doi.org/10.1350/ijps.2011.13.3.219

Goldberg, L. R. (1993). The structure of phenotypic personality traits. *The American Psychologist*, *48*(1), 26–34.

Gottfredson, L. S. (1991). The evaluation of alternative measures of job performance. In *Performance assessment for the workplace, Vol. 1; Vol. 2: Technical issues* (pp. 75–126). Washington, DC, US: National Academy Press.

Hodgson, P. (1987). A practical guide to successful interviewing. McGraw-Hill: Maidenhead.

Hoekstra, H. A., & Van Sluijs, E. (2003). Managing competences: Implementing Human Resource Management. Nijmegen, The Netherlands: GITP International.

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative Big Five trait taxonomy. In O. P. John, R. W. Robbins, &L. A. Pervin (Eds.), Personality handbook: Theory and research (pp.114–158). New York: Guilford.

Judge, T. A., Rodell, J. B., Klinger, R. L., Simon, L. S., & Crawford, E. R. (2013). Hierarchical representations of the five-factor model of personality in predicting job performance: integrating three organizing frameworks with two theoretical perspectives. *The Journal of Applied Psychology*, *98*(6), 875–925.

Kassin, S. M., & Kiechel, K. L. (1996). The Social Psychology of False Confessions: Compliance, Internalization, and Confabulation. *Psychological Science*, *7*(3), 125–128. Retrieved from JSTOR.

Kebbell, M., Hurren, E., & Mozerelle, P. (2006). An investigation into the effective and ethical interviewing of suspected sex offenders. Final report to the Australian Criminology Research Council.

Kebbell, M., Alison, L., Hurren, E., & Mazerolle, P. (2010). How do sex offenders think the police should interview to elicit confessions from sex offenders? *Psychology, Crime & Law*, *16*(7), 567–584. https://doi.org/10.1080/10683160902971055

Kelly, C. E., Abdel-Salam, S., Miller, J. C., & Redlich, A.D. (2015). Social identity and the perceived effectiveness of interrogation methods. *Investigative Interviewing: Research and Practice, 7(*2). 24-42.

Köhnken, G., Milne, R., Memon A, & Bull, R. (1999). The cognitive interview: a meta-analysis. Psychol Crime Law, 5(1-2), 3–27. doi:10.1080/10683169908414991

Krug, S. E. (1980). *Clinical Analysis Questionnaire manual*. Champaign, IL: Institute for Personality and Counseling.

Kurke, M. I., & Gettys, V.S. (1995). Human Factors Psychology for Law Enforcement Agencies. In M. I. Kurke & E. M. Scrivner (Eds.), *Police Psychology into the 21st Century* (pp. 467-496). New York: Psychology Press.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159-174. doi:10.2307/2529310

McCrae, R. R. & John, O. P. (1992). An Introduction to the Five-Factor Model and Its Applications. *Journal of Personality, 60*, 175–215.

Mandrekar, J. N. (2011). Measures of Interrater Agreement. *Journal of Thoracic Oncology, 6* (1), 6-7.

Meissner, C., Redlich, A., Michael, S., Evans, J., Camilletti, C., Bhatt, S., & Brandon, S. (2014). Accusatorial and information-gathering interrogation methods and their effects on true and false confessions: a meta-analytic review. *Journal of Experimental Criminology, 10*(4), 459–486.

Memon, A., Holley, A., Milne, R., Koehnken, G., & Bull, R. (1994). Towards understanding the effects of interviewer training in evaluating the cognitive interview. *Applied Cognitive Psychology, 8*(7), 641-659.

Memon, A., Meissner, C.A., Fraser, J. (2010). The cognitive interview: a meta-analytic review and study space analysis of the past 25 years. *Psychol Public Policy Law, 16*, 340–372. doi:10.1037/a0020518

Milne, R. & Bull, R. (1999). *Investigative Interviewing: Psychology and Practice*. Chichester: Wiley.

McGurk, B.J., Carr, M.J. & McGurk, D. (1993). Investigative interviewing courses for police officers: an evaluation. Police Research Series: Paper No.4. Home Office: London.

