Unproctored Internet Testing in Personnel Selection:
A Test of Cheating Deterrence Strategies

by

Emrah Eren

A thesis submitted to the Faculty of Graduate and Postdoctoral
Affairs in partial fulfillment of the requirements
for the degree of

Master of Arts

in

Psychology

Carleton University
Ottawa, Ontario

© 2018

Emrah Eren

Abstract

Unproctored Internet testing (UIT) has become increasingly popular in the context of personnel recruitment and selection. Despite the widespread use of UIT, its fairness and overall impact on the hiring process is still called into question (Ryan & Ployhart, 2014). In consideration of the issues surrounding UIT, the purpose of this thesis was to examine job applicant cheating behaviour. Job candidates ($n=110$) were recruited through a series of job advertisements. Mild deception was used to cover up the true purpose of the research. Participants were administered a cognitive ability and cheating detection test as part of the regular recruitment process, and upon completion were invited to complete a follow-up survey. Contrary to popular practice, cheating deterrence strategies in the form of warning messages did not prevent candidate cheating. Furthermore, cheating was associated with overall lower cognitive ability. These results are discussed in relation to the Theory of Planned Behaviour, as well as the implications for future research.

Acknowledgments

I would like to begin by thanking my thesis supervisor, Dr. Pychyl, for not only believing in me the first time but for taking me under his wing again for when I returned to Carleton to complete my Master's degree. This has been a long journey and he was there as my support system from day one. The countless discussions we had together over the past year, and the expert, structured feedback I received from him made the completion of my thesis possible. I have learned and grown under your guidance and the competencies I have developed during this process have been added to my skill set as invaluable learning.

Another big thank you to Etelle Bourassa who also played a big role in my return to Carleton and encouraged me to complete my program. I would also like to thank my committee members, Dr. Campbell and Dr. Sears. Your valuable insight and feedback contributed to an even stronger Master's thesis. Thank you all for taking the time to be a part of this journey.

I would also like to extend my appreciation and complete gratitude to Suzanne Simpson and Lorraine McKay (my current employers of over 5 years) for believing in my research topic and providing me with the support I needed to return to school and complete my degree. Thank you for seeing the value in the work I was trying to accomplish. Another person I would like to acknowledge is Shamarukh Chowdhury for her numerous reviews and strong guidance on my statistical approach - this was a key contribution to the overall quality of my work and I am forever thankful for the time and effort she put into helping me bring my vision into fruition. I also want to thank my good friend Colin Kemp for helping me set up my data files - one of my first challenges on this journey which he helped me overcome and was essential in this important and foundational work. Another big thank you to Glen Howell for reviewing my statistical

Unproctored Internet Testing

analysis section - Glen's attention to detail and his extensive statistical knowledge was instrumental in the precision of the work I was able to achieve.

I will also always be grateful to my parents and sister for encouraging and supporting me to go back to school even though I felt it was too late - thank you for believing in me and supporting me throughout this journey.

And above all, I would like to thank my beautiful wife, Masumé, without whom none of this would have been possible. I want to express my profound gratitude for her unwavering support from day one, and how she inspired me to return to school and achieve this important milestone. I know that the countless hours she took care of our life and our baby girl while I was plugged to my computer working on my thesis paper and defence was invaluable to my success and the single most important reason I was able to complete this journey and achieve my goals.

Table of Contents

Acknowledgments.....	2
Table of Contents.....	4
List of Tables	6
List of Figures	7
List of Appendices	8
UIT – Introduction and Background.....	10
UIT – Advantages.....	12
UIT – Issues and Concerns	13
UIT – Cheating	15
Academic Cheating Research	20
Definition and Prevalence.....	20
Theory of Planned Behaviour	21
Theory of Planned Behaviour and Cheating.....	23
Contextual Factors	26
Individual Differences	28
Implications for UIT	29
Cheating and Response Distortion – How do they differ?.....	31
UIT – Common Types of Cheating and Prevention Methods.....	35
UIT – Cheating Detection.....	37
Literature Review - Conclusion	38
Hypotheses Based on Current Review	39
Methods	40
Participants.....	40
Procedure	42
Measures and Materials	45
Results.....	49
Preliminary Analysis.....	49

Unproctored Internet Testing

Main Analyses and Hypothesis Testing.....	53
Discussion.....	58
References.....	65
Appendices.....	77

List of Tables

Table 1. UIT Issues and Concerns	13
Table 2. Common UIT Cheating Methods: Prevention and Detection.....	35
Table 3. Sample Characteristics: Means, Frequencies, and Standard Deviations	41
Table 4. Scale Descriptive and Standardized Skewness and Kurtosis Statistics	52
Table 5. Frequency of Cheating Behaviour across the Experimental Conditions	54
Table 6. Frequency of Cheating Behaviour across the Time Since Graduation Groups	55

List of Figures

Figure 1. The impact of different testing conditions on cheating risk. (Figure retrieved from Landers & Sackett, 2012)	17
Figure 2. Antecedents and mitigators of cheating on unproctored employment tests (figure retrieved from Kantrowitz & Dainis, 2014).....	19
Figure 3. Modified TPB Framework for College Cheating. (figure retrieved from Mayhew et al. (2009)).....	26
Figure 4. Model 1: Consistency Testing	37
Figure 5. Model 2: Successive Hurdles	37
Figure 6. Visual Diagram of Study Procedure.....	44

List of Appendices

Appendix A.....	77
Appendix B.....	79
Appendix C.....	82
Appendix D.....	86
Appendix E.....	88
Appendix F.....	89
Appendix G.....	91
Appendix H.....	92
Appendix I.....	95
Appendix J.....	101
Appendix K.....	103

Unproctored Internet Testing in Personnel Selection: A Test of Cheating Deterrence Strategies

Unproctored Internet testing (UIT) occurs when an exam is administered remotely through the Internet without a traditional human proctor overseeing the test taker and administration process (Tippins, 2009a). UIT is becoming increasingly popular as technology continues to shape the way in which online exams are designed and delivered. Pearlman (2009) estimates that UIT accounts for the majority of personnel tests administered in the U.S. private sector. More recent research from CEB Global suggests that 83% of employment tests are administered through UIT (Kantrowitz & Dainis, 2014).

Major U.S. corporations such as Cisco (Guo & Drasgow, 2010) and Bank of America (Hense, Golden, & Burnett, 2009) administer (or have administered in the past) unproctored Internet tests to support their talent acquisition and management processes. In fact, Cisco went as far as delivering their professional certification through UIT, leading to a nation-wide scandal (see Guo & Drasgow, 2010)¹. UIT is also commonly deployed as a test delivery strategy in Canada, especially in high volume hiring programs. For example, Canada's Public Service leverages the efficiency and convenience of UIT in the initial stages of its recruitment process (see Government of Canada, 2017), although usage statistics in the U.S. suggest higher prevalence in the private sector (Nesnidol, 2016).

UIT is an area of concern for I/O psychologists because of the demand from organizations to use unproctored employment tests for talent acquisition and management. Organizations want to know whether they *should* use UIT, and research has yet to produce conclusive evidence for its fairness and legal defensibility. As some researchers have pointed

¹ <http://www.livemint.com/Companies/zaPzFn6aIL3R3YhO7LWpGN/Study-confirms-widespread-cheating-on-job-examinations.html>

out, however, the question is not whether we *should* be *using* UIT but *how* we should be using it more effectively (Kantrowitz & Dainis, 2014; Ryan & Ployhart, 2014). In other words, the ship has sailed. All types of organizations are now using UIT, yet there are real challenges to this use such as test security, perceptions of fairness, as well as cheating and technology use, to which practitioners and researchers should turn their attention. In consideration of the issues surrounding UIT (see Ryan & Ployhart, 2014), the purpose of this thesis was to examine applicant cheating behaviour in the context of an actual selection program that utilizes UIT.

My thesis begins with an overview of UIT. In this section, I explain how and why UIT became popular, its main advantages, and the issues and concerns that threaten test validity (e.g., cheating). Next, I present and summarize the extant literature on academic cheating and the implications of this research to UIT. I then present the common types of cheating and cheating detection methods used in UIT and discuss the application of UIT to pre-employment testing. Finally, based on this literature review, I present the current study hypotheses, method, and results, which investigates the relations between applicant cheating behaviour and cheating deterrence strategies in the context of personnel selection and UIT.

UIT – Introduction and Background.

Unproctored Internet Testing (UIT) is the result of a decades-long evolution in the testing and assessment industry (Arthur, Glaze, Villado, & Taylor, 2010; Burke & Normand, 1987; Ryan & Ployhart, 2014). While early research focused on the efficacy and validity of computer-based psychological testing (see Burke & Normand, 1987; Kleinmuntz & McLean, 1968), research between the years of 1990 to 2000 focused on the validity of computer-based testing for educational and employment purposes (Schmitt, Gilliland, Landis, & Devine, 1993). Internet-based testing emerged as a natural extension of computer-based testing (Lievens & Harris,

2003), which out of convenience to test takers and cost savings for administrators incorporated unproctored Internet testing into its delivery platform. It is probable that UIT caught the attention of I/O researchers after major organizations who had access to the required technology and infrastructure started using UIT in their personnel selection processes. It is clear from the literature review conducted for this thesis that UIT did not emerge in the context of organizational or business psychology research. Initial reactions ranged from researchers who cautioned against the use of UIT, for example Greenberg (1999), to Baron and Austin (2000; as cited in (Lievens & Harris, 2003), who developed an unproctored cognitive ability test which they administered using item response theory principles to prevent cheating. Studies looking at the differences between proctored vs. unproctored testing did not emerge until Tippins, Beaty, Drasgow, Gibson, Pearlman, and Segall's (2006) paper - the first modern critique of UIT. Tippins et al. (2006) formed a panel of six experts to discuss the issues related to UIT and came to consensus on several items. Summarized below are the key issues the panel agreed on. For the full list of consensus items see Tippins et al., (2006):

1. Type of test (e.g., personality vs cognitive) will determine whether UIT should be used.
2. UIT alone should never be used for high-stakes examinations.
3. Verification testing should be used in high-stakes testing environments.
4. There is disagreement among I/O psychologists over the validity of inferences based on scores from UIT exams.
5. Cheating is likely when the stakes are high.
6. UIT will lead to test/question exposure.

While several advances have been made since Tippins et al. (2006) explored the issues surrounding UIT (e.g., virtual proctoring techniques, biodata verification, and test validity research), researchers have not produced *conclusive* evidence that UIT is fair and ethical in the

context of personnel selection. In the next two sections, I explore in greater detail some of the key advantages, as well as the key issues and concerns surrounding UIT.

UIT – Advantages

Regardless of the risks associated with UIT, researchers and HR practitioners have noted several advantages for the test-taker and administering organization. With regards to the delivery of unproctored web-based assessments, practitioners note the ease of real-time question modification, expanded coverage of psychological dimensions, test questions that are more realistic, and increased measurement accuracy (Scott & Lezotte, 2012). Driving the use of UIT from an organizational perspective are the savings associated with front-end hiring and test maintenance costs, reductions in time-to-hire, more efficient management of candidate volumes, broader and more diverse candidate pools, and the portrayal of the organization as “high-tech” (Pearlman, 2009; Reynolds, Wasko, Sinar, Raymark, & Jones, 2009; Tippins, 2009b; Tippins et al., 2006). When compared to traditional paper and pencil testing, UIT is also environment-friendly, reducing paper and CO2 emissions. There are also several advantages for test takers such as the convenience of taking a test from home, at any time, and without the restrictions of travel and scheduling conflicts. From this perspective, UIT is considered a candidate-centric test delivery method.

Although the label UIT implies the absence of a human proctor, there are several non-traditional test monitoring activities associated with UIT. For example, organizations will typically deploy alternative proctoring methods such as virtual proctoring or remote surveillance through which test administrators have the capability to monitor several testing sessions at once. Furthermore, test administrators have forged other methods such as verification testing, biodata authentication, and response pattern analysis to prevent, detect, and discourage cheating. The

Unproctored Internet Testing

main disadvantage with virtual proctoring and other remote cheating prevention methods are the costs associated with administering them. These methods are typically deployed for high-stakes exams (e.g., medical device certification) in which the test taker is in a remote geography and travel to the centre would cost even more. Personnel selection tests used for pre-employment screening rarely deploy virtual proctoring methods.

UIT – Issues and Concerns

As summarized in *Table 1* and discussed by others such as Tippins et al. (2006) and Pearlman (2009), UIT has its fair share of issues and concerns. While the primary focus of this Master's thesis is on the issue of cheating in the context of UIT, there are several other issues that affect the validity and reliability of assessments administered through UIT.

Table 1. UIT Issues and Concerns

Cheating
- Validity of assessment and score inflation
- Cheaters becoming employees
- Incompetence and unethicity
Threats to fairness perceptions of process
- Disgruntled candidates may challenge the organization
- Unpleasant feeling toward administering organization
Threat to test security
- Leakage of test content to the Internet
- Hacking of testing site to retrieve content
Lack of standardized testing condition
- No control over testing environment
- Candidates might be interrupted
Impersonal nature of remote computerized testing
- Nobody to provide guidance or answer questions
Internet connection and hardware failures
- Remote areas still have slow Internet
- Computer hardware may be outdated

Job candidates who are administered an unproctored assessment are aware that others taking the exam have the opportunity to cheat should they desire to do so. For cheaters, this is less of a concern because they are exploiting the system and putting others at an unfair

Unproctored Internet Testing

disadvantage. Perceptions of fairness are most critical for the honest test-taker. If honest test takers fail an unproctored selection test and overhear that others cheated and passed the test, then they are more likely to challenge the test administering organization. Perceptions of unfairness will also affect test takers' attitudes toward the hiring organization (Pearlman, 2009). If one considers the reality that high volume testing programs reach thousands of candidates per year, then ensuring the integrity of the testing program is a critical public relations matter for the UIT administering organization.

Test security has always been an issue for test developers and administrators, even before the introduction of UIT (see previous versions of the American Educational Research Association, Standards for Educational and Psychological Testing, 2014). Leaked test content, whether from a photocopy machine or computer server, seriously jeopardizes the integrity of testing programs. UIT has introduced a series of threats to test security because of the ease in which a test-taker can record/capture, screenshot, or share an assessment with others during or after the examination. Capable hackers can also exploit the online testing system and access the assessment and test-taker data, both serious violations to the standards for psychological testing.

Test validity is also affected by the conditions in which test takers complete an exam. Establishing validity requires consistent (i.e., reliable) testing conditions to ensure variations in test scores are a true reflection of ability and not other factors such as room lighting, size of computer screen, slow Internet connection, hardware failures, and interruptions (Lievens & Harris, 2003). While environmental concerns have always been an issue for UIT research, there is very little evidence for the effects of testing conditions (i.e., proctored in the classroom vs. unproctored at home or elsewhere) on test performance (Templer & Lange, 2008; Weiner & Morrison Jr, 2009).

On the other hand, the topic of cheating continues to dominate the UIT research agenda (Guo & Drasgow, 2010; Hense, Golden, & Burnett, 2009; Hollister & Berenson, 2009; Hylton, Levy, & Dringus, 2016; Kaminski & Hemingway, 2009; Kantrowitz & Dainis, 2014). Even though it appears researchers have accepted the reality that UIT is now an integral part of the recruitment process, especially for high volume hiring, there is some skepticism with regards to the efficacy, fairness, and ethicality of UIT. For example, research has yet to produce conclusive answers to the following questions: How often and why do job applicants cheat? How do contextual and individual difference variables relate to cheating behaviour in personnel testing? And, how effective are cheating deterrence strategies? While some of these questions have been addressed more thoroughly in the academic cheating literature (summarized below), it is crucial that cheating is examined more systematically in the UIT literature. The following section provides an overview of the UIT cheating literature.

UIT – Cheating

Cheating in the context of personnel testing is defined as “behavior from the test taker that intends to deceive the test provider by enhancing the ability or trait that the test tries to measure” (Tendeiro, Meijer, Schakel, & Maij-de Meij, 2013). In the UIT definition of cheating, response distortion to personality tests would also be considered cheating. Common methods of cheating in unproctored Internet tests include the use of a surrogate test-taker (i.e., having someone else take the test), utilizing unauthorized devices or material during the test (e.g., smartphone, calculator, textbook), and using the Internet to search for test answers (Kantrowitz & Dainis, 2014; Pearlman, 2009).

Cheating in the context of personnel selection creates negative consequences for the hiring organization, as well as their job applicant pool. Most critically, the occurrence of

Unproctored Internet Testing

cheating calls into question the fairness of the hiring process because cheaters will pass and honest test takers may be screened out of the selection process. Furthermore, score inflation due to cheating will affect the validity of the test being used, rendering the selection program liable to legal challenges (Pearlman, 2009).

Test administrators deliver a variety of different exams through UIT (e.g., general ability, personality, and skills/knowledge testing). The decision to use UIT is largely influenced by the stakes involved in passing or failing an exam. Low-stakes exams occur in situations where the outcome of passing or failing an exam has negligible effects on oneself, others, and the exam delivering institutions (e.g., self-assessment or diagnostic test; Pearlman, 2009). On the other hand, high stakes exams will typically result in a decision that has the potential to significantly affect oneself, others, and the exam delivering institutions (e.g., passing or failing a pilot's exam). For example, because of the high-stakes involved in credentialing programs, UIT is typically limited to providing candidates with sample exam questions. For high-stakes exams such as entry to practice and certification, practitioners prefer proctored exams to ensure public safety and legal defensibility of the certification program (Tippins et al., 2006). Testing for recruitment and selection is considered medium to high stakes, depending on the position and skills being tested. Relevant to when and how practitioners use UIT, research examining the contextual features that promote cheating on unproctored Internet tests show that applicants are more likely to cheat if the stakes are high and there is little or no penalty for cheating (see Figure 1; Landers & Sackett, 2012).

Figure 1. The impact of different testing conditions on cheating risk. (Figure retrieved from Landers & Sackett, 2012)

There is no agreed upon standard for assessing risk in employment testing: levels differ from low stakes (e.g., self-assessment for performance management) to high stakes (e.g., assessment for certification, promotion or hiring decision). Practitioners measure the magnitude of risk based on the impact on others. A key question is whether passing an exam without the required qualifications would pose a risk to others. If the position being recruited for is low-risk, then pre-employment testing for that position can be considered low to medium stakes. Putting aside some of the higher-risk industries such as aviation, healthcare, public safety, and national defence, *pre-employment testing* is generally considered low to medium stakes. Furthermore, if one examines how most organizations are using UIT – that is, mainly without remote proctoring methods or confirmation testing – then one might be inclined to categorize pre-employment tests as low stakes.

Researchers who initially investigated UIT were skeptical that it could be used as a valid method for screening job candidates, primarily because of the assumption that test takers would cheat (Pearlman, 2009; Tippins, 2009a, 2009b; Tippins et al., 2006; Weiner & Morrison Jr,

2009). Early research focused on understanding the risks associated with UIT, with an emphasis on candidate cheating, including test exposure. Despite the strong interest in cheating behaviour, relatively few experimental studies have investigated cheating behaviour with actual job applicants in a personnel selection setting.

The few studies that do exist have found very minimal cheating behaviour, generally in the range of 6% to 10% (Arthur, Glaze, Villado, & Taylor, 2009, 2010; Kantrowitz & Dainis, 2014; Lievens & Burke, 2011; Wright, Meade, & Gutierrez, 2014). Most notably, Kantrowitz and Dainis (2014) used pre-existing applicant data ($n=4026$) from SHL² and examined inconsistent test scores across an initial unproctored test followed by a proctored verification test of deductive reasoning. The test was computer adaptive and linked to a large item bank of 300 questions. Applicant data were drawn from five different United States based organizations and across 11 different job families with varying job complexity. Kantrowitz and Dainis' (2014) study results suggest that 6.4% of the applicant pool demonstrated cheating behaviour. The authors tie the low prevalence of cheating to the use of speeded computer adaptive testing administered with a follow-up verification test. The weakness of this approach is that the authors based the inference on cognitive ability, when in fact, there is no way to know if their cheating approach actually influenced their unproctored test score. In other words, this result does not mean that only 6.4% of candidates cheated but that only 6.4% of candidates cheated effectively, probably through the use of a surrogate test taker.

Kantrowitz and Dainis (2014) further propose the model presented in Figure 2 to explain how (and why) cheating occurs on personnel selection tests administered through UIT.

Kantrowitz and Dainis (2014) suggest that several individual and contextual factors form the

² SHL (now Gartner) is a personnel testing and assessment company. <https://www.cebglobal.com>

Unproctored Internet Testing

antecedents of cheating behaviour on unproctored employment tests. With regards to the individual difference factors, the authors theorize that higher scores on conscientiousness, integrity, and risk aversion, may be related to lower levels of cheating on UIT. On the other hand, they suggest that difficult economic factors, perceptions of low job availability, and greater desire for the job are what form the contextual factors that may increase cheating behaviour among job applicants. While Kantrowitz and Dainis' (2014) research is helpful in explaining the prevalence of cheating in the context of UIT, it is apparent from the available research that there is a gap in the literature on the individual and contextual factors that lead to cheating behaviour on personnel selection tests. In the next section, I summarize the extant literature investigating cheating in the academic setting.

Figure 2. Antecedents and mitigators of cheating on unproctored employment tests (figure retrieved from Kantrowitz & Dainis, 2014).

Academic Cheating Research

HR practitioners investigating UIT would prefer to rely on personnel selection research to inform recommendations for best practice. Unfortunately, cheating has been studied less frequently, and albeit with less rigor, in the corporate setting than it has in the academic context. The earliest paper on UIT and cheating dates to Greenberg (1999), whereas cheating research in the academic setting dates to more than 100 years ago (Harris, 1916). Therefore, the purpose of this section is to identify relevant findings from academic cheating research and explore the implications for UIT. The following research questions guided the literature review for this section:

- How is cheating defined in the academic literature?
- How prevalent is cheating on university campuses?
- Which theories have been used to explain cheating behaviour?
- What are the contextual/situational factors that influence cheating?
- Which individual differences are correlated to academic cheating behaviour?

Academic Cheating – Definition and Prevalence

Cheating is broadly defined as “breaking the rules to get ahead academically, professionally, or financially” (Callahan, 2004). In this broader definition of cheating, response distortion to personality tests, lying during interviews, fudging credit applications, and falsifying work histories, would all be considered cheating. Despite the societal impact of breaking professional and financial rules, researchers have focused very narrowly on one form of cheating – academic. In fact, cheating has been most extensively studied within the context of universities and there is consensus among researchers that academic dishonesty in the form of cheating is widespread (e.g., Bloodgood, Turnley, & Mudrack, 2008; Carpenter, Harding, Finelli,

Montgomery, & Passow, 2006; Carpenter, Harding, Finelli, & Passow, 2004; Harding, Carpenter, Finelli, & Passow, 2004; McCabe & Trevino, 1997; McCabe, Trevino, & Butterfield, 2001).

Over the past thirty years, research has consistently shown that approximately 40% to 80% of students will cheat at least once in their academic careers (Mayhew, Hubbard, Finelli, Harding, & Carpenter, 2009). There is also evidence that cheating behaviour, especially in the form of test cheating, has steadily increased over time (Spiller & Crown, 1995) and that student cheating varies by academic department (McCabe & Trevino, 1997), with engineering students reporting the highest levels (82% to 96%) of academic dishonesty and cheating (Harding, Mayhew, Finelli, & Carpenter, 2007). More recent research from Stogner, Miller, and Marcum (2013) suggest that 40% of students engage in e-cheating and that e-cheaters are also more likely to cheat on traditional exams and assignments.

Cheating in the academic context has been defined both narrowly (e.g., copying someone else's answers during an exam or using cheat notes during a test) and more broadly to include several forms of academic dishonesty (e.g., falsifying a bibliography, or learning what was on a test from someone else; McCabe et al., 2001). Regardless of how academic cheating has been defined, there is agreement among researchers that cheating constitutes unethical behaviour (Buckley, Wiese, & Harvey, 1998; Carpenter et al., 2004; Harding et al., 2004; Harper, 2006; Spiller & Crown, 1995).

Theory of Planned Behaviour

While researchers have explored several theoretical frameworks to explain cheating behaviour, (i.e., cognitive dissonance theory; Tang & Zuo, 1997; Vinski & Tryon, 2009), deterrence theory (Buckley et al., 1998), and rational choice theory (Buckley et al., 1998) the

most widely applied framework has been the theory of planned behaviour (TPB) (e.g., Alleyne & Phillips, 2011; Beck & Ajzen, 1991; Harding et al., 2007; Lonsdale, 2017; Mayhew et al., 2009; Stone, Jawahar, & Kisamore, 2009, 2010; Whitley, Nelson, & Jones, 1999). The TPB encapsulates the idea that behaviour is the outcome of rational decision-making in which the agent has pre-conceived behavioural beliefs and expectations with regards to the results of the behaviour (Ajzen, 1991). In other words, behavioural achievement is the result of perceived behavioural control and the strength of behavioural intention. As such, a core tenet of the TPB is the formation of intentions along with an individual's drive to engage in behaviour and the motivational factors which play a role in the behavioural execution of intentions. The strength of the intention determines the likelihood of behaviour in instances where the behaviour is within volitional control (Ajzen, 1991). The TPB posits that three main constructs play a role in forming an intention: 1) the attitude toward behaviour (behavioural beliefs); 2) subjective norms (normative beliefs); and 3) perceived behavioural control (control beliefs).

The TPB (Ajzen, 1988, 1991) – which is an extension of Ajzen and Fishbein's (1977) Theory of Reasoned Action (TRA), emerged as an integrated theory of action to explain the empirical gap between attitudes and behaviour. Specifically, social psychologists were concerned with empirical studies showing that attitudes alone were insufficient in predicting behaviour (Armitage, & Conner, 2001). The TRA initially posited that attitudes and subjective norms led to the formation of intentions, and that intentions paved the way for action. While this initial model was successful in explaining behaviours that were relatively simple in nature (and considered under volitional control), the theory was less adequate in explaining behaviours when there were constraints on action (e.g., lack of skill, knowledge, resources, money, etc.) (Armitage, & Conner, 2001).

Led by the realization that situational or individual constraints can prevent follow through on intentions, Azjen (1991) extended the Theory of Reasoned Action (TRA) by including the construct of perceived behavioural control (PBC), which unlike the other two TRA constructs was theorized to have a direct effect on both the formation of the intention *and* the conversion to behaviour. The PBC construct was used to predict behaviour when other factors impeded on the intention to follow through. According to Ajzen (1991), belief that one has control over behaviour and the subsequent outcomes inevitably influences the strength of one's intention. In other words, PBC is the individual's perception of how difficult or easy it will be to perform a given task or activity (Ajzen, 1991).

The attitudes toward the behaviour describe one's perception of how they perceive the behaviour – for example, whether they feel the act is positive, beneficial, or detrimental to themselves or others. Evidence of positive attitudes toward a behaviour predict the formation of intentions (Azjen, 1991). Subjective norms refer to the normative expectations of others with regards to the behaviour, and the individual's general perception of how acceptable/unacceptable or encouraged/discouraged the behaviour is by their peers, family, and friends. In sum, the TPB is a robust theory that predicts a wide range of intentions and subsequent behaviours, and has been applied to more than 750 theory, review, and empirical publications (Stone, Jawahar, Kisamore, 2010).

Theory of Planned Behaviour and Cheating

A review of the TPB literature on cheating and academic dishonesty suggest that cheating is intentional and follows the general framework of decision-making and behaviour posited by the TPB. Mayhew, Hubbard, Finelli, Harding, and Carpenter (2009) examined college cheating behaviour using a slightly modified version of the TPB which included additional variables such

as moral obligation and past cheating behaviour (see Figure 3). To examine the predictive validity of their modified TPB framework, Mayhew et al. (2009) used structural equation modeling and found evidence for their hypothesized model. In fact, all pathways were significant, and there was a strong relation between high school cheating and college cheating behaviour. It is important to note that Mayhew et al. (2009) also examined how individuals at various stages of moral development (based on Kohlberg's theory) formed an intention to cheat. Study results suggest path coefficients vary in magnitude depending on which stage of moral development the student was in (i.e., transition vs. consolidation). Mayhew et al. (2009) provide strong evidence that TPB constructs are related to cheating, and that past cheating and moral obligation are additional predictors which can be incorporated into the TPB model.

Stone, Jawahar, and Kisamore (2010) examined the relative strength of the TPB model in predicting academic cheating among a student sample. In addition to the TPB constructs, the authors also examined the influence of two personality traits – prudence and adjustment, which are measured in the Hogan Personality Index (HPI; Hogan & Hogan, 1995). High scores on prudence are associated with favourable behaviours such as following protocol, attention to detail, and consistency, whereas lower scores are associated with irresponsible behaviours such as acting impulsively and carelessly (Hogan & Hogan, 1995; Stone, Jawahar, & Kisamore, 2010). Prudence is similar to the five-factor trait of conscientiousness. On the other hand, low scores on the trait of Adjustment indicate lower subjective well-being, poor stress management, and poor temperament, whereas higher scores are associated with self-confidence and an ability to work effectively under pressure (Hogan & Hogan, 1995; Stone, Jawahar, & Kisamore, 2010). Low scores on Adjustment are similar to neuroticism (Stone, Jawahar, & Kisamore, 2010).

Stone, Jawahar, and Kisamore (2010) conducted structural equation modelling, one full mediation model, and one partial model, to examine the individual pathways between each trait, TPB constructs, and behaviour. Perceived Behavioural Control (PBC) was the strongest predictor of intentions and actual cheating behaviour, independent of the intention-behaviour relationship. While all TPB constructs were significantly related to intentions, the evidence for the influence of subjective norms on intentions was the weakest, although a stronger link between subjective norms and PBC suggest that one's circle of influence may alter their perceptions of how easy or difficult it is to cheat. Based on the authors mediation model, TPB fully mediated the relationship between the two HPI traits (prudence and adjustment) and intentions to cheat. Prudence had a direct effect on behaviours but not on intentions to cheat whereas adjustment had no effect on either TPB construct. In sum, the authors found strong evidence for the TPB in explaining the variance in intentions to cheat and actual academic misconduct.

Considering the TPB and past research results, we can conclude that if one can prevent candidates from forming an intention to cheat, then the overall frequency and magnitude of cheating should decrease. On the other hand, the independent pathway between PBC and cheating behaviour suggests that students who are very confident in their cheating skills will cheat regardless of the other TPB variables. An implication of this finding is that warning messages which suggest that cheating will be difficult may be effective in lowering PBC and actual cheating behaviour. In the following sections, I examine the contextual and individual factors that influence cheating, keeping in mind that some of the factors are related to the TPB constructs.

Figure 3. Modified TPB Framework for College Cheating. (figure retrieved from Mayhew et al. (2009))

Academic Cheating – Contextual Factors

Researchers investigating academic cheating delineate between contextual and individual differences when examining factors that influence cheating behaviour (McCabe & Trevino, 1997). With regards to contextual factors that influence cheating, research has investigated group membership (e.g., fraternity; McCabe & Trevino, 1997), classroom instruction (Alt, 2014; Bloodgood et al., 2008), perception of peer behaviour (McCabe & Trevino, 1997), cheating deterrence strategies (Carpenter et al., 2006), and contextual values that promote or deter cheating (e.g., honour/integrity codes; McCabe & Trevino, 1997). In research which surveyed 6000 students from 31 campuses across the United States, McCabe and Trevino (1993) found that the strongest correlate to cheating was perceptions of peer behaviour. Students cheat because they learn from their peers that this is acceptable behaviour (McCabe et al., 2001). Moreover, if students believe that their peers and classmates are cheating than they will be inclined to “level the playing field” by cheating themselves (McCabe et al., 2001).

Perception of peer behaviour and its influence on cheating is examined in this thesis through the Theory of Planned Behaviour’s (TPB) subjective norms construct. From a contextual

perspective, the social pressure to engage (or not) in certain behaviours is theorized to directly influence our intentions and how easy or difficult we perceive the action to be (or our Perceived Behavioural Control: PBC) (Ajzen, 1991). There is less evidence in TPB research to support a direct (independent) effect between subjective norms and behaviour, and in this regard, PBC has been shown to mediate the relationship (Stone, Jawahar, & Kisamore, 2010).

In addition, McCabe and Trevino (1997) found that the strongest contextual factor related to cheating prevention was the existence of honour and integrity codes. According to McCabe and Trevino (1997), most academic institutions have honour and integrity codes that outline the policies and procedures that relate to cheating and academic dishonesty. While utilized to varying degrees by academic institutions, honour and integrity codes can form the basis of student cheating attitudes if they become an integral part of campus culture (McCabe & Trevino, 1997). Relating this finding back to Mayhew et al.'s modified TPB framework, honour and integrity codes might instill a sense of moral obligation, which in turn can be used to deter cheating.

Although early research into academic cheating behaviour emphasized the influence of contextual factors (McCabe & Trevino, 1997), there has been a shift in the past few decades towards an examination of individual differences. In fact, there is a consensus in the literature that individual differences determine academic cheating behaviour to a significant extent (Gallant & Drinan, 2006)³.

³ Meta-analytic findings suggest that, taken together, contextual factors are stronger predictors of academic dishonesty than individual differences (Whitley et al., 1999), although more recent research has identified the TPB construct of Perceived Behavioural Control (PBC) as the strongest predictor of academic cheating behaviour.

Academic Cheating – Individual Differences

An examination of the academic cheating literature provides evidence that the following individual differences may contribute to increased cheating behaviour: being male (Whitley, Nelson, & Jones, 1999), having an external attribution bias (Davis, Grover, Becker, & McGregor, 1992), scoring high on sensation seeking (DeAndrea, Carpenter, Shulman, & Levine, 2009) and psychopathy (Coyne & Thomas, 2008), as well as scoring low on measures of the following constructs: intelligence and religiosity (Bloodgood et al., 2008), conscientiousness (Giluk & Postlethwaite, 2015), agreeableness (Giluk & Postlethwaite, 2015), self-esteem (Buckley et al., 1998), moral development (Mayhew et al., 2009)⁴, and integrity (Lucas, 2005).

Whitley, Nelson, and Jones (1999) conducted a meta-analysis on sex differences and academic cheating. The authors utilized Gilligan's (1982) differential socialization theory of gender differences as an organizing and explanatory framework. The theory proposes that differences in moral orientation among males and females produce differential behavioural outcomes (Whitley, Nelson, & Jones, 1999). In the context of cheating, Whitley, Nelson, and Jones (1999) propose that males will be more likely than females to cheat because of their tendency to engage in riskier behaviour. According to the authors, females are socialized to higher standards of moral behaviour, and are more concerned how their behaviour affects others, hence they will be less inclined to engage in academic misconduct such as cheating (Whitley, Nelson, & Jones, 1999). Across 42 studies, the results of their research reveal a mean effect size of moderate magnitude (equivalent to $r=.21$) for sex differences and cheating attitudes. Their analysis on actual cheating behaviour did not reveal significant differences (Whitley, Nelson, & Jones, 1999).