Moston, S., Stephenson, G. M., & Williamson, T. M. (1992). The effects of case characteristics on suspect behaviour during police questioning. *British Journal of Criminology*, *32*(1), 23–40.

Murphy, K. R. & Dzieweczynski, J.L. (2005). Why don't measures of broad dimensions of personality perform better as predictors of job performance? *Human Performance, 18*(4), 343–357.

O'Neill, M. & Milne, B. (2014). Success Within Criminal Investigations: Is Communication Still a Key Component? In R. Bull (Ed), *Investigative Interviewing*. New York: Springer-Verlag.

O'Connor, T., & Carson, W. (2005). Understanding the psychology of child molesters: A key to getting confessions. The Police Chief, 72, December, 1-7.

Ono, M., Sachau, D. A., Deal, W. P., Englert, D. R., & Taylor, M. D. (2011). Cognitive Ability, Emotional Intelligence, and the Big Five Personality Dimensions as Predictors of Criminal Investigator Performance. *Criminal Justice and Behavior, 38*(5), 471–491.

Ozer, D. J., & Benet-Martínez, V. (2005). Personality and the Prediction of Consequential Outcomes. *Annual Review of Psychology*, *57*(1), 401–421. https://doi.org/10.1146/annurev.psych.57.102904.190127

Pearse, J., & Gudjonsson, G. H. (1999). Measuring influential police interviewing tactics: A factor analytic approach. *Legal and Criminological Psychology*, *4*(2), 221–238. http://doi.org/10.1348/135532599167860

Pitts, M. J., & Miller-Day, M. (2007). Upward turning points and positive rapport-development across time in researcher—participant relationships. *Qualitative Research*, *7*(2), 177–201. https://doi.org/10.1177/1468794107071409

Posten, H. O. (1984). Robustness of the Two-Sample T-Test. In D. Rasch & M. L. Tiku (Eds.), *Robustness of Statistical Methods and Nonparametric Statistics* (pp. 92–99). https://doi.org/10.1007/978-94-009-6528-7_23

Prien, K. O., Prien, E. P., & Wooten, W. (2003). Interrater Reliability in Job Analysis: Differences in Strategy and Perspective. *Public Personnel Management*, *32*(1), 125–141. https://doi.org/10.1177/009102600303200107

Russano, M. B., Meissner, C. A., Narchet, F. M., & Kassin, S. M. (2005). Investigating true and false confessions within a novel experimental paradigm. *Psychological Science*, *16*(6), 481–486. https://doi.org/10.1111/j.0956-7976.2005.01560.x

Russano, M., & Narchet, F. (2013). Interrogation practices and beliefs: Does high-value target experience matter? Paper presented at the Annual Conference of the American Psychology-Law Society, Portland.

Korkman, J., Santilla, P., & Sandnabba, N. K. (2006). Dynamics of verbal interaction between interviewer and child in interviews with alleged victims of child sexual abuse. *Scandinavian Journal of Psychology*, *47*, 109–19.

Salgado, J. F. (1997). The five factor model of personality and job performance in the European Community. *Journal of Applied Psychology*, *82*(1), 30–43.

Sanchez, J. I., & Levine, E. L. (2000). Accuracy or consequential validity: which is the better standard for job analysis data? *Journal of Organizational Behavior*, *21*(7), 809–818. https://doi.org/10.1002/1099-1379(200011)21:7<809::AID-JOB28>3.0.CO;2-O

Schollum, M. (2005). Investigative Interviewing: the literature. New Zealand. Police: Wellington

Schreiber, N. & Fisher, R. P. (2005). *Police Interviewing Techniques: Types of Questions, Positive and Negative Techniques in a South Florida Sample*. Paper presented at the American Psychology – Law Society, La Jolla, CA.

Schmidt, F. L., & Hunter, J. (2004). General mental ability in the world of work: occupational attainment and job performance. *Journal of Personality and Social Psychology*, *86*(1), 162–173.

Singh, P. (2008). Job analysis for a changing workplace. *Human Resource Management Review*, *18*(2), 87–99. https://doi.org/10.1016/j.hrmr.2008.03.004

Smets, L. (2009). Reliability and correlational validity of police interview competences: Assessing the stability of the police interview competency inventory. In M. Cools, et al. (Eds.), Governance of security research papers series: 2. Readings on criminal justice, criminal law & policing (pp. 311-328). Apeldoorn-Antwerpen: Maklu.

Snook, B., Eastwood, J., Stinson, M., Tedeschini, J., & House, J. C. (2010). Reforming Investigative Interviewing in Canada. *Canadian Journal of Criminology and Criminal Justice*, *52*(2), 215–229.

Snook, B., & Keating, K. (2011). A field study of adult witness interviewing practices in a Canadian police organization. *Legal & Criminological Psychology*, *16*(1), 160–172. https://doi.org/10.1348/135532510X497258

St-Yves, M. (2009). Police interrogation in Canada: From the quest for confession to the search for the truth. In T. Williamson, B. Milne, and P. Savage (Eds.), *International Developments in Investigative Interviewing* (pp.92-110). Cullompton: Willan.

St-Yves, M., Griffiths, A., Cyr, M., Gabbert, F., Carmans, M., Sellie, C., … Powell, M. (2014). Training in Investigative Interviewing: observations and challenges. In M. St-Yves & A. Griffiths (Eds.), *Investigative Interviewing: The Essential Handbook of best practices* (pp. 245–269). Retrieved from http://research.gold.ac.uk/10672/

Swanson, C.R., Chamelin, N.C. & Territo, L. (2002). Criminal Investigation. McGraw-Hill Higher Education: United States.

Tett, R. P., Jackson, D. N., & Rothstein, M. (1991). Personality Measures as Predictors of Job Performance: A Meta-Analytic Review. *Personnel Psychology*, *44*, 703–742.

Thwaites, R. (2017). (Re)Examining the Feminist Interview: Rapport, Gender "Matching," and Emotional Labour. *Frontiers in Sociology*, *2*. https://doi.org/10.3389/fsoc.2017.00018

Thomas, D. J. (2011). *Police Psychology: A New Specialty and New Challenges for Men and Women in Blue*. ABC-CLIO.

Tracey, W. R. (2016). The Human Resources Glossary, Third Edition: The Complete Desk Reference for HR Executives, Managers, and Practitioners. CRC Press.

Varela, J. G., Boccaccini, M. T., Scogin, F., Stump, J., & Caputo, A. (2004). Personality Testing in Law Enforcement Employment Settings: A Metaanalytic Review. *Criminal Justice and Behavior*, *31*(6), 649–675. https://doi.org/10.1177/0093854804268746

Vartanian, O., Wertz, C. J., Flores, R. A., Beatty, E. L., Smith, I., Blackler, K., … Jung, R. E. (2018). Structural correlates of Openness and Intellect: Implications for the contribution of personality to creativity. *Human Brain Mapping*, *39*(7), 2987–2996. https://doi.org/10.1002/hbm.24054

Villalba, D. (2014). The Effect of Rapport Building in Police Interrogations: Can Rapport Improve the Diagnosticity of Confession Evidence? *FIU Electronic Theses and Dissertations*. https://doi.org/10.25148/etd.FI14110772

Wachi, T., Watanabe, K., Yokota, K., Otsuka, Y., & Lamb, M. E. (2016). The relationship between police officers' personalities and interviewing styles. *Personality and Individual Differences*, *97*, 151–156. https://doi.org/10.1016/j.paid.2016.03.031

Warren, A. R., Woodall, C. E., Thomas, M., Nunno, M., Keeney, J. M., Larson, S. M., & Stadfeld, J. A. (1999). Assessing the effectiveness of a training program for interviewing child witnesses. *Applied Developmental Science*, *3*, 128-135.

Wicklander, D.E. & Zulawski, D.E. (2003). *Interview and interrogation techniques - a training course*. Illinois: Wicklander -Zulawski and Associates Inc.