⁴ Interestingly, the relationship between cheating and moral development is statistically non-linear.

Several of the individual differences related to cheating are examined in the context of UIT to determine if they also influence cheating behaviour in personnel testing. Specifically, I examine the influence of intelligence, conscientiousness, and agreeableness. I chose to investigate these individual differences because they have been shown to predict job performance and are typically incorporated into personnel assessments. Further to the research cited above, I also examined time since graduation as a demographic characteristic which may influence cheating behaviour. This is based on the assumption that recent graduates will hold subjective norms influenced by their peer groups (i.e., recent graduates who may hold favourable attitudes toward cheating).

It is evident from the literature review conducted for this Master's thesis that individual differences play a critical role in the outcome of cheating behaviour. Unfortunately, there is little value from an applied perspective if the individual differences associated with cheating are characteristics that we cannot change or measure in the context of a personnel selection program (e.g., gender, religiosity). The strongest benefit from the individual differences literature on cheating is that we can be more cognizant of cheating behaviour when targeting specific employment groups such as recent graduates or college students. On the other hand, contextual factors such as the constructs of the TPB are situation specific characteristics that practitioners can control without disadvantaging particular groups.

Academic Cheating – Implications for UIT

Cheating research within academic settings has important implications for personnel selection research, especially when we consider hiring for corporate level positions within organizations. The prevalence of academic cheating is particularly worrying because of research showing a positive correlation between cheating behaviour in college and unethical work

behaviour (Carpenter et al., 2004; Harding et al., 2004; Harper, 2006; Nonis & Swift, 2001).

Even more worrying is that cheating behaviour shares an unappealing relation with almost every individual characteristic known to correlate positively with work performance (e.g., intelligence, conscientiousness, and integrity). If academic cheating research is purporting that 40%-80% of students cheat at least once in their academic careers, then this statistic is at odds with the UIT researchers claiming prevalence rates as low as 6.4% (Kantrowitz & Dainis, 2014). Most corporate level positions require academic degrees, and major companies target university campuses for recruitment; correspondingly then, approximately 40%-80% of applicants applying for a job would have cheated at some point in their academic careers. It is then also plausible that these individuals might be tempted to cheat on employment tests if they are unproctored.

Similarly, Burke (2009) suggests that job applicants would be likely to cheat on UIT exams simply because data suggest that 45% of job applicants falsify employment histories on resumes.

The academic literature on cheating does fall short, however, on some aspects of UIT related issues and concerns. First, very little cheating research in the academic context has focused on response distortion to personality tests, as well other measures (e.g., integrity) or forms (e.g., work history) used in the personnel selection process. As such, the scope of cheating behaviour in the academic context is much narrower than how it has been conceptualized by UIT researchers. Furthermore, despite the acknowledgment that cheating is widespread in the academic context, very little research has focused on prevention and detection methods. In fact, most modern cheating prevention/detection methods have emerged in the context of UIT research and not the academic cheating literature. In the following section, I discuss research which has examined the relation between response distortion and cheating. Following this, I summarize the available research on cheating prevention and detection.

Cheating and Response Distortion – How do they differ?

While cheating and response distortion are often referred to as malfeasant candidate behaviours (Arthur et al., 2010), cheating in the test and measurement literature (not to be confused with the UIT literature) is defined in the context of cognitive and knowledge tests; for example, when test takers use devices (e.g., calculators, smart phones) that are not allowed, use surrogate test takers who might perform better on the test, or identify test content through Google searches and use this content during the exam (Arthur et al., 2010). In cognitive test cheating, the test taker is trying to appear more intelligent - which may cover up their actual ability to learn, solve complex problems, and perform work related activities. On the other hand, response distortion occurs in the context of non-cognitive testing (e.g., personality/attitude measures) and is theorized to be the outcome of a test-taker's desire for social acceptance (e.g., social desirability) and an intentional impression management tactic. In response distortion, the test taker is trying to appear more suitable to the position, organization, and team for which they are being assessed.

There is some commonality between the reasons for cheating and response distortion when applied to employment settings. In both cases, job applicants are trying to deceive the organization by presenting themselves as more suitable for the position than they actually are. Similar to cheating, the ethical implications of response distortion depend on the context in which the measure is administered. If for employment, then distorting responses for social desirability should carry similar ethical implications.

Millham (1974) investigated the relation between need for approval and cheating. The author conducted a study in which he ran participants through two conditions – success and failure of an IQ test. After failing or succeeding on the IQ test, which to the participant was

presented based on a norm referencing procedure, participants were provided the opportunity to change their test results by falsifying the number of items they correctly responded to. Millham (1974) measured social desirability using the Marlowe-Crowne Social Desirability Scale (Crowne & Marlowe, 1960). Millham (1974) found that participants high in need for approval were more likely to cheat following a failure than they were a success. Millham's (1974) study also found that fear of failing or scoring below the norm is a greater motivation for cheating than the desire to succeed or surpass others. In other words, individuals cheat to maintain status quo, not necessarily exceed it. Relating Millham's (1974) findings to the theory of planned behaviour, participants must find the perception of scoring below the norm as carrying greater consequences to the self than their desire to succeed, validating the notion that perceptions of consequences strongly impact the formation of cheating intentions.

The similarity between cheating and response distortion becomes more nuanced when the intentionality of the acts are considered. It has been theorized that some dimensions of social desirability are not only other-deceiving but self-deceiving as well (Arthur et al., 2010; Zettler, Hilbig, Moshagen, & de Vries, 2015). In other words, impression management in some instances is involuntary. You may deceive yourself into believing you are more conscientious than you actually are simply because you have been socialized to believe that its great to have that personality characteristic. To be other-deceiving, the individual would have to intentionally decide to respond in a way that is different from who they are. For example, an individual might respond more strongly to conscientiousness related questions because they assume this is what the job provider is seeking.

Another major difference between cheating and response distortion is that the presence or absence of a proctor has no bearing on response distortion in personality testing (Arthur et al.,

2010). Arthur, Glaze, Villado, and Taylor (2009) investigated whether response distortion was more prevalent in UIT based personality tests. Their study results showed that the prevalence of cheating on a speeded UIT cognitive ability test was around 7.7%. Arthur, and colleagues found that using a speeded test to measure cognitive ability in a simulated pre-employment (high stakes, participants as job applicants scenario) or employment testing (low stakes, participants as job incumbents scenario) helped prevent cheating in the higher stakes condition. With regards to personality tests, they found that response distortion increased in the high stakes condition, regardless of whether it was proctored or not. The magnitude of score inflation was greater on Agreeableness, Conscientiousness, and Emotional Stability compared to Extraversion and Openness. Detrick, Chibnall, and Call (2010) found similar results for the NEO-PI when used in the recruitment context for Police officers.

More recently researchers have called into question the idea that social desirability scales actually measure impression management. On the contrary, the argument has been put forth that high scores on impression management measure true virtue either in the form of self-control (Uziel, 2010, 2014) or honesty (de Vries, Zettler, & Hilbig, 2014) when administered in low-stakes settings. Building off Uziel (2010, 2014) and de Vries, Zettler and Hilbig's (2014) research, Zettler, Hilbig, Moshagen, and de Vries (2015) found that high-scores on an impression management scale (Balanced Inventory of Desirable Responding) were actually related to *less* cheating behaviour in a low-stakes testing condition. The research also replicated de Vries, Zettler, and Hilbig's (2014) findings that higher scores on impression management are aligned with higher scores on Big Five factors such as agreeableness and conscientiousness – two of the strongest predictors of job performance. While the findings of Zettler, Hilbig, Moshagen, and de Vries (2015) study are very interesting from a personality measurement perspective, they carry

less weight or impact on personnel selection research. Although individual perceptions of stakes might differ across job applicants, it is very difficult to categorize personality testing during personnel selection as low stakes. Moreover, an individual's propensity to distort responses on a personality measure when they know the results will be used for evaluation have been well established. What is not so clear in personnel research is whether higher scores on impression management (collected in higher-stake settings) are correlated with increased cheating behaviour. Similarly, Zettler, Hilbig, Moshagen, and de Vries (2015) also call for more research aimed at delineating honest responders from socially desirable responders.

That there are such few studies linking response distortion on personality to cheating on cognitive (or knowledge) tests is rather surprising given the demand for these types of assessments in the employment setting. Personality, cognitive, and knowledge assessments are used in nearly all aspects of the talent management and acquisition cycle (selection, promotion, team formation, certification, and licensure). For this reason alone, it is crucial that researchers examine the relations between falsifying responses and cheating across all types of assessments. Furthermore, researchers need to identify the consequential effects of each and create measures to both prevent and detect cheating in these different contexts. For the purpose of my Master's thesis, I measured Impression Management (IM) in the follow-up survey administered to job candidates who volunteered to take part in this study. Candidates were assured that the results of the survey questions would not be used in the hiring process, hence, the expectation was that candidates would respond honestly and the IM score could be used a true measure of their general tendency to self-enhance.

UIT – Common Types of Cheating and Prevention Methods

As technology changes the testing and assessment landscape, individuals dedicated to breaking the rules have forged numerous cheating strategies that exploit the weaknesses presented by UIT. Research shows that UIT is susceptible to many forms of candidate cheating. Aware of these cheating methods, HR practitioners and administrators have designed UIT programs that safeguard against these methods (see Table 2). UIT experts recommend the use of identity checks (e.g., bio data verification), speeded tests, statistical and response pattern analysis, and software which monitor and limit end-user capabilities (Kantrowitz & Dainis, 2014). Virtual proctoring companies that provide innovative technology services recommend webcam/audio monitoring, bio-data verification, and even retina or fingerprint scans for high-stake exams (for example, see <https://www.caveon.com/>).

Table 2. Common UIT Cheating Methods: Prevention and Detection

Cheating Method	Prevention Method
Logging in with a phony username to retrieve test content and practice.	- Analysis of duplicate IPs accessing the testing system - Biodata verification
Taking the test with assistance from others.	- Remote webcam monitoring (Hylton et al., 2016)
Exploiting software (web-browser).	- Software that limits computer functionality
Using a handheld device.	- Remote webcam monitoring
Having someone else take the test / using surrogate test taker (i.e., proxy testing, surrogate testing).	- Biodata verification (e.g., retina or fingerprint scans)
Using the Internet or other devices to search for test answers.	- Software that limits computer functionality - Inserting test items that detect cheating - Tracking eye movement - Speeded tests - Use of non-multiple choice tests
Distorting responses on personality measures	- Forced-choice responses - Empirical keying - Warnings of verification testing - Response elaboration

Even more effective than the prevention methods listed above, practitioners recommend the use of warning messages and having candidates agree to a code of honour, terms of use, and copyright statements (Kantrowitz & Dainis, 2014). In line with the TPB's assertion that beliefs regarding the consequences will influence the formation of an intention, deterrence strategies may be an efficient and cost-effective way of reducing cheating behaviour. If candidates feel like the consequences (e.g., professional embarrassment) will outweigh the benefits of cheating (e.g., securing a new job), then the behaviour will be reduced. Below is a list of deterrence techniques that researchers and practitioners suggest minimize the likelihood of cheating in UIT see (Burke, 2009; Kantrowitz & Dainis, 2014; Lievens & Burke, 2011; Lievens & Harris, 2003; Nye, Do, Drasgow, & Fine, 2008; Pearlman, 2009; Tippins, 2009a, 2009b):

- Warning candidates that cheating detection methods are used.
- Informing candidates that there could be a follow-up test.
- Clearly articulating the consequences of cheating.
- Explaining that honesty is in the candidate's best interest.
- The use of an "honesty contract" (Burke, 2009) that requires applicants to agree that they will respond honestly and without obtaining the assistance of others.

Test administrators understand that candidates will find other ways of cheating that go undetected with current measures. To ensure that cheaters do not progress through the recruitment process, program designers have proposed the following two models (see Figure 4 and Figure 5) that utilize UIT in a multi-hurdle process (Weiner, Burke, Fetzner, & Gibby, 2010). The ultimate goal is to prevent cheaters from advancing further in the recruitment process, ensuring the validity and fairness of the process. The drawback of the consistency model is that cost savings associated with UIT will not be recognized unless a very high percentage of

Unproctored Internet Testing

candidates are eliminated at the UIT stage (which is also problematic from a testing best practice perspective). The successive hurdle model is more feasible for practitioners because there is no proctored retest; however, the decision to implement a proctored retest will depend on the volume of testing and pass rates on the UIT.

Figure 4. Model 1: Consistency Testing⁵

Figure 5. Model 2: Successive Hurdles⁶

UIT – Cheating Detection

Cheating in unproctored employment tests is most commonly detected through the use of statistical techniques which entail post-hoc aggregation and analysis of testing data. These statistical methods include the use of item response theory (IRT), statistical simulations⁷, pairwise z-score testing, logit analysis, invariance analysis, response latency changes, score inflation, and score deflation after follow-up testing (Carstairs & Myors, 2009; Guo & Drasgow,

⁵ Figure retrieved from Weiner, Burke, Fetzer, & Gibby, (2010).

⁶ Figure retrieved from Weiner, Burke, Fetzer, & Gibby, (2010).

⁷ According to simulation research conducted by Landers and Sackett (2012), “a large enough increase in applicant pool size will overcome most problems induced by cheating, however extensive this cheating may be” (p. 227).

2010; Kantrowitz & Dainis, 2014; Landers & Sackett, 2012; Nye et al., 2008; Tendeiro, Meijer, Schakel, & Maij-de Meij, 2013; Wright et al., 2014). The analytic techniques are useful for understanding cheating patterns at the aggregate level but fail to provide test administrators with cheating detection in real time.

The most common real-time cheating detection method is the use of virtual proctors who oversee the test-taker through webcam technology. While virtual proctors are commonly used in high-stakes exam settings, they are rarely used in pre-employment testing contexts, mainly because of the prohibitive cost. A less common approach used in practice is the insertion of detection questions that when answered correctly reveal candidate cheating. In comparison to post-hoc techniques, detection questions can alert test administrators in real time when candidates cheat. Test administrators can then retest those individuals who scored the item correctly (instead of re-testing the entire candidate pool).

Cheating detection is an area of opportunity for researchers and practitioners to leverage innovative testing and assessment tools. Testing software platforms provide seamless integrations with virtual proctoring tools and provide a robust and practical platform for enabling other cheating detection and prevention methods. UIT programs should incorporate these technologies into their day-to-day testing activities to ensure the integrity of their testing programs.

Literature Review - Conclusion

Since the widespread use of the Internet to acquire talent, organizations have been using unproctored Internet tests to assess job applicants (Kantrowitz & Dainis, 2014). Formerly in the domain of proctored testing, unproctored ability testing posed a serious challenge to test developers and practitioners. Ryan and Ployhart (2014) highlight unproctored Internet testing

Unproctored Internet Testing

research as an area that science has been unable to produce valid answers for, especially for practitioners who are expected to provide best practice recommendations to organizations. Couple this concern with the pre-existing sensitivities that surround ability testing (i.e., discrimination, fairness, and accessibility, just to name a few), and the field of testing and assessment faces a serious threat. The threat is particularly true in the context of cognitive ability testing – an area of substantial validity research (see Schmidt and Hunter, 1998) – that could be put at risk if empirical linkages between test scores and performance are no longer as strong as they were when testing was proctored. Although no major lawsuits have surfaced sufficient enough to influence organizations to implement strong measures to prevent cheating, this does not mean that researchers should wait until something serious happens to further the empirical investigation of UIT. As such, the overarching purpose of this Master's thesis was to examine the effectiveness of cost-efficient strategies that could be used to deter or even prevent cheating in employment testing. Furthermore, I examined contextual and individual difference factors that enable cheating behaviour and provide recommendations on how to detect and prevent cheating across different employment groups. Participants of this research were actual job applicants recruited for various positions within a talent management company located in Canada.

Hypotheses Based on Current Review

Novel hypotheses that extend the research in this area:

- Hypothesis 1: Using the Theory of Planned Behaviour as a guiding framework, I hypothesized that cheating behaviour on personnel screening tests would be reduced when the UIT design and communication materials target attitude modification, emphasize moral responsibility, and clearly describe the detrimental consequences of cheating.

Unproctored Internet Testing

- Hypothesis 2: I hypothesize that recent university and college graduate job applicants will be more likely to cheat.
- Hypothesis 3: I hypothesize that cheaters will have higher scores on impression management than non-cheaters.

Hypotheses that replicate past research:

- Hypothesis 4: I hypothesize that cheaters will have more lenient attitudes, subjective norms, and moral obligations toward cheating, and demonstrate stronger perceived behavioural control over cheating.
- Hypothesis 5: I hypothesize that cheaters will score lower on cognitive ability, conscientiousness, and agreeableness when compared to non-cheaters.

Methods

Participants

A sample of job applicants ($n = 110$) were recruited from a database of job applicants seeking employment (see *Table 3*). Job advertisements for six different entry to mid-level positions were posted on a multi-national job board site between November 2017 to February 2018. Throughout the study duration, 199 job applicants applied for the available positions. In total, 110 applicants ($n=55$ Control, $n=55$ Treatment) volunteered and completed the online testing process, and 41 applicants ($n=22$ Control, $n=19$ Treatment) followed through to complete the online questionnaire and survey measures.