Yeschke, C. L. (2003). The art of investigative interviewing: a human approach to testimonial evidence. Boston: Butterworth-Heinemann.

Ziegler, M., Bensch, D., Maaß, U., Schult, V., Vogel, M., & Bühner, M. (2014). Big Five facets as predictor of job training performance: The role of specific job demands. *Learning and Individual Differences*, *29*, 1–7. https://doi.org/10.1016/j.lindif.2013.10.008

Appendix A: Witness Perception Scale

Witness Perception Questionnaire (Adopted from Duke (2013) and Villalba (2014))						
Read the following statements and mark the appropriate choice by putting (X). (1=Strongly disagree 5=Strongly agree)		1	2	3	4	5
1	The interviewer seems like a good-natured person.					
2	The interviewer seems to have a compassionate attitude towards people.					
3	The interviewer seems to be generally sincere.					
4	The interviewer was friendly towards me.					
5	The interviewer treated me with respect					
6	The interviewer paid careful attention to what I had to say.					
7	I felt obliged to answer the interviewer's question.					
8	The interviewer acted with calmness.					
9	The interviewer showed a personal interest and tried to create a personal conversation.					
10	The interviewer was attentive to me.					
11	The interviewer thinks highly of me.					
12	The interviewer thinks that I am a good person.					
13	The interviewer values my opinion.					
14	The interviewer respects my intelligence.					
15	The interviewer was honest with me.					
16	The interviewer was impatient.					
17	The interviewer had good intentions towards me.					

18	The interviewer has empathy.					
19	The interviewer doesn't care about what happens to me.					
20	The interviewer showed a formal and non-accessible attitude					
21	I can't trust the interviewer to tell me the truth.					
22	The interviewer acted as a trouble-shooter.					
23	The interviewer showed a condemning attitude against me.					
24	During the interview I felt like the interviewer and I understood each other well.					
25	I would be willing to do another interview with the interviewer.					
26	The interviewer helped me to feel comfortable enough to share information during the interview.					
27	The interviewer and I got along well during the interview.					
28	Communication went smoothly between the interviewer and me.					
29	I connected with the interviewer in a positive way.					
30	I was motivated to help the interviewer.					
31	I was cooperative during the interview.					
32	I was honest with the interviewer					
33	I became frightened.					
34	I became aggressive.					
35	I became paralysed.					
36	I was under stress.					
37	I felt confident during the interview?					
38	I felt insulted as a human being.					

39	I became obliging.					
40	I felt acknowledged as a human being.					
41	The interviewer allowed me to provide all the information about the criminal event					
42	The interviewer gave me time to fully answer the questions.					
43	The interviewer was thorough with his/her questioning.					
44	The interviewer was successful in accessing all the information about the event that was in my memory					