Table 3. Sample Characteristics: Means, Frequencies, and Standard Deviations

Sample Characteristics (survey respondents only)	Data
Age	$n=39$ $M=32.44, SD=7.92$
Sex	$n=40$
Male	$n=26$ (65%)
Female	$n=14$ (35%)
Highest Level of Education	$n=40$
Community College	$n=8$ (20.0%)
University Undergraduate Degree	$n=11$ (27.5%)
University Graduate School Degree	$n=17$ (42.5%)
Other	$n=4$ (10%)
Student Status	$n=40$
Not a student	$n=34$ (85%)
Full-time	$n=5$ (12.5%)
Part-time	$n=1$ (2.5%)
Years since graduation	$n=34$
Less than 5 years	$n=18$ (53%)
5 or more years	$n=16$ (47%)
Indigenous Belonging	$n=40$
Yes	$n=0$ (0%)
No	$n=40$ (40%)
Visible Minority	$n=40$
Yes	$n=20$ (50%)
No	$n=20$ (50%)
Visible Minority Group*	$n=27$
Black	$n=3$
Latino/Hispanic	$n=1$
Middle Eastern	$n=2$
South Asian	$n=15$
South East Asian	$n=1$
West Asian	$n=1$
Other	$n=4$

* A discrepancy is noted between those who said yes to being a visible minority ($n=20$), and those who then indicated they belonged to a visible minority group ($n=27$). A possible explanation is that some individuals may not think they are a 'visible' minority, even though they identify with an ethnic origin.

Procedure

After applying for the position through an online job board, all applicants who met the minimum criteria for the position (i.e., education, language, and experience) were invited to take the online cognitive ability test. Job applicants who provided consent to be contacted by the researcher were invited to take part in the research as well. The hiring organization also obtained consent from job applicants to collect and share their non-identifiable testing data for the purpose of this research (for a visual representation of the study procedure, please see Figure 6). Mild deception was used in the researcher's consent form to cover up the true purpose of the research (i.e., investigation of cheating behaviour). The personnel selection data collected as part of the normal hiring procedure was shared with the primary researcher only and included the results of a cognitive ability test (The Wonderlic Cognitive Ability – Pretest; WPT-Q) and cheating detection questions. The hiring organization provided written permission to the researcher to use the relevant testing data for the purpose of this research. The online consent form and subsequent instructions clearly differentiated between the components of the application process that were used for research purposes and the components that contributed to a hiring decision.

All job applicants who met the minimum requirements, regardless of whether they participated in this research, were first administered the online cognitive ability test (WPT-Q) and cheating detection questions (CDT) and were randomly assigned (see Appendix I for randomization plan) to the control or treatment group condition. The randomization plan was slightly altered to include more control group participants near the end of the study. This modification ensured that the conditions had equal sample sizes. Participants were then invited to complete the online questionnaire package, which included a personality questionnaire (The OCEAN 2.0), an impression management test (the BIDR-IM), a questionnaire designed to

Unproctored Internet Testing

capture their demographic characteristics, attitudes and intentions toward cheating, and the moral obligation they feel regarding cheating behaviour (all measures are described in detail below with cross-references to appropriate appendices). In addition to the measures and instructions administered to the control group, the treatment group received test instructions and UIT design components that targeted the participants' attitudes and sense of moral responsibility regarding the consequences of cheating (testing the TPB model to prevent cheating). The treatment also included a mandatory honesty pledge, which if not taken would bump the participant out of the testing system. The TPB cheating prevention messages (i.e., the treatment) were delivered before the WPT-Q and cheating detection questions. Upon completion, participants were brought to a webpage that debriefed them on the study objectives (Appendix J) and this was followed with a request for secondary consent (Appendix K). Noting that deception was used in the initial consent form to conceal the true purpose of the thesis, secondary consent was obtained after the true purpose of the research was revealed to study participants.

Figure 6. Visual Diagram of Study Procedure.

Measures and Materials

Cognitive Ability. Cognitive ability was measured using the Wonderlic Cognitive Ability – Pretest (WPT-Q). The WPT-Q test combines a wide variety of item and problem types to measure what is commonly referred to as general intelligence (Wonderlic Test Manual, 2002). The questions include: word comparisons, disarranged sentences, sentence parallelism, following directions, number comparisons, number series, analysis of geometric figures, and story problems requiring either mathematics or logic solutions (Wonderlic Test Manual, 2002, p. 5). The Wonderlic also correlates strongly with other tests of intelligence including cognitive, verbal, quantitative, spatial, and reading comprehension. The WPT-Q demonstrates strong reliability and validity and has decades of normative data to support its use in an employment context (Bell, Matthews, Lassiter, & Leverett, 2002; Dodrill, 1983; Dodrill, & Warner, 1988; Edinger, Shipley, Watkins, & Hammett, 1985; Hawkins, Faraone, Pepple, Seidman, & Tsuang, 1990; Wonderlic Test Manual, 2002).

Cheating Behaviour. Cheating behaviour was detected using a series of memory questions hereon referred to as the Cheating Detection Test (CDT). This version of the CDT was developed by the author of this thesis for the purpose of detecting cheating among job candidate test takers at the hiring organization. The sole purpose of the CDT is to detect cheating – the item scores do not contribute to a pass or fail but responding correctly to two of the items suggest cheating behaviour. The CDT is comprised of 4 memory tasks (see Appendix C). Test takers are presented with an array of images which they have only a few seconds to memorize. On a subsequent screen, they are asked to arrange the images in the same order they were presented previously. The first two memory tasks increase in difficulty and are meant to provide the candidate an opportunity to learn the task. The third and fourth items are called the impossible

questions (or IQs) and cannot be answered correctly unless the test taker cheats by manipulating the web browser or uses handheld technology such as a cell phone, camera, or tablet. If the candidates answer any one of the IQs correctly their WPT-Q test result was flagged with “caution”. For the purpose of this Master’s thesis, participants who scored either one of the impossible questions correctly (i.e., they correctly arranged 7 out of the 9 images correctly) were grouped into the ‘Cheater’ group.

Personality. Personality was measured using the OCEAN.20, a 20-item five-factor measure derived from the Trait Self-Descriptive (TSD) inventory (see Appendix D; O’Keefe, Kelloway, & Francis, 2012). The OCEAN.20 is a short measure of personality based on the Big Five framework and suitable for organizational research (O’Keefe, Kelloway, & Francis, 2012). The OCEAN.20 measures openness, conscientiousness, extraversion, agreeableness, and neuroticism. Previous research has shown that conscientiousness and agreeableness are strongly correlated with cheating behaviour (Giluk & Postlethwaite, 2015), such that individuals who score high on these traits are less likely to cheat. The purpose of including the OCEAN.20 was to replicate previous research findings, explore the relations between cheating, conscientiousness, and agreeableness across different demographic groups, and examine the relations between cheating and the remaining five factor traits. Scale reliabilities for the OCEAN.20 from the current study were as follows:

- Openness (4-items): Cronbach’s $\alpha = .66$;
- Conscientiousness (4-items): Cronbach’s $\alpha = .82$
- Extraversion (4-items): Cronbach’s $\alpha = .74$
- Agreeableness (4-items): Cronbach’s $\alpha = .87$
- Neuroticism (4-items): Cronbach’s $\alpha = .76$

Impression Management. Impression management was measured using the Balanced Inventory Desirable Responding (BIDR) scale's Impression Management (IM) subscale (Paulhus, 1991). The BIDR's IM scale (see Appendix E) is a 20-item measure of social desirability and assesses whether respondents are being truthful or are misrepresenting themselves in order to manage their self-image (Zettler, Hilbig, Moshagen, & de Vries, 2015). The BIDR has been utilized extensively in organizational and personality research and has demonstrated strong reliability across many different samples (Zettler, Hilbig, Moshagen, & de Vries, 2015). IM was measured to infer the accuracy of conscientiousness and agreeableness scores on the OCEAN.20 and to understand if Cheaters as a group score higher on this characteristic. It was assumed that participants who score high on IM – in other words, participants who respond in socially desirable ways – would inevitably respond favourably to conscientiousness related items given the context of this research (i.e., personnel selection). IM was also measured to examine the relationship between socially desirable responding and cheating behaviour. In the current study, the BIDR IM scale initially resulted in a Cronbach's α of .80. Item 14 'I never read sexy books or magazines' was removed from the scale because of very low item-total correlation ($r=.07$). The Cronbach's $\alpha = .81$ for the corrected 19-item measure.

Cheating Attitudes. The Theory of Planned Behaviour (TPB) constructs related to cheating behaviour (i.e., attitudes, intentions, subjective norms, perceived behavioural control, and moral obligation) were measured with an adapted and modified version of Harding et al.'s (2007) *The Perceptions and Attitudes toward Cheating among Engineering Students Survey - Version 2* (PACES-2; see Appendix F). The PACES-2 was originally developed by Harding et al. (2007) to assess the constructs of the theory of planned behavior. Two additional versions of

the PACES-2 were adapted from Stone, Jawahar, and Kisamore's (2010) and Imran and Nordin's (2013) studies to build the version included in this Master's thesis. Two additional items were developed for the Attitudes subscale to measure attitudes towards surrogate test taker cheating. To align with the objectives of my thesis, the three versions of the PACES-2 were further adapted to the context of personnel selection. In sum, the modified measure was comprised of 29-items: 5-items measuring Intentions (IN), 7-items measuring Attitudes Towards Cheating (ATC), 8-items measuring Subjective Norms (SN), 6-items measuring Perceived Behavioural Control (PBC), and 3-items measuring Moral Obligation (MO). All items are scored on a five-point Likert type scale, ranging from strongly disagree to strongly agree. Harding et al., (2007) noted reasonable internal consistency scores (ranging from .67 to .90) for each subscale of the PACES-2. In the current study, The Cronbach's alphas for the adapted PACES-2 subscales are as follows:

- Intentions (5-item): Cronbach's $\alpha = .80$
- Attitudes (7-item): Cronbach's $\alpha = .69$
- Subjective Norms (7-item): Cronbach's $\alpha = .86$: Removed one reverse coded item: 'NO ONE who is important to me (e.g., my family, friends, colleagues, etc.) thinks it is OK to cheat on an in-class test or exam' which had a very low item-total correlation $r = .07$, and resulted in a very low initial Cronbach's $\alpha = .79$.
- Perceived Behavioural Control (6-item): Cronbach's $\alpha = .74$
- Moral Obligation (2-item): Cronbach's $\alpha = .86$: Removed one item: 'It would NOT be morally wrong for me to cheat on an employment test.' which had a fair item-total correlation, $r = .34$, but low Cronbach's $\alpha = .70$ if the item was retained.

Warning Messages. The treatment for this research was the difference in instructions and warning messages provided to the participants in the treatment group. The treatment group instructions included the TPB deterrence strategy and content such as warning messages that

target cheating intentions, attitudes towards cheating, subjective norms related to cheating, perceived control of cheating behaviour, and the sense of moral obligation one feels to not engage in cheating behaviour. The warning messages also included the possibility of a proctored follow-up test, cheating detection measures and questions, and post-hoc analytic techniques (see Appendix G). The TPB deterrence messaging was included in the communication material as early as the invitation email. The control group instructions included the basic test taking instructions and guidelines but no TPB deterrence content or instructions.

Demographic Survey. Demographic information related to gender, age, and ethnicity was collected (see Appendix H). Demographic information was collected to describe the sample and compare mean scores across the battery of tests and survey measures. The primary purpose of collecting and examining the demographic data was to ensure that confounding variables, should any exist, were controlled for during the analysis.

Results

Preliminary Analysis

Prior to conducting the main analysis, the data were reviewed for missing data, univariate outliers, and normality of distributions. First, Little's MCAR test for missing values was conducted. The result of the missing values analysis showed that the few points of missing data were missing at random (Little's MCAR test $\chi^2 = 264.16$, $df = 445$, $p = 1.0$). The total percentage of missing values did not exceed 4.9% ($n = 2$) across any specific item. No action was taken to impute data points or remove cases from the data set. Next, all reverse coded items were recoded into new variables and scale means were computed using IBM SPSS Statistics Subscription Software. Scale reliabilities for all measures and subscales were computed and items with very

low item-total correlations were removed from the computation of scale means (see Measures section for more details on which items were removed).

Standardized scores for all quantitative scale variables were computed and saved to check for univariate outliers (i.e., scores that were + or – 3.29 SD of the mean). Several scales contained univariate outliers that were adjusted: the Conscientiousness and Agreeableness scales of the OCEAN.20 contained one outlier each, while the Moral Obligation scale of the PACES 2 contained two univariate outliers. All outlier scores were adjusted to the scale score value that corresponded to + or - 3.29 SD of their respective scale means. Next, standardized skewness and kurtosis statistics were computed to verify if the scores for all scale variables were normally distributed. Each skewness and kurtosis statistic was standardized by dividing its score by its corresponding standard error (see *Table 4*). Standardized skewness and kurtosis statistics that were greater than + or – 1.98 SD were considered to have asymmetric and/or non-normal distributions. In addition, Q-Q plots were examined for all scale variables to determine whether the assumption of normality of residuals was met. While the distribution of residuals for the Wonderlic Cognitive Ability Test (WPT-Q), the BIDR Impression Management test, PACES 2 Perceived Behavioural Control scale, and OCEAN.20 Openness, Conscientiousness, and Extraversion scales met the assumptions of normality, residuals for the remaining scale variables (i.e., PACES 2 Attitudes, Intentions, Norms, and Moral Obligation, and the OCEAN.20 Agreeableness and Neuroticism scales) were non-normally distributed and followed a similar pattern of deviations from the line of best fit. Although the BIDR Impression Management scale was significantly negatively skewed ($-.74$, $SE = .37$, $z = -1.98$, $p = .047$), the Shapiro-Wilk test of normality ($S-W = 0.95$, $DF = 40$, $p = .08$) suggest that normality is a reasonable assumption. Visually, the distribution of the histogram was bell shaped and although there were slight

deviations from the line of best fit in the Q-Q plot, the remaining points adhered closely to the line of best fit also suggesting normality. The PACES 2 Cheating Attitudes scale was significantly positively skewed while the PACES 2 Cheating Intentions, Subjective Norms, and Moral Obligation scales were significantly positively skewed and leptokurtic, and deviated from the line of best fit at higher and lower values of the scale. As displayed in *Table 4*, Analysis of the standardized skewness and kurtosis statistics, Shapiro-Wilk tests of normality, along with the examination of Q-Q plots were used to determine the suitable statistical approach (parametric vs non-parametric) to test each hypothesis. It is crucial to note that the assumption of independence of errors was met in the present study because participants did not know each other, and all participants took part in this online study separately. The small sample size, specifically on the follow-up survey measures ($n = 41$) requires a conservative interpretation of the significant and non-significant findings.

Table 4. Scale Descriptive and Standardized Skewness and Kurtosis Statistics

Scale Name	Scale Statistics	Standardized Statistics	Shapiro-Wilk Test of Normality
WPT-Q	$n = 101$ $M = 20.67$ $SD = 4.83$	<i>Skewness</i> $Z = -0.74, p = .46$ <i>Kurtosis</i> $Z = -1.51, p = .13$	$(S-W = 0.98, DF = 101, p = .12)$
OCEAN.20 - Openness	$n = 41$ $M = 4.70$ $SD = 1.06$	<i>Skewness</i> $Z = -1.01, p = .31$ <i>Kurtosis</i> $Z = 0.73, p = .47$	$(S-W = 0.98, DF = 41, p = .51)$
OCEAN.20 - Conscientiousness	$n = 41$ $M = 5.74$ $SD = 0.90$	<i>Skewness</i> $Z = -1.09, p = .28$ <i>Kurtosis</i> $Z = -0.71, p = .48$	$(S-W = 0.95, DF = 41, p = .095)$
OCEAN.20 - Extraversion	$n = 41$ $M = 4.64$ $SD = 1.22$	<i>Skewness</i> $Z = 0.47, p = .64$ <i>Kurtosis</i> $Z = -1.07, p = .29$	$(S-W = 0.96, DF = 41, p = .11)$
OCEAN.20 – Agreeableness*	$n = 41$ $M = 6.27$ $SD = 0.64$	<i>Skewness</i> $Z = -2.71, p = .007$ <i>Kurtosis</i> $Z = 0.76, p = .45$	$(S-W = 0.89, DF = 41, p = .001)$
OCEAN.20 – Neuroticism*	$n = 41$ $M = 2.74$ $SD = 1.29$	<i>Skewness</i> $Z = 2.26, p = .02$ <i>Kurtosis</i> $Z = 0.37, p = .71$	$(S-W = 0.93, DF = 41, p = .02)$
BIDR - Balanced Inventory Desirable Responding: Impression Management Scale	$n = 40$ $M = 5.15$ $SD = 0.90$	<i>Skewness</i> $Z = -1.98, p = .047$ <i>Kurtosis</i> $Z = 0.27, p = .79$	$(S-W = 0.95, DF = 40, p = .08)$
PACES 2 – Intentions*	$n = 40$ $M = 1.39$ $SD = 0.62$	<i>Skewness</i> $Z = 4.96, p = .0001$ <i>Kurtosis</i> $Z = 4.24, p = .0001$	$(S-W = 0.69, DF = 40, p = .000)$
PACES 2 – Attitudes*	$n = 40$ $M = 1.62$ $SD = 0.55$	<i>Skewness</i> $Z = 2.63, p = .009$ <i>Kurtosis</i> $Z = 0.47, p = .64$	$(S-W = 0.90, DF = 40, p = .002)$
PACES 2 – Subjective Norms*	$n = 40$ $M = 1.60$ $SD = 0.74$	<i>Skewness</i> $Z = 3.35, p = .0008$ <i>Kurtosis</i> $Z = 1.23, p = .22$	$(S-W = 0.81, DF = 40, p = .000)$
PACES 2 – Perceived Behavioural Control	$n = 40$ $M = 2.34$ $SD = 0.98$	<i>Skewness</i> $Z = 0.66, p = .51$ <i>Kurtosis</i> $Z = -1.05, p = .29$	$(S-W = 0.95, DF = 40, p = .07)$
PACES 2 – Moral Obligation*	$n = 40$ $M = 1.44$ $SD = 0.83$	<i>Skewness</i> $Z = 5.23, p = .0001$ <i>Kurtosis</i> $Z = 4.21, p = .0001$	$(S-W = 0.60, DF = 40, p = .000)$

*Asymmetrical and/or non-normal distribution

Main Analyses and Hypothesis Testing

Statistical Approach. Two main statistical approaches were applied to test the hypotheses for this study. First, chi-square tests of independence were used to determine whether the distribution of frequencies across the Cheater vs. Non-Cheater group changed as a function of experimental condition and time since graduation. I used a chi-square test for this analysis because of the categorical nature of both variables – Condition (Treatment vs Control) and Cheating Group (Cheater vs Non-Cheater). Second, independent samples *t*-tests were used to compare scale means (i.e., the BIDR – Impression Management Scale, the PACES 2 Perceived Behavioural Control, OCEAN.20 Conscientiousness, and the WPT-Q Cognitive Ability Scale), across the Cheater vs. Non-Cheater groups. I used the Levene’s Test of Equality of Variances to test whether the assumption of homogeneous variances was violated. For Levene’s tests with statistically significant results, I used the SPSS *t*-test for “equal variances not assumed.” When the assumption of normality was violated for a comparison between groups, I used the non-parametric Mann-Whitney U Test of independence.

Hypothesis 1: Impact of cheating deterrence strategy on cheating behaviour. Strict warning messages, honesty pledges, and other cheating deterrence strategies are commonly used in practice and are argued to reduce cheating behaviour. Based on previous research and testing best practice, I hypothesized that a cheating deterrence strategy that targeted the constructs of the Theory of Planned Behaviour (TPB) would significantly reduce the frequency of cheating behaviour. To test this hypothesis, the sample was split into two groups: Treatment and Control - the treatment being the deterrence and warning messages received by candidates before taking the personnel selection tests. A chi-square analysis was then conducted between the two experimental conditions – Treatment and Control – to examine whether the cheating deterrence

measures reduced actual cheating behaviour frequency as measured by the Cheating Detection Test. The frequency of cheating behaviour across the experimental conditions is presented in *Table 5*. Contrary to the hypothesis and popular practice among testing professionals, the results suggest that cheating deterrence in the form of warning messages and an honesty pledge did not significantly reduce cheating behaviour in the treatment group, $\chi^2 (1, n=108) = 2.47, p = .12, r = .15$ ⁸. The observed statistical power of the test was 34%. A very surprising finding is that 25.5% of the treatment group – all of whom pledged to take the test honestly – demonstrated cheating behaviour.

Table 5. Frequency of Cheating Behaviour across the Experimental Conditions

	Control Group	Treatment Group	Totals
Cheater Group	<i>n</i> = 21 (39.6%)	<i>n</i> = 14 (25.5%)	<i>n</i> = 35 (32.4%)
Non-Cheater Group	<i>n</i> = 32 (60.4%)	<i>n</i> = 41 (74.5%)	<i>n</i> = 73 (67.6%)
Totals	<i>n</i> = 53 (49.1%)	<i>n</i> = 55 (50.9%)	<i>n</i> = 108 (100%)

Hypothesis 2: Time since graduation and cheating behaviour. One of the core tenets of the TPB is that subjective norms play a critical role in the formation of intentions and behaviour. Previous research (McCabe & Trevino, 1997; McCabe, Trevino, & Butterfield, 2001) suggests that university students are more likely to cheat on exams because cheating is very common among campus culture and student peer groups. It was theorized that time since graduation would represent a strong proxy measure to subjective norms and a useful metric in its relation to cheating behaviour because recruiters commonly target recent university/college

⁸ The statistical power to identify a medium effect size ($\Phi=0.3$) with the current sample ($n=108$) was 88%. In other words, there was a 12% chance that this analysis would find a nonsignificant result if in fact there was a true difference in the population.

graduates for employment. Based on this reasoning, I hypothesized that recent graduates would be more likely to cheat. To test this hypothesis, the sample was split into two groups: Group 1) Less than 5 years since graduation, and Group 2) More than 5 years since graduation. A chi-square analysis was then conducted between the Cheater vs. Non-Cheater groups to examine whether there would be a difference in cheating behaviour. While the results of this analysis should be interpreted with caution because of the small sample size ($n = 33$) and low frequency counts in the cross-tabulation, the results were non-significant, $\chi^2 (1, n=33) = .76, p = .38, r = .15$. Given the small sample who completed the follow-up survey, there was only enough observed statistical power (14%) to detect a large effect size⁹. Frequency of cheating behaviour across the two groups is presented below in *Table 6*.

Table 6. Frequency of Cheating Behaviour across the Time Since Graduation Groups

	Less than 5 years since graduating	More than 5 years since graduating	Totals
Cheater Group	$n = 4$ (23.5%)	$n = 6$ (37.5%)	$n = 10$ (30.3%)
Non-Cheater Group	$n = 13$ (76.5%)	$n = 10$ (62.5%)	$n = 23$ (69.7%)
Totals	$n = 17$ (51.5%)	$n = 16$ (48.5%)	$n = 33$ (100%)

Hypothesis 3: Relationship between impression management and cheating. It was hypothesized that job candidates who cheat may also try to present themselves more favourably on a measure of impression management. An independent-samples *t*-test was conducted to compare Impression Management across the Cheater and Non-Cheater groups. Levene's test indicated equal variances ($F(37) = 3.54, p = .07$). Contrary to the hypothesis, the Cheater group

⁹ The statistical power to identify a large effect size ($\Phi=0.5$) with the current sample ($n=33$) was 82%. In other words, there was an 18% chance that this analysis would find a nonsignificant result if in fact there was a true difference in the population.

($n = 14$, $M = 5.02$, $SD = 1.16$) did not significantly differ from the Non-Cheater group ($n = 25$, $M = 5.25$, $SD = 0.75$) on impression management: $t(37) = 0.74$, $p = 0.47$, $d = .25$. A possible explanation for this finding is that the BIDR – IM scale was not measured in the context of the hiring process but was included in the online questionnaire which participants knew would not inform the hiring process.

Hypothesis 4: Relationship between Theory of Planned Behaviour (TPB) constructs and cheating. The TPB constructs were measured using the PACES 2 scales. I hypothesized that attitudes, subjective norms, feelings of moral obligation, and perceived behavioural control would relate to cheating behaviour such that cheaters would have more lenient cheating attitudes and norms, and a stronger sense of behavioural control over cheating. Only one PACES 2 subscale – Perceived Behavioural Control (PBC) – had a distribution which satisfied the assumption of normality and homogeneity of variances. An independent-samples t -test was conducted to compare participants' PBC across the Cheater and Non-Cheater participant groups. Levene's test indicated equal variance ($F(37) = 0.86$, $p = .36$). Contrary to the hypothesis that PBC would increase cheating behaviour, the Cheater group ($n = 14$, $M = 2.59$, $SD = 0.92$) did not significantly differ from the Non-Cheater group ($n = 25$, $M = 2.22$, $SD = 1.02$) on PBC: $t(37) = -1.12$, $p = 0.27$, $d = .38$. However, the medium effect size for this analysis should be noted. Because of the very small and unequal sample size, it is crucial that this result be interpreted with caution.

I conducted non-parametric Mann-Whitney U Tests of independence¹⁰ for the remaining PACES 2 subscales which indicated non-normal distributions: Intentions, Attitudes, Subjective

¹⁰ The Mann-Whitney U test is an appropriate test when the distribution of scores violate the assumption of normality.

Norms, and Moral Obligation. Because my hypotheses were directional, I divided the Mann-Whitney Exact p value by 2 to calculate the one-tailed significance value of the test¹¹. The Mann-Whitney test indicated that intentions to cheat were stronger for cheaters ($Mdn = 1.5$) than for non-cheaters, $Mdn = 1.0$, $U = 109.0$, $p = .03$, $r = .34$. Similarly, the Mann-Whitney test indicated that cheaters held more lenient attitudes toward cheating ($Mdn = 1.64$) than non-cheaters, $Mdn = 1.43$, $U = 114.5$, $p = .04$, $r = .29$. On the other hand, the Mann-Whitney test indicated there were no significant differences in subjective norms across the cheater ($Mdn = 1.93$) and non-cheater ($Mdn = 1.29$) groups, $U = 123.0$, $p = .07$, $r = .25$. Finally, there was no significant difference between the Cheater ($Mdn = 1.00$) and Non-Cheater ($Mdn = 1.00$) groups on the construct of Moral Obligation, $U = 129.0$, $p = .09$, $r = .28$.

Hypothesis 5: Impact of individual difference factors on cheating behaviour. A review of the literature on academic cheating behaviour suggests that cheaters score lower on cognitive ability (Bloodgood et al., 2008), conscientiousness, and agreeableness (Giluk & Postlethwaite, 2015). The OCEAN.20 Conscientiousness scale and the WPT-Q cognitive test were found to have normal distributions for my sample that were appropriate for group comparisons. On the other hand, the OCEAN.20 Agreeableness scale distribution was highly skewed and produced a non-normal distribution. Two separate independent-samples t -tests were conducted to compare participants on the measures of cognitive ability and conscientiousness across the Cheater and Non-Cheater participant groups. There was a significant difference with a large effect size in the WPT-Q cognitive ability scores across the Cheater ($n = 33$, $M = 18.82$, $SD = 4.68$) and Non-Cheater ($n = 68$, $M = 21.57$, $SD = 4.68$) groups: $t(99) = 2.77$, $p = .007$, $d = .59$, such that Cheaters as a group scored lower on cognitive ability than Non-Cheaters. Levene's

¹¹ I calculated the effect size for each test with the following formula: $r = Z / (\text{sqrt of } N)$.

test indicated equal variance ($F(99) = 0.064, p = .80$). On the other hand, there was no significant difference in the OCEAN.20 Conscientiousness scores across the Cheater ($n = 14, M = 5.73, SD = 0.87$) and Non-Cheater ($n = 26, M = 5.73, SD = 0.94$) groups: $t(38) = -0.005, p = .99, d = .001$. Levene's test indicated equal variance ($F(38) = 0.32, p = .58$). Of note, the Conscientiousness analysis had a much smaller sample size but the means for both groups were equivalent.

I conducted a non-parametric Mann-Whitney U Test of independence on the OCEAN.20 Agreeableness subscale which had indicated a non-normal distribution. The Mann-Whitney test indicated that the personality construct of Agreeableness did not differ significantly across the Cheater ($Mdn = 6.63$) and Non-Cheater ($Mdn = 6.25$) groups, $U = 157.00, p = .25, r = .11$.

Discussion

Although several studies have examined cheating within the context of personnel selection (see Kantrowitz & Dainis, 2014 for a review), only a few studies have rigorously examined cheating behaviour in an ongoing recruitment context and with the use of an experimental research design. Specifically, previous research has relied mainly on self-reported cheating data or archival test data. Even fewer published studies have examined the relation between cheating in personnel contexts and other individual factors. This Master's thesis addresses a gap in the unproctored Internet testing literature and has implications regarding the fairness and validity of using such tests in an employment testing context. Furthermore, exploring the relations between individual difference factors and cheating behaviour will provide practitioners with useful information on how they should design and implement UIT programs for different job groups and positions.

Several hypotheses with regards to Unproctored Internet Testing (UIT) were examined for the purpose of replicating past research, as well as exploring novel relationships which could extend the research in the area of personnel selection and testing. The primary research objective of my Master's thesis was to examine whether cheating deterrence strategies actually reduced cheating behaviour. To ensure the cheating deterrence strategies I deployed were effective, I targeted the psychological dimensions which the Theory of Planned Behaviour (TPB) has shown to influence subsequent behaviour and/or behaviour change. To ensure the generalizability of my findings to applied settings, I modeled the deterrence strategies deployed in my research to follow a similar format and flow commonly used by testing specialists¹².

Contrary to UIT best practice and my hypothesis that cheating deterrence would reduce actual cheating behaviour, results from my study question the effectiveness of typically deployed cheating deterrence strategies such as warning messages and honesty pledges. A possible explanation for this finding is that the method of cheating deterrence used in my Master's thesis, as well as in general practice, is insufficient to create behaviour change because of the format (e.g., text/written), timing (e.g., immediate/right before the testing process), and context (e.g., impersonal/internet) of the messages. For example, the question remains as to whether there would have been a stronger impact if the cheating deterrence messages were delivered in video format and the pledge was made in-person or digitally written to a specific person a few days before the actual exam.

Based on the analyses of the Theory of Planned Behaviour (TPB) constructs, there is still hope for the use of UIT in personnel selection settings. Specifically, intentions to cheat and

¹² The author has first-hand experience working in an applied personnel testing setting and has conducted numerous research studies and projects related to psychological testing and assessment in personnel settings.

lenient attitudes toward cheating were significant predictors of actual cheating behaviour, while the impact of perceived behavioural control, although non-significant statistically, appeared to have practical significance through a moderate effect size. The TPB construct Subjective Norms – which previous academic cheating research had identified a strong relation with cheating – did not appear to have a statistically significant relation with the type of cheating measured in this study.

A secondary yet equally important objective of my study was to explore individual differences and their relation to cheating behaviour. Two novel hypotheses which were formulated to extend the research in UIT and cheating were not supported by the data collected through my Master's thesis. It was hypothesized that time since graduation from college or university would play a significant role in cheating outcomes. The rationale for this hypothesis was based on the TPB construct of subjective norms. I hypothesized that if the candidates were recent graduates, then their circle of influence were mainly students who may regard cheating more leniently or even normatively. However, similar to the finding regarding subjective norms, time since graduation did not predict cheating behaviour in this sample of job applicants. Finally, previous research had suggested that conscientiousness, agreeableness, and impression management were related to cheating behaviour as demonstrated in various experimental and academic settings. The analysis conducted for this Master's thesis did not identify significant differences between cheaters and non-cheaters across these individual difference variables.

The finding that Cheaters scored significantly lower on the WPT-Q cognitive ability test in the context of personnel selection is a significant and relevant finding. UIT and personnel selection research have rarely (if ever) focused on the difference between cognitive ability levels across cheater and non-cheater groups. The relation between intelligence and cheating was

uncovered through the literature review on academic cheating research. This finding has several implications, but the most critical one is related to on-the-job performance. Recruiters deploy cognitive tests to ensure that new recruits have the minimum cognitive ability level required to carry out the job. If individuals with lower cognitive ability are more prone to cheating, then it follows that some individuals with lower intelligence will be placed into roles that exceed their limit of competence. This misalignment could lead to turnover – a costly outcome for the hiring organization - and/or safety issues if the job carries risk to self or others. An immediate question that comes to mind is why Cheaters as a group scored lower on the WPT-Q cognitive ability test if they cheated on the Cheating Detection Test that was administered just prior to it. One might assume that they would have found a way to cheat on the WPT-Q as well. One possible explanation is that the type of cheating captured in the Cheating Detection Test is within the category of using assisted/restricted technology. There is no opportunity to cheat with assisted technology on the WPT-Q because the test questions are problem-based and timed, and do not include a memory test component. Currently, assisted technology cannot be used or manipulated to solve these types of problems.¹³ Furthermore, the test vendors who distribute the WPT-Q ensure that no items are leaked or available online and randomly administer different test forms to prevent cheating from repetition and exposure. Job applicants may have cheated on the WPT-Q by using a surrogate test taker but this does not mean the surrogate would necessarily score any better than the actual test taker. Furthermore, acquiring a surrogate would seem like an effortful form of cheating that involves the disclosure of something most individuals, even if they are cheaters, may not feel comfortable doing. The ease and method of cheating on the CDT

¹³ While this is a limitation of current technology, developments in the field of artificial intelligence and machine learning may deem this limitation obsolete in a few years.

may also explain the higher than expected frequency of cheating (32.4%) across the control and treatment groups.

Limitations. The current study is limited by several factors. Results for the between group differences between Cheaters and Non-Cheaters, especially for the TPB constructs and individual difference factors, might have resulted in more statistically significant findings if the sample size was large enough. It is also difficult to interpret the 32.4% cheating rate based on the sample size, an unusually high rate according to previous research. With a large enough sample size, it is very likely that this rate could have regressed toward the true mean. It is important to note, however, that job applicant pools are typically not that large to begin with. Unless there is a significant number of positions being hired for, most applicant pools are not that large, especially for specialized or professional roles. Hence, from a practical/applied perspective, the size of the sample for this study can be generalized to what we commonly experience in real world recruitment settings.

Another limitation is regarding my analysis of the individual difference factors, specifically conscientiousness, agreeableness, and impression management. If these variables were measured as part of the hiring process – and not in a subsequent survey as in the current study – additional relations may have been uncovered because of score inflation. In other words, if job candidates felt that the results of the survey questionnaires would influence the hiring decision, they might have tried to respond in more socially desirable ways. The fact that participants did not respond in socially desirable ways suggests that true scores of personality and impression management are not related to cheating behaviour, at least in this sample of participants.

Implications and Direction for Future Research. Future research may want to investigate the differences between types of deterrence strategies, using context and timing to strengthen the effectiveness of the message. If attitudes and intentions to cheat are predictive of cheating behaviour, then changing the behaviour through well-thought-out deterrence strategies is still possible. While testing and assessment professionals have expended significant energy on the development of innovative exam questions, less attention has been given to innovations in cheating prevention. In the context of UIT, the emerging challenge is to deliver persuasive and convincing cheating deterrence messages, those that target attitude and intention change specifically. Based on the findings of my Master's thesis, programs targeting the moral undertone of cheating behaviour may not demonstrate the same level effectiveness. The same implications apply to Subjective Norms, as both constructs were unrelated to cheating behaviour in this sample.

The academic cheating research reviewed for this Master's thesis also provides an evidence-based framework through which UIT applications can be audited and examined. Although the data collected for this Master's thesis did not support the hypothesis that recent graduates are more likely to cheat, other individual difference factors that have been demonstrated to correlate with cheating among student samples (e.g., intelligence) showed a significant relation to cheating and organizations should be made aware that this relationship exists. Applying what we know through the TPB, practitioners should design UIT programs and administration strategies that target TPB constructs in more innovative ways – perhaps utilizing online video technology and other tools that will demonstrate greater effectiveness.

Conclusion. The findings of this Master's thesis will have implications for UIT research and personnel testing. Most importantly, it is anticipated that the findings of this research will

initiate scholarly discussion regarding the fairness and effectiveness of UIT. The unusually high cheating frequency (32.4%) captured in this research indicates that cheating behaviour is underreported in self-report studies. As such, testing practitioners should not rely too heavily on previous UIT studies to anticipate the frequency of cheating behaviour in their UIT programs. Researchers should also explore cheating detection methods to understand more accurately the frequency of cheating on unproctored Internet exams, especially for assessments in which technology can be used to cheat. It is also evident that more research is needed to develop innovative and effective cheating deterrence strategies. Finally, future research might want to put more focus on the role of cognitive ability in understanding and predicting cheating behaviour.

The literature review conducted for this Master's thesis also uncovered a wide gap in the published literature between academic cheating and UIT cheating research. Although cheating behaviour in the academic setting has been researched for over 100 years, some of the main theories (e.g., TPB) and consistent research findings (e.g., individual characteristics related to cheating) have not been applied or utilized in UIT research. Indeed, this difference between the two streams of research need not exist, especially because student samples are the main target of online pre-employment tests upon or even before graduation. Decades of academic cheating research should not be overlooked in the context of personnel selection research. Hence, my Master's thesis provides one step forward in the convergence of these two streams of research and the testing of one critical framework used in academic research, the TPB, in a personnel selection context.

References

- Ajzen, I. (1988). *Attitudes, personality, and behavior*. Milton, Keynes: Open University Press.
- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50(2), 179-211.
- Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behaviour*. Englewood Cliffs, NJ: Prentice-Hall.
- Ajzen, I., & Madden, T. (1986). Prediction of goal directed behaviour: Attitudes, intentions and perceived behavioural control. *Journal of Experimental Social Psychology*, 22, 453-474.
- Alleyne, P., & Phillips, K. (2011). Exploring academic dishonesty among university students in Barbados: an extension to the theory of planned behaviour. *Journal of Academic Ethics*, 9(4), 323-338.
- Alt, D. (2014). Assessing the connection between students' justice experience and attitudes toward academic cheating in higher education new learning environments. *Journal of Academic Ethics*, 12(2), 113-127.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, & Joint Committee on Standards for Educational and Psychological Testing. (2014). *Standards for educational and psychological testing*. Washington, DC: AERA.
- Armitage, C. J., & Conner, M. (2001). Efficacy of the theory of planned behaviour: A meta-analytic review. *British Journal of Social Psychology*, 40(4), 471-499.

- Arthur, W., Glaze, R. M., Villado, A. J., & Taylor, J. E. (2009). Unproctored Internet-based tests of cognitive ability and personality: Magnitude of cheating and response distortion. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 2(1), 39-45. doi: 10.1111/j.1754-9434.2008.01105.x
- Arthur, W., Glaze, R. M., Villado, A. J., & Taylor, J. E. (2010). The magnitude and extent of cheating and response distortion effects on unproctored Internet-based tests of cognitive ability and personality. *International Journal of Selection and Assessment*, 18(1), 1-16.
- Beck, L., & Ajzen, I. (1991). Predicting dishonest actions using the theory of planned behavior. *Journal of Research in Personality*, 25(3), 285-301.
- Bilker, W. B., Hansen, J. A., Brensinger, C. M., Richard, J., Gur, R. E., & Gur, R. C. (2012). Development of abbreviated nine-item forms of the Raven's Standard Progressive Matrices Test. *Assessment*, 19(3), 354-369.
- Bloodgood, J. M., Turnley, W. H., & Mudrack, P. (2008). The influence of ethics instruction, religiosity, and intelligence on cheating behavior. *Journal of business ethics*, 82(3), 557-571.
- Buckley, M. R., Wiese, D. S., & Harvey, M. G. (1998). An investigation into the dimensions of unethical behavior. *Journal of Education for Business*, 73(5), 284-290. doi: 10.1080/08832329809601646
- Burke, E. (2009). Preserving the integrity of online testing. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 2(1), 35-38. doi: 10.1111/j.1754-9434.2008.01104.x

Burke, M. J., & Normand, J. (1987). Computerized psychological testing: Overview and critique.

Professional Psychology: Research and Practice, 18(1), 42-51. doi:

<http://dx.doi.org/10.1037/0735-7028.18.1.42>

Callahan, D. (2004). *The cheating culture: Why more Americans are doing wrong to get ahead*.

Orlando, FL: Harcourt.

Carpenter, D. D., Harding, T. S., Finelli, C. J., Montgomery, S. M., & Passow, H. J. (2006).

Engineering students' perceptions of and attitudes towards cheating. *Journal of Engineering Education*, 95(3), 181-194.

Carpenter, D. D., Harding, T. S., Finelli, C. J., & Passow, H. J. (2004). Does academic

dishonesty relate to unethical behavior in professional practice? An exploratory study.

Science and Engineering Ethics, 10(2), 311-324.

Carstairs, J., & Myors, B. (2009). Internet testing: A natural experiment reveals test score

inflation on a high-stakes, unproctored cognitive test. *Computers in Human Behavior*, 25(3), 738-742. doi: 10.1016/j.chb.2009.01.011

Coyne, S. M., & Thomas, T. J. (2008). Psychopathy, aggression, and cheating behavior: A test of

the Cheater-Hawk hypothesis. *Personality and Individual Differences*, 44(5), 1105-1115.

doi: <http://dx.doi.org/10.1016/j.paid.2007.11.002>

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of

psychopathology. *Journal of Consulting Psychology*, 24(4), 349-354.

- Davis, S. F., Grover, C. A., Becker, A. H., & McGregor, L. N. (1992). Academic dishonesty: Prevalence, determinants, techniques, and punishments. *Teaching of Psychology, 19*(1), 16-20.
- de Vries, R. E., Zettler, I., & Hilbig, B. E. (2014). Rethinking trait conceptions of social desirability scales: Impression management as an expression of honesty-humility. *Assessment, 21*(3), 286-299.
- DeAndrea, D. C., Carpenter, C., Shulman, H., & Levine, T. R. (2009). The relationship between cheating behavior and sensation-seeking. *Personality and Individual Differences, 47*(8), 944-947. doi: <http://dx.doi.org/10.1016/j.paid.2009.07.021>
- Detrick, P., Chibnall, J. T., & Call, C. (2010). Demand effects on positive response distortion by police officer applicants on the revised NEO personality inventory. *Journal of Personality Assessment, 92*(5), 410-415. doi: 10.1080/00223891.2010.497401
- Donnellan, M. B., Oswald, F. L., Baird, B. M., & Lucas, R. E. (2006). The mini-IPIP scales: Tiny-yet-effective measures of the Big Five factors of personality. *Psychological Assessment, 18*(2), 192-203.
- Gallant, T. B., & Drinan, P. (2006). Organizational theory and student cheating: Explanation, responses, and strategies. *The Journal of Higher Education, 77*(5), 839-860.
- Gilligan, C. (1982). In a different voice: Psychological theory and women's development. Cambridge, MA: Harvard University Press.

Giluk, T. L., & Postlethwaite, B. E. (2015). Big Five personality and academic dishonesty: A meta-analytic review. *Personality and Individual Differences*, 72, 59-67.

Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Deary, F. De Fruyt & F. Ostendorf (Eds.), *Personality Psychology in Europe* (Vol. 7, pp. 7-28). Tilburg, The Netherlands: Tilburg University Press.

Government of Canada. Post Secondary Recruitment – Tests Retrieved March 15, 2017, 2017, from <https://www.canada.ca/en/public-service-commission/jobs/services/recruitment/graduates/post-secondary-recruitment/tests.html>

Greenberg, C. I. (1999). Technological innovations and advancements for psychologists working with organizations. *The Psychologist-Manager Journal*, 3(2), 181-190. doi: <http://dx.doi.org/10.1037/h0095868>

Guo, J., & Drasgow, F. (2010). Identifying cheating on unproctored Internet tests: The Z-test and the likelihood ratio test. *International Journal of Selection and Assessment*, 18(4), 351-364.

Harding, T. S., Carpenter, D. D., Finelli, C. J., & Passow, H. J. (2004). *The influence of academic dishonesty on ethical decision making in the workplace: A study of engineering students*. Paper presented at the ASEE Annual Conference and Exposition, Salt Lake City, UT.

- Harding, T. S., Mayhew, M. J., Finelli, C. J., & Carpenter, D. D. (2007). The theory of planned behavior as a model of academic dishonesty in engineering and humanities undergraduates. *Ethics & Behavior*, 17(3), 255-279.
- Harper, M. G. (2006). High tech cheating. *Nurse Education in Practice*, 6(6), 364-371.
- Harris, G. (1916). The ethics of college students. *Harvard Theological Review*, 9(02), 190-200.
- Hense, R., Golden, J. H., & Burnett, J. (2009). Making the case for unproctored Internet testing: Do the rewards outweigh the risks? *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 2(1), 20-23. doi: 10.1111/j.1754-9434.2008.01100.x
- Hogan, R., & Hogan, J. (1995). *Hogan personality inventory manual*. Tulsa, OK: Hogan Assessment Systems. [retrieved from <http://www.hoganassessments.com/sites/default/files/HPI%20Tech%20Manual%20-%20S.pdf>]
- Hollister, K. K., & Berenson, M. L. (2009). Proctored versus unproctored online exams: Studying the impact of exam environment on student performance. *Decision Sciences Journal of Innovative Education*, 7(1), 271-294. doi: 10.1111/j.1540-4609.2008.00220.x
- Hylton, K., Levy, Y., & Dringus, L. P. (2016). Utilizing webcam-based proctoring to deter misconduct in online exams. *Computers & Education*, 92-93(January-February 2016), 53-63. doi: <http://dx.doi.org/10.1016/j.compedu.2015.10.002>

- Imran, A. M., & Nordin, M. S. (2013). Predicting the underlying factors of academic dishonesty among undergraduates in public universities: A path analysis approach. *Journal of Academic Ethics, 11*(2), 103-120.
- Kaminski, K. A., & Hemingway, M. A. (2009). To proctor or not to proctor? Balancing business needs with validity in online assessment. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 2*(1), 24-26. doi: 10.1111/j.1754-9434.2008.01101.x
- Kantrowitz, T. M., & Dainis, A. M. (2014). How secure are unproctored pre-employment tests? Analysis of inconsistent test scores. *Journal of Business and Psychology, 29*(4), 605-616.
- Kleinmuntz, B., & McLean, R. S. (1968). Diagnostic interviewing by digital computer. *Behavioral Science, 13*(1), 75-80.
- Landers, R. N., & Sackett, P. R. (2012). Offsetting performance losses due to cheating in unproctored Internet-based testing by increasing the applicant pool. *International Journal of Selection and Assessment, 20*(2), 220-228.
- Lievens, F., & Burke, E. (2011). Dealing with the threats inherent in unproctored Internet testing of cognitive ability: Results from a large-scale operational test program. *Journal of Occupational and Organizational Psychology, 84*(4), 817-824. doi: 10.1348/096317910x522672
- Lievens, F., & Harris, M. M. (2003). Research on Internet recruiting and testing: Current status and future directions. *International Review of Industrial and Organizational Psychology, 18*, 131-166.

- Lonsdale, D. (2017). Intentions to cheat: Ajzen's planned behavior and goal-related personality facets. *The Journal of Psychology, 151*(2), 113-129. doi: 10.1080/00223980.2016.1241737
- Mayhew, M. J., Hubbard, S. M., Finelli, C. J., Harding, T. S., & Carpenter, D. D. (2009). Using structural equation modeling to validate the theory of planned behavior as a model for predicting student cheating. *The Review of Higher Education, 32*(4), 441-468.
- McCabe, D. L., & Trevino, L. K. (1993). Academic dishonesty: Honor codes and other contextual influences. *The Journal of Higher Education, 64*(5), 522-538.
- McCabe, D. L., & Trevino, L. K. (1997). Individual and contextual influences on academic dishonesty: A multicampus investigation. *Research in Higher Education, 38*(3), 379-396.
- McCabe, D. L., Trevino, L. K., & Butterfield, K. D. (2001). Cheating in academic institutions: A decade of research. *Ethics & Behavior, 11*(3), 219-232.
- Millham, J. (1974). Two components of need for approval score and their relationship to cheating following success and failure. *Journal of Research in Personality, 8*(4), 378-392.
- Nesnidol, S. (2016). *Public sector perceptions of unproctored Internet testing*.
- Nonis, S., & Swift, C. O. (2001). An examination of the relationship between academic dishonesty and workplace dishonesty: A multicampus investigation. *Journal of Education for Business, 77*(2), 69-77.

- Nye, C. D., Do, B.-R., Drasgow, F., & Fine, S. (2008). Two-step testing in employee selection: Is score inflation a problem? *International Journal of Selection and Assessment*, 16(2), 112-120.
- O'Keefe, D. F., Kelloway, E. K., & Francis, R. (2012). Introducing the OCEAN. 20: A 20-Item five-factor personality measure based on the trait self-descriptive inventory. *Military Psychology*, 24(5), 433-460.
- Paulhus, D. L. (1988). Assessing self deception and impression management in self-reports: The Balanced Inventory of Desirable Responding.
- Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 17-59). San Diego, CA: Academic Press.
- Pearlman, K. (2009). Unproctored Internet testing: Practical, legal, and ethical concerns. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 2(1), 14-19. doi: 10.1111/j.1754-9434.2008.01099.x
- Reynolds, D. H., Wasko, L. E., Sinar, E. F., Raymark, P. H., & Jones, J. A. (2009). UIT or not UIT? That is not the only question. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 2(1), 52-57. doi: 10.1111/j.1754-9434.2008.01108.x
- Rios, J. A., & Liu, O. L. (2017). Online proctored versus unproctored low-stakes Internet test administration: Is there differential test-taking behavior and performance? *American Journal of Distance Education*, 1-14. doi: 10.1080/08923647.2017.1258628

Ryan, A. M., & Ployhart, R. E. (2014). A century of selection. *Annual Review of Psychology*, 65(20.1), 693-717.

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124(2), 262-274.

Schmitt, N., Gilliland, S. W., Landis, R. S., & Devine, D. (1993). Computer-based testing applied to selection of secretarial applicants. *Personnel Psychology*, 46(1), 149-165.

Scott, J. C., & Lezotte, D. V. (2012). Web-based assessments. In N. Schmitt (Ed.), *The Oxford handbook of personnel assessment and selection* (pp. 485-513). New York, NY: OUP.

Spiller, M. S., & Crown, D. F. (1995). Changes over time in academic dishonesty and unethical business practices. *Journal of Education for Business*, 68(4), 207-211.

Stogner, J. M., Miller, B. L., & Marcum, C. D. (2013). Learning to e-cheat: A criminological test of Internet facilitated academic cheating. *Journal of Criminal Justice Education*, 24(2), 175-199.

Stone, T. H., Jawahar, I. M., & Kisamore, J. L. (2009). Using the theory of planned behavior and cheating justifications to predict academic misconduct. *Career Development International*, 14(3), 221-241.

Stone, T. H., Jawahar, I. M., & Kisamore, J. L. (2010). Predicting academic misconduct intentions and behavior using the theory of planned behavior and personality. *Basic and Applied Social Psychology*, 32(1), 35-45.

Unproctored Internet Testing

- Tang, S., & Zuo, J. (1997). Profile of college examination cheaters. *College Student Journal*, 31, 340-346.
- Templer, K. J., & Lange, S. R. (2008). Internet testing: Equivalence between proctored lab and unproctored field conditions. *Computers in Human Behavior*, 24(3), 1216-1228. doi: 10.1016/j.chb.2007.04.006
- Tendeiro, J. N., Meijer, R. R., Schakel, L., & Maij-de Meij, A. M. (2013). Using cumulative sum statistics to detect inconsistencies in unproctored Internet testing. *Educational and Psychological Measurement*, 73(1), 143-161. doi: 10.1177/0013164412444787
- Tippins, N. T. (2009a). Internet alternatives to traditional proctored testing: Where are we now? *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 2(1), 2-10. doi: 10.1111/j.1754-9434.2008.01097.x
- Tippins, N. T. (2009b). Where is the unproctored Internet testing train headed now? *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 2(1), 69-76. doi: 10.1111/j.1754-9434.2008.01111.x
- Tippins, N. T., Beaty, J., Drasgow, F., Gibson, W. M., Pearlman, K., Segall, D. O. (2006). Unproctored Internet testing in employment settings. *Personnel Psychology*, 59(1), 189-225. doi: 10.1111/j.1744-6570.2006.00909.x
- Vinski, E. J., & Tryon, G. S. (2009). Study of a cognitive dissonance intervention to address high school students' cheating attitudes and behaviors. *Ethics & Behavior*, 19(3), 218-226.

- Weiner, J. A., & Morrison Jr, J. D. (2009). Unproctored online testing: Environmental conditions and validity. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 2(1), 27-30. doi: 10.1111/j.1754-9434.2008.01102.x
- Whitley, B. E., Nelson, A. B., & Jones, C. J. (1999). Gender differences in cheating attitudes and classroom cheating behavior: A meta-analysis. *Sex Roles*, 41(9-10), 657-680.
- Wright, N. A., Meade, A. W., & Gutierrez, S. L. (2014). Using invariance to examine cheating in unproctored ability tests. *International Journal of Selection and Assessment*, 22(1), 12-22.
- Zettler, I., Hilbig, B. E., Moshagen, M., & de Vries, R. E. (2015). Dishonest responding or true virtue? A behavioral test of impression management. *Personality and Individual Differences*, 81, 107-111.

Appendices

Appendix A

Subject: Invitation to participate in a research project on internet-based employment testing.

Dear Sir or Madam,

In response to your application for [Job Title] at [Company Name] we are also pleased to invite you to participate in an online research study that is conducted by one of our research consultants as part of his Master's thesis. Please see his letter below.

Sincerely,
[Company Name] Recruitment Team

Dear Sir or Madam,

My name is Emrah Eren and I am a Master's student in the Psychology Department at Carleton University. I am working on a research project under the supervision of Prof. Timothy A. Pynchl.

I am writing to you today to invite you to participate in a study entitled "Internet-based Employment Testing: Job Candidate Behaviours and Attitudes". The purpose of this research is to study the impact of online employment testing on job candidate behaviour and attitudes.

This study involves the completion of a series of psychological measures and questionnaires regarding your attitudes and behaviours and will take approximately 15-20 minutes to complete. You can complete the survey at a location of your choice. You will need access to a computer that is connected to the internet. We also ask that you grant us permission to use your [COMPANY NAME] pre-employment testing data from the general abilities test that you are scheduled to take.

Strong measures will be taken to protect your identity in this research. The personal information you choose to disclose in response to the questionnaires such as your gender, ethnicity, and age will **NOT** be shared with [COMPANY NAME] or any third party. **The results of this research project and/or your decision to participate will not in any way be related to (or impact) your job competition process with [COMPANY NAME].** Your identifiable information will be held in complete confidence by the primary researcher and will not inform the hiring process or decision. This will be done by keeping all responses confidential and stored on an encrypted drive and allowing you to request that certain responses not be included in the final project. Again, this research is being conducted separately from [COMPANY NAME]'s recruitment process and you can withdraw from the research at any time without consequences

Several measures will be taken to keep your survey responses secure, and to ensure that no one

Unproctored Internet Testing

from [COMPANY NAME] will have access to your data. All of your responses will be stored on a survey platform that no one from [COMPANY NAME] other than the primary researcher will have access to. No identifiable information will be linked to your survey responses and the file linking your identifiable information to your unique participant code will be destroyed immediately after data collection. All other information pertaining to your results will be stored on an encrypted drive on the researcher's personal computer.

You will have the right to end your participation in the study at any time, for any reason, up until February 15, 2018. If you choose to withdraw, all the information you have provided will be destroyed.

In appreciation of your time and efforts in completing this questionnaire, you will be provided with the option to enter a draw for a chance to win a cash prize: First prize-\$100; second prize-\$50; and third prize-\$25. Approximately 150 participants will be entered into a draw for 1 of 3 cash prizes. (You will be entered into the draw and the compensation is yours to keep, even if you choose to withdraw.)

All research data will be encrypted. Research data will only be accessible by the researcher and the research supervisor.

The ethics protocol for this project was reviewed by the Carleton University Research Ethics Board, which provided clearance to carry out the research. (Clearance expires on: insert date here.)

CUREB-B:

If you have any ethical concerns with the study, please contact Dr. Andy Adler, Chair, Carleton University Research Ethics Board-B (by phone at 613-520-2600 ext. 4085 or via email at ethics@carleton.ca).

If you would like to participate in this research project, or have any questions, please click this [link](#), reply to this email, or contact emraheren@gmail.com

Sincerely,

Emrah Eren

NOTE: This must be sent from your official university email address.

Appendix B

Online Consent Form:

Title: Internet-based Employment Testing: Job Candidate Behaviours and Attitudes

Ethics Clearance #: Project # 107337

Date of ethics clearance: October 11, 2017

Ethics Clearance for the Collection of Data Expires: October 11, 2018

Purpose: This is a study on Internet-based employment testing. This study aims to study the impact of online employment testing on job candidate behaviour and attitudes. **The researcher for this study is Emrah Eren in the Psychology Department at Carleton University.**

Research Personnel: Emrah is working under the supervision of Timothy A. Pychyl in the Psychology Department of Carleton University.

Task requirements: This study involves the completion of psychological measures and questionnaires that will take approximately 15-20 minutes to complete. You can complete the survey at a location of your choice. You will need access to a computer that is connected to the internet. We also ask that you grant us permission to use your [COMPANY NAME] pre-employment testing data from the general abilities tests that you are scheduled to take.

Potential Risk/Discomfort: There are no known risks or discomforts associated with this type of research. Should you experience any discomfort during this research, please contact Emrah at 1-613-614-6853 or emraheren@cmail.carleton.ca.

The results of this research project and/or your decision to participate will not in any way be related to (or impact) your job competition process with [COMPANY NAME].

Permission to Use Data: By agreeing to participate in this research, you are granting the researcher of this study permission to use your data for publications, presentations, and teaching purposes. Strong measures will be taken to protect your identity in this research. The personal information you choose to disclose in response to the questionnaires such as your gender, ethnicity, and age will **NOT** be shared with [COMPANY NAME] or any third

party. Your identifiable information will be held in complete confidence by the primary researcher and will not inform the hiring process or decision. This will be done by keeping all responses confidential and stored on encrypted drives and allowing you to request that certain responses not be included in the final project. This research is being conducted separately from [COMPANY NAME]'s recruitment process. All data will be presented in aggregate and no identifiable information will be revealed publicly.

Anonymity, Confidentiality, and Right to withdraw: You have the right to end your participation in the survey or cognitive task at any time, for any reason, up until you hit the "submit" button. You can withdraw by exiting the survey at any time before completing it. If you withdraw from the study, all information you provided will be immediately destroyed. You may withdraw from the study by emailing the researcher by February 15, 2018.

All research data will be encrypted. The company running the online survey is Qualtrics based in the United States (US). The survey company will keep a copy of the survey responses on its servers in the US and this anonymous data will be subject to US laws on data privacy. This data will also be encrypted and will be deleted once the survey is complete. Research data will be accessible by the researcher, the research supervisor and the survey company. No names or IP addresses will be linked to any of the data provided.

Several measures will be taken to keep your survey responses secure, and to ensure that no one from [COMPANY NAME] will have access to your data. All of your responses will be stored on a survey platform that no one from [COMPANY NAME] other than the primary researcher will have access to. No identifiable information will be linked to your survey responses and the file linking your identifiable information to your unique participant code will be destroyed immediately after data collection. All other information pertaining to your results will be stored on an encrypted drive on the researcher's personal computer.

You will have the right to end your participation in the study at any time, for any reason, up until February 15, 2018. If you choose to withdraw, all the information you have provided will be destroyed. Once the project is completed, all research data will be kept for five years and potentially used for other research projects on this same topic. At the end of five years, all research data will be deleted.

Compensation: In appreciation of your time and efforts in completing this questionnaire, you will be provided with the option to enter a draw for a chance to win a cash prize: First prize-\$100; second prize-\$50; and third

prize-\$25. Approximately 150 participants will be entered into a draw for 1 of 3 cash prizes. (You will be entered into the draw and the compensation is yours to keep, even if you choose to withdraw.)

If you would like a copy of the finished research project, you are invited to contact the researcher to request an electronic copy which will be provided to you as long as the safety of all participants will not be comprised by doing so.

The ethics protocol for this project was reviewed by the Carleton University Research Ethics Board, which provided clearance to carry out the research.

CUREB-B contact info for Concerns:

If you have any ethical concerns with the study, please contact Dr. Andy Adler, Chair, Carleton University Research Ethics Board-B (by phone at 613-520-2600 ext. 4085 or via email at ethics@carleton.ca).

Researcher contact information:

Name: Emrah Eren
Department: Psychology
Carleton University
Tel: 1-613-614-6853
Email: emraheren@cmail.carleton.ca
TimPychyl@cunet.carleton.ca

Supervisor contact

Name: Timothy A. Pychyl
Department: Psychology
Carleton University
Tel: 613-520-2600 x 1403
Email:

By clicking "submit", you consent to participate in the research study as described above.

Appendix C

Cheating Detection Test

CDT Question 1

Page 1:

You will have 7 seconds to memorize the order of the following images...

Image displays for 7 seconds...

Page 2:

Place the images in the same order they were presented on the previous screen:

CDT Question 2

Page 3:

You will have 5 seconds to memorize the order of the following images:

Image displays for 5 seconds...

Page 4:

Place the images in the same order they were presented on the previous screen:

CDT Question 3:

Page 5:

You will have 3 seconds to memorize the order of the following images:

Image displays for 3 seconds...

Page 6:

Place the images in the same order they were presented on the previous screen;

CDT Question 4

Page 7:

You will have 3 seconds to memorize the order of the following images:

Image displays for 3 seconds...

Page 8:

Place the images in the same order they were presented on the previous screen:

Appendix D

The OCEAN.20 Occupational Personality Questionnaire¹

This questionnaire asks about a variety of subjects such as your attitudes towards other people, what you like doing, and how you would feel in particular circumstances. The inventory takes about 5 minutes to complete, however there is no time limit.

Please read the following instructions carefully.

On the following page, you will find a number of adjectives and statements. Using the following rating scale, decide how well each of them describes you.

1 - Extremely Uncharacteristic
2 - Quite Uncharacteristic
3 - Slightly Uncharacteristic
4 - Neither Characteristic Nor Uncharacteristic
5 - Slightly Characteristic
6 - Quite Characteristic
7 - Extremely Characteristic

Here is an example: 'Cautious'

Decide if 'cautious' is characteristic of you. If you decide 'cautious' is 'quite uncharacteristic' of you, then you should select the number 2, which corresponds to the 'quite uncharacteristic' rating.

Please reply to all adjectives and statements. Give your first impression of how characteristic each adjective and statement is of you. Don't spend too long on deciding what your answer should be. Answer all questions, even if you are not entirely sure of your answer.

Answer honestly. Please respond as honestly and accurately as you can.

¹ This survey was developed by O'Keefe, Kelloway and Francis (2012) and it based on the Trait Self-Descriptive Inventory-Personality Inventory, which is the property of the Director General Military Personnel Research and Analysis, Canadian Department of National Defence. Questions and/or comments can be addressed to damian.okeefe@smu.ca,

Unproctored Internet Testing

1		2		3		4		5		6		7		
Extremely Uncharacteristic		Quite Uncharacteristic		Slightly Uncharacteristic		Neither Characteristic Nor Uncharacteristic		Slightly Characteristic		Quite Characteristic		Extremely Characteristic		
1	Silent							1	2	3	4	5	6	7
2	Neat							1	2	3	4	5	6	7
3	Sympathetic							1	2	3	4	5	6	7
4	Organized							1	2	3	4	5	6	7
5	Withdrawn							1	2	3	4	5	6	7
6	Kind							1	2	3	4	5	6	7
7	Quiet							1	2	3	4	5	6	7
8	I have thought a lot about the origins of the universe							1	2	3	4	5	6	7
9	I like to keep all my belongings neat and organized							1	2	3	4	5	6	7
10	I often have headaches when things are not going well							1	2	3	4	5	6	7
11	I am always generous when it comes to helping others							1	2	3	4	5	6	7
12	Sometimes I get so upset, I feel sick to my stomach							1	2	3	4	5	6	7
13	I am highly interested in all fields of science							1	2	3	4	5	6	7
14	I like to have a place for everything and everything in its place							1	2	3	4	5	6	7
15	I am fascinated with the theory of evolution							1	2	3	4	5	6	7
16	When I am under great stress I often feel like I am about to break down							1	2	3	4	5	6	7
17	I always treat other people with kindness							1	2	3	4	5	6	7
18	My feelings are easily hurt							1	2	3	4	5	6	7
19	I am very shy person							1	2	3	4	5	6	7
20	I would enjoy being a theoretical scientist							1	2	3	4	5	6	7

Appendix E

Balanced Inventory Desirable Responding Version 6

Reverse scored items: 1,3,5,7,9,11,13,15,17,19.

Using the scale below as a guide, write a number beside each statement to indicate how true it is.

+	+	+	+	+	+	+
1	2	3	4	5	6	7
not true			somewhat			very true

- ___ 1. I sometimes tell lies if I have to.
- ___ 2. I never cover up my mistakes.
- ___ 3. There have been occasions when I have taken advantage of someone.
- ___ 4. I never swear.
- ___ 5. I sometimes try to get even rather than forgive and forget.
- ___ 6. I always obey laws, even if I'm unlikely to get caught.
- ___ 7. I have said something bad about a friend behind his/her back.
- ___ 8. When I hear people talking privately, I avoid listening.
- ___ 9. I have received too much change from a salesperson without telling him or her.
- ___ 10. I always declare everything at customs.
- ___ 11. When I was young I sometimes stole things.
- ___ 12. I have never dropped litter on the street.
- ___ 13. I sometimes drive faster than the speed limit.
- ___ 14. I never read sexy books or magazines.
- ___ 15. I have done things that I don't tell other people about.
- ___ 16. I never take things that don't belong to me.
- ___ 17. I have taken sick-leave from work or school even though I wasn't really sick.
- ___ 18. I have never damaged a library book or store merchandise without reporting it.
- ___ 19. I have some pretty awful habits.
- ___ 20. I don't gossip about other people's business.

Appendix F

The Perceptions and Attitudes toward Cheating among Engineering Students Survey
PACES -2 (adapted from Harding et al. (2007) and Stone, Jawahar, and Kisamore's (2010):
Modified for Personnel Selection Testing)

Instructions: Using the five point scale below, please indicate your level of agreement or disagreement with the items listed below.

1	2	3	4	5
Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

(R) Reverse Coded

(N) New item developed for this Master's thesis.

Intention

1. I will try to cheat on an employment test while seeking a new job.
2. I intend to cheat on an employment test while seeking a new job.
3. I do NOT plan to cheat on an employment test while seeking a new job. (R)
4. I will NOT cheat on an employment test while seeking a new job. (R)
5. If I had the opportunity, I would cheat on an employment test while seeking a new job.

Attitude toward cheating

6. It is always wrong to cheat. (R)
7. Cheating is necessary to level the playing field.
8. Job applicants should go ahead and cheat if they know they can get away with it.
9. It is necessary to report the cheating of others. (R)
10. I would let another job applicant cheat off my test if he/she asked. (N)
11. I would take an employment test for someone else if he/she asked. (N)
12. I would allow someone else take my employment test if I knew he/she could pass. (N)

Subjective Norms

13. If I cheated on an employment test, most of the people who are important to me (e.g., family, friends, colleagues, etc.) would approve my behaviour.
14. The people in my life whose opinions I value (e.g., my family, friends, colleagues, etc.) would be willing to cheat on an employment test if they were in my situation.
15. Most people who are important to me (e.g., my family, friends, colleagues, etc.) would be willing to cheat on an employment test if they were in my situation.
16. The people in my life whose opinions I value (e.g., my family, friends, colleagues, etc.) would NOT approve if I cheated on an employment test. (R)
17. Most people who are important to me (e.g., my family, friends, colleagues, etc.) think I should NOT cheat on an employment test. (R)
18. People whose opinions I value (e.g., my family, friends, colleagues, etc.) expect me to cheat on an employment test.
19. Most people who are important to me (e.g., my family, friends, colleagues, etc.) will look down on me if I cheat on an employment test. (R)

Unproctored Internet Testing

20. NO ONE who is important to me (e.g., my family, friends, colleagues, etc.) thinks it is OK to cheat on an in-class test or exam. (R)

Perceived Behavioural Control

21. If I wanted to cheat on an employment test, it would be easy.
22. It would be easy for me to cheat on the employment test I took during this job application.
23. I believe that I have the skills needed to cheat on an employment test.
24. It is mostly up to me whether or not I successfully cheat on an employment test.
25. Even if I wanted to, I could NOT cheat on an employment test. (R)
26. It is difficult to cheat and not get caught. (R)

Moral Obligation

27. Cheating on an employment test is against my principles. (R)
28. I would feel guilty if I cheated on an employment test. (R)
29. It would NOT be morally wrong for me to cheat on an employment test.

Appendix G

TPB Deterrence Strategy (Treatment Group Instructions)

Message to be included before the WPT-Q and CDT:

Instructions that target TPB Perceived Consequences

Cheating in any form will not be tolerated. Once completed, your assessment will be screened to detect patterns of cheating as some questions have been designed to report suspicious activity. You may also be invited to take a follow-up test in a proctored environment. No part of this online test may be reproduced, distributed, or transmitted in any form or by any means, including recording, scanning, or other electronic or mechanical methods. Unauthorized reproduction of any portion of this online test is a violation of copyright laws.

Instructions that target TPB Attitudes, Subjective Norms, Instructions that target TPB Perceived Behavioural Control, and Moral Obligation:

Cheating is dishonest behaviour that will disadvantage you, this personnel selection process, and others who are taking the test honestly by following the instructions. Numerous research studies have shown that most job applicants (90% - 95%) honestly complete their examinations when taking the test without human proctoring. Furthermore, it has been shown that it is very difficult to cheat on online tests that measure cognitive ability because there is rarely enough time to cheat, and when there is, these questions can be analyzed to detect cheating. Trying to cheat can even earn you lower scores than you would if you were taking the test honestly. It is simply easier to be true to yourself and your moral principles.

Also note that there could be a follow-up test administered to verify your scores.

Instructions that target TPB Intentions:

I, _____, pledge to take the online assessment honestly, by following the test administration instructions and without the aid of someone else or any assistive technology.

Appendix H
Demographic Questionnaire

Please respond to the following demographic questions. Your participation is voluntary and this information will **not** be used in the selection process. The information you provide will be kept strictly confidential and used only for the purpose of this Master's thesis research.

What is your age? (drop down)

16 to 65 and above

What is your gender?

- ☐ Male
- ☐ Female
- ☐ Other: _____

What is the highest level of education you completed?

- ☐ Community College
- ☐ University Undergraduate Degree
- ☐ University Graduate School Degree
- ☐ Other: _____

Are you currently a student?

- ☐ Yes, Full-time (skip next question)
- ☐ Yes, Part-time (skip next question)
- ☐ No

How long has it been since you graduated from your educational institution?

- ☐ 1 year or less
- ☐ 1-2 years
- ☐ 2-3 years
- ☐ 4-5 years
- ☐ 5 or more years

Do you identify as belonging to an Indigenous group?

- ☐ Yes
- ☐ No

Unproctored Internet Testing

Do you identify as a visible minority?

- ☐ Yes
☐ No

If you are a member of a visible minority group please indicate which ethnic heritage you identify with:

- ☐ *Black (Haitian, African, etc)*
- ☐ *East Asian (Chinese, Korean, etc)*
- ☐ *West Asian (Iranian, etc)*
- ☐ *South Asian (Indian, Pakistani, etc)*
- ☐ *South East Asian (Vietnamese, Filipino, Cambodian, etc)*
- ☐ *Middle Eastern (Lebanese, Syrian, etc)*
- ☐ *Latino/Hispanic (Mexican, Columbian, etc)*
- ☐ *Other: (please specify below)*

If you chose other, please specify which ethnic heritage group you identify with:

Do you identify as a person with a disability?

- ☐ Yes
☐ No

If yes, please indicate your disability:

- ☐ *Blind or Partially Sighted*
- ☐ *Hearing Impairment*
- ☐ *Deaf*
- ☐ *Mental Impairment*
- ☐ *Learning Impairment*
- ☐ *Other: (please specify below)*

If you chose other, please specify your disability:

From which Province or Territory are you taking this online test?

AB
BC
MB
NB
NL
NS
NT
NU

Unproctored Internet Testing

ON

PE

QC

SK

YT

Appendix I

Randomization Plan

100. Control Group _____
101. Treatement Group _____
102. Treatement Group _____
103. Treatement Group _____
104. Control Group _____
105. Control Group _____
106. Control Group _____
107. Treatement Group _____
108. Control Group _____
109. Control Group _____
110. Control Group _____
111. Treatement Group _____
112. Control Group _____
113. Control Group _____
114. Treatement Group _____
115. Treatement Group _____
116. Control Group _____
117. Control Group _____
118. Control Group _____
119. Treatement Group _____
120. Treatement Group _____
121. Control Group _____
122. Treatement Group _____
123. Control Group _____
124. Treatement Group _____
125. Treatement Group _____
126. Control Group _____
127. Treatement Group _____
128. Treatement Group _____
129. Control Group _____
130. Control Group _____
131. Treatement Group _____
132. Treatement Group _____
133. Control Group _____
134. Treatement Group _____
135. Treatement Group _____
136. Control Group _____
137. Control Group _____
138. Treatement Group _____
139. Treatement Group _____
140. Control Group _____
141. Treatement Group _____

Unproctored Internet Testing

142. Treatement Group _____
143. Treatement Group _____
144. Control Group _____
145. Control Group _____
146. Control Group _____
147. Control Group _____
148. Treatement Group _____
149. Treatement Group _____
150. Control Group _____
151. Treatement Group _____
152. Treatement Group _____
153. Control Group _____
154. Treatement Group _____
155. Treatement Group _____
156. Control Group _____
157. Control Group _____
158. Control Group _____
159. Treatement Group _____
160. Control Group _____
161. Treatement Group _____
162. Treatement Group _____
163. Control Group _____
164. Control Group _____
165. Treatement Group _____
166. Control Group _____
167. Control Group _____
168. Treatement Group _____
169. Control Group _____
170. Treatement Group _____
171. Control Group _____
172. Control Group _____
173. Treatement Group _____
174. Treatement Group _____
175. Treatement Group _____
176. Control Group _____
177. Control Group _____
178. Treatement Group _____
179. Control Group _____
180. Treatement Group _____
181. Control Group _____
182. Treatement Group _____
183. Control Group _____
184. Control Group _____
185. Control Group _____
186. Control Group _____
187. Control Group _____

Unproctored Internet Testing

188. Control Group _____
189. Control Group _____
190. Control Group _____
191. Treatement Group _____
192. Control Group _____
193. Control Group _____
194. Control Group _____
195. Control Group _____
196. Treatement Group _____
197. Treatement Group _____
198. Control Group _____
199. Treatement Group _____
200. Treatement Group _____
201. Control Group _____
202. Treatement Group _____
203. Treatement Group _____
204. Treatement Group _____
205. Control Group _____
206. Treatement Group _____
207. Treatement Group _____
208. Treatement Group _____
209. Treatement Group _____
210. Control Group _____
211. Control Group _____
212. Treatement Group _____
213. Treatement Group _____
214. Control Group _____
215. Control Group _____
216. Treatement Group _____
217. Treatement Group _____
218. Treatement Group _____
219. Control Group _____
220. Treatement Group _____
221. Treatement Group _____
222. Control Group _____
223. Treatement Group _____
224. Treatement Group _____
225. Treatement Group _____
226. Treatement Group _____
227. Control Group _____
228. Control Group _____
229. Treatement Group _____
230. Control Group _____
231. Treatement Group _____
232. Control Group _____
233. Treatement Group _____

Unproctored Internet Testing

234. Treatement Group _____
235. Treatement Group _____
236. Treatement Group _____
237. Treatement Group _____
238. Control Group _____
239. Treatement Group _____
240. Treatement Group _____
241. Control Group _____
242. Control Group _____
243. Treatement Group _____
244. Treatement Group _____
245. Control Group _____
246. Treatement Group _____
247. Treatement Group _____
248. Control Group _____
249. Treatement Group _____
250. Control Group _____
251. Treatement Group _____
252. Control Group _____
253. Treatement Group _____
254. Control Group _____
255. Treatement Group _____
256. Treatement Group _____
257. Control Group _____
258. Control Group _____
259. Control Group _____
260. Control Group _____
261. Control Group _____
262. Control Group _____
263. Control Group _____
264. Control Group _____
265. Control Group _____
266. Control Group _____
267. Control Group _____
268. Treatement Group _____
269. Control Group _____
270. Control Group _____
271. Treatement Group _____
272. Treatement Group _____
273. Control Group _____
274. Treatement Group _____
275. Control Group _____
276. Treatement Group _____
277. Control Group _____
278. Control Group _____
279. Control Group _____

Unproctored Internet Testing

280. Control Group _____
281. Treatement Group _____
282. Treatement Group _____
283. Treatement Group _____
284. Treatement Group _____
285. Treatement Group _____
286. Control Group _____
287. Control Group _____
288. Control Group _____
289. Treatement Group _____
290. Treatement Group _____
291. Treatement Group _____
292. Control Group _____
293. Treatement Group _____
294. Treatement Group _____
295. Control Group _____
296. Control Group _____
297. Control Group _____
298. Control Group _____
299. Control Group _____
300. Treatement Group _____
301. Control Group _____
302. Control Group _____
303. Treatement Group _____
304. Treatement Group _____
305. Control Group _____
306. Control Group _____
307. Control Group _____
308. Treatement Group _____
309. Treatement Group _____
310. Control Group _____
311. Treatement Group _____
312. Control Group _____
313. Treatement Group _____
314. Treatement Group _____
315. Treatement Group _____
316. Treatement Group _____
317. Treatement Group _____
318. Treatement Group _____
319. Treatement Group _____
320. Control Group _____
321. Treatement Group _____
322. Control Group _____
323. Treatement Group _____
324. Control Group _____
325. Control Group _____

Unproctored Internet Testing

- 326. Treatement Group_____
- 327. Treatement Group_____
- 328. Control Group_____
- 329. Treatement Group_____
- 330. Control Group_____
- 331. Treatement Group_____
- 332. Treatement Group_____
- 333. Control Group_____
- 334. Treatement Group_____
- 335. Control Group_____
- 336. Control Group_____
- 337. Treatement Group_____
- 338. Control Group_____
- 339. Control Group_____
- 340. Treatement Group_____
- 341. Treatement Group_____
- 342. Control Group_____
- 343. Control Group_____
- 344. Control Group_____
- 345. Control Group_____
- 346. Treatement Group_____
- 347. Treatement Group_____
- 348. Treatement Group_____
- 349. Control Group_____
- 350. Treatement Group_____
- 351. Treatement Group_____

252 subjects randomized into 2 blocks

To reproduce this plan, use the seed 16112

along with the number of subjects per block/number of blocks
and (case-sensitive) treatment labels as entered originally.

Randomization plan created on 5/17/2017, 9:10:46 AM

Appendix J

Debriefing Page / Letter of Explanation

[PLEASE READ CAREFULLY]

Letter of Explanation: Internet-based Employment Testing and Cheating Behaviour

Research suggests that people behave differently when taking job related exams remotely and in unproctored environments. There is also research showing that job candidates with different personality and cognitive characteristics are more likely to cheat during unproctored internet test.

Who Benefits From This Research? Your participation in this research will primarily benefit Emrah Eren in obtaining his Master's degree from Carleton University. Your participation will also support the generation of scientific knowledge that will advance the field of online employment testing. Finally, this research will also benefit [COMPANY NAME] because a summary of the results will be provided to [COMPANY NAME] and will inform the products they develop in the field of online testing.

Use of Mild Deception & Purpose of this Research

A mild form of deception was deployed to prevent you from realizing the true purpose of this research, which was to explore how different exam taking conditions and warning messages are related to cheating behaviour, the extent to which these characteristics influence test taking behaviours, and the intentions to continue engaging in those behaviours in different test taking conditions. It was necessary to withhold the true purpose of this research from you to ensure that you behaved as naturally as you would in an online personnel testing environment.

Participants experienced one of two conditions in this research: in one condition, participants did not receive cheating warning messages. In the other condition, participants received clear warnings and were instructed to make a pledge not to cheat. The purpose of this research is to examine the difference between these two experimental conditions and determine if the warning messages actually work. Cheating (and/or suspicious candidate behaviour) was detected through a series of memory questions administered during your employment testing process. We would kindly ask that you keep the true purpose of this research confidential and do not disclose the details of this recruitment process to your friends, family, and/or colleagues.

If you have any concerns related to this research please do not hesitate to contact Emrah at emraheren@gmail.com or if you have any ethical concerns with the study, please contact Dr. Andy Adler, Chair, Carleton University Research Ethics Board-B (by phone at 613-520-2600 ext. 4085 or via email at ethics@carleton.ca

Consent

Unproctored Internet Testing

Because of the deception used in this research, I have to ask that you provide your consent again after understanding the true purpose of this research. You may choose to withdraw from this research and still participate in the cash prize draw. If you agree to participate in this research please click next, read the new consent form, then click submit. Your time and thoughtful responses are greatly appreciated.

Sincerely,
Emrah Eren, Master of Arts Student
Department of Psychology
Carleton University

Appendix K

Online Consent Form:

Title: Internet-based Employment Testing: Job Candidate Behaviours and Attitudes

Ethics Clearance #: Project # 107337

Date of ethics clearance: October 11, 2017

Ethics Clearance for the Collection of Data Expires: October 11, 2018

Purpose: This is a study on Internet-based employment testing and job candidate cheating behaviour. This study aims to study the impact of different warning messages on job candidate cheating behaviour and attitudes. **The researcher for this study is Emrah Eren in the Psychology Department at Carleton University.**

Who Benefits From This Research? Your participation in this research will primarily benefit Emrah Eren in obtaining his Master's degree from Carleton University. Your participation will also support the generation of scientific knowledge that will advance the field of online employment testing. Finally, this research will also benefit [COMPANY NAME] because a summary of anonymous results will be provided to [COMPANY NAME] and will inform the products they develop in the field of online testing.

Research Personnel: Emrah is working under the supervision of Timothy A. Pychyl in the Psychology Department of Carleton University.

Task requirements: This study involves the completion of psychological measures and questionnaires that will take approximately 15-20 minutes to complete. You can complete the survey at a location of your choice. You will need access to a computer that is connected to the internet. We also ask that you grant us permission to use your [COMPANY NAME] pre-employment testing data from the general abilities tests that you are scheduled to take.

Potential Risk/Discomfort: There are no known risks or discomforts associated with this type of research. Should you experience any discomfort during this research, please contact Emrah at 1-613-614-6853 or emraheren@cmail.carleton.ca.

The results of this research project and/or your decision to participate, will not in any way be related to or impact your job competition process with [COMPANY NAME].

Permission to Use Data: By agreeing to participate in this research, you are granting the researcher of this study permission to use your data for publications, presentations, and teaching purposes. Strong measures will be taken to protect your identity in this research. The personal information you choose to disclose in response to the questionnaires such as your gender, ethnicity, and age will **NOT** be shared with [COMPANY NAME] or any third party. Your identifiable information will be held in complete confidence by the primary researcher and will not inform the hiring process or decision. This will be done by keeping all responses confidential and stored on encrypted drives and allowing you to request that certain responses not be included in the final project. This research is being conducted separately from [COMPANY NAME]'s recruitment process. All data will be presented in aggregate and no identifiable information will be revealed publicly.

Several measures will be taken to keep your survey responses secure, and to ensure that no one from [COMPANY NAME] will have access to your data. All of your responses will be stored on a survey platform that no one from [COMPANY NAME] other than the primary researcher will have access to. No identifiable information will be linked to your survey responses and the file linking your identifiable information to your unique participant code will be destroyed immediately after data collection. All other information pertaining to your results will be stored on an encrypted drive on the researcher's personal computer.

Mild Deception

A mild form of deception was deployed to prevent you from realizing the true purpose of this research, which was to explore how different exam taking conditions and warning messages are related to cheating behaviour, the extent to which these characteristics influence test taking behaviours, and the intentions to continue engaging in those behaviours in different test taking conditions. It was necessary to withhold the true purpose of this research from you to ensure that you behaved as naturally as you would in an online personnel testing environment.

Anonymity, Confidentiality, and Right to withdraw: You have the right to end your participation in the survey or cognitive task at any time, for any reason, up until you hit the "submit" button. You can withdraw by exiting the survey at any time before completing it. If you withdraw from the study, all information you provided will be immediately destroyed. You may withdraw

from the study be emailing the researcher by February 15, 2018.

All research data will be encrypted. The company running the online survey is Qualtrics based in the United States (US). The survey company will keep a copy of the survey responses on its servers in the US and this anonymous data will be subject to US laws on data privacy. This data will also be encrypted and will be deleted once the survey is complete. Research data will be accessible by the researcher, the research supervisor and the survey company. No names or IP addresses will be linked to any of the data provided.

You will have the right to end your participation in the study at any time, for any reason, up until February 15, 2017. If you choose to withdraw, all the information you have provided will be destroyed. Once the project is completed, all research data will be kept for five years and potentially used for other research projects on this same topic. At the end of five years, all research data will be deleted.

Compensation: In appreciation of your time and efforts in completing this questionnaire, you will be provided with the option to enter a draw for a chance to win a cash prize: First prize-\$100; second prize-\$50; and third prize-\$25. Approximately 150 participants will be entered into a draw for 1 of 3 cash prizes. (You will be entered into the draw and the compensation is yours to keep, even if you choose to withdraw.)

If you would like a copy of the finished research project, you are invited to contact the researcher to request an electronic copy which will be provided to you as long as the safety of all participants will not be comprised by doing so.

The ethics protocol for this project was reviewed by the Carleton University Research Ethics Board, which provided clearance to carry out the research.

CUREB-B contact info for Concerns:

If you have any ethical concerns with the study, please contact Dr. Andy Adler, Chair, Carleton University Research Ethics Board-B (by phone at 613-520-2600 ext. 4085 or via email at ethics@carleton.ca).

Researcher contact information:

Name: Emrah Eren

Supervisor contact

Name: Timothy A. Pychyl

Unproctored Internet Testing

Department: Psychology
Carleton University
Tel: [REDACTED]
Email: emraheren@cmail.carleton.ca
TimPychyl@cunet.carleton.ca

Department: Psychology
Carleton University
Tel: 613-520-2600 x 1403
Email:

By clicking "submit", you consent to participate in the research study as described above.