Appendix B: Researcher Rating Scale

1. Interview length (minutes)

_____

2. Overall interview outcome

○ Brief account

○ Partial account

○ Comprehensive account

3. Introduce self

○ Yes

○ No

4. Explains interview purpose

	5	4	3	2	1	
Clear&Professional	○	○	○	○	○	Not explained

5. Provides details of routine and route map

	5	4	3	2	1	
Clear&Professional	○	○	○	○	○	Not given

6. Evidence of Rapport Building

	5	4	3	2	1	
Excellent	○	○	○	○	○	No rapport attempted

PERSONALITY AND INTERVIEWING

7. Encourages interviewee to give an uninterrupted account

	5	4	3	2	1	
Frequently	○	○	○	○	○	Never

8. Development of topics

	5	4	3	2	1	
Apparent	○	○	○	○	○	Not apparent

9. Dealing with difficulty

	5	4	3	2	1	
Maintains interview	○	○	○	○	○	Flustered/Early Close

10. Appropriate Structure/ witness led sequencing

	5	4	3	2	1	
Always	○	○	○	○	○	Never

11. Appropriate use of questions

	5	4	3	2	1	
Always	○	○	○	○	○	Never

12. Keeps interview to relevant topics

	5	4	3	2	1	
Always (1)	○	○	○	○	○	Never

13. Summaries and links

	5	4	3	2	1	
Always	○	○	○	○	○	Never

14. Covers points to prove

	5	4	3	2	1	
All points	○	○	○	○	○	None

15. Clarification

	5	4	3	2	1	
Problem solving	○	○	○	○	○	Confrontational

16. Exploration of information

	5	4	3	2	1	
Apparent	○	○	○	○	○	Not apparent

17. Use of pauses and silence

	5	4	3	2	1	
Apparent	○	○	○	○	○	Not apparent

PERSONALITY AND INTERVIEWING

To what extent did the interviewer demonstrate the following behaviour

18. Self confidence

	5	4	3	2	1	
Clear&Confident	○	○	○	○	○	Nervous&Unsure

19. Open mind

	5	4	3	2	1	
listens to and explores explanation	○	○	○	○	○	persists with own view

20. Flexibility

	5	4	3	2	1	
Responds to new information	○	○	○	○	○	Persists with own view

21. Communication skills

	5	4	3	2	1	
Clear appropriate speech	○	○	○	○	○	Complex/ unclear speech

22. Active listening

	5	4	3	2	1	
Always	○	○	○	○	○	Never

23. Summarizes Interview

	5	4	3	2	1	

Comprehensive ○ ○ ○ ○ ○ None

24. Invites interviewee to Add/Alter or Correct

○ Yes

○ No

Appendix C: Aptitudes Scale (Adopted from DeFruyt et al., 2006)

	Characteristically I	**1** Hardly characteristic	**2** Barely characteristic	**3** More or less characteristic	**4** Characteristic	**5** Very characteristic
1.	...am patient					
2.	...am assertive					
3.	...am calm					
4.	…have good intuition					
5.	...am understanding					
6.	...am communicative					
7.	...am empathic					
8.	…pay attention to details					
9.	…am careful					
10.	…am honest					
11.	…am sincere					
12.	…am sensitive and understanding					
13.	…am thick-skinned					
14.	…take action					
15.	...am tender					
16.	…am flexible					
17.	…am confident					
18.	…am open-minded					
19.	…am polite					
20.	…am even-tempered					
21.	..have a firm tone of voice					
22.	…am tactful					

	Characteristically I	**1** Hardly characteristic	**2** Barely characteristic	**3** More or less characteristic	**4** Characteristic	**5** Very characteristic
23.	…am tenacious					
24.	…have the ability to calm other people					
25.	...am persuasive					
26.	…work with relentless zeal					
27.	...am able to act gently					
28.	…have sympathy					
29.	...am able to keep my head cool					
30.	…have perseverance					
31.	...am quiet					
32.	…having a tongue of my own					
33.	...am rigid					
34.	...am quick to understand					
35.	...am complaisant					
36.	...am authoritative					
37.	…maintain concentration					
38.	...am self-controlled					
39.	...am fluent in social contacts					
40.	... am aware of my own thoughts and actions					
41.	…am persistent (stick to my task)					
42.	...am able to put things in perspective					

	Characteristically I	**1** Hardly characteristic	**2** Barely characteristic	**3** More or less characteristic	**4** Characteristic	**5** Very characteristic
43.	...am talkative					
44.	…am thorough					
45.	…am driven					
46.	...am able to handle pressure					
47.	...am offensive					
48.	...am good-hearted					
49.	...am tongue-tied					
50.	...am able to respond quickly and appropriately					

Appendix D: Demographics Questionnaire

1. What is your gender?

- Male
- Female

2. How old are you?

_____

* 3. What is your ethnicity?

- White
- Black
- Indigenous person
- Asian
- Arab
- Latin American
- Other

4. Is English your first language?

- Yes
- No

5. In which program are you studying currently?

_____

6. In what year of the program are you currently in?

- 1st year
- 2nd year
- 3rd year
- 4th year
- 5th or more

7. Are you an international student?

- Yes
- No

8. Have you ever received any training, course, or seminar on Interviewing?

- Yes
- No

If yes, please specify: