

6-1-2018

Demographic and Psychiatric Correlates of Performance Validity Profiles of Individuals Assessed Subsequent to Motor Vehicle Accidents

Shayna Hannah Nussbaum
University of Windsor

Follow this and additional works at: <https://scholar.uwindsor.ca/etd>

Recommended Citation

Nussbaum, Shayna Hannah, "Demographic and Psychiatric Correlates of Performance Validity Profiles of Individuals Assessed Subsequent to Motor Vehicle Accidents" (2018). *Electronic Theses and Dissertations*. 7477.
<https://scholar.uwindsor.ca/etd/7477>

This online database contains the full-text of PhD dissertations and Masters' theses of University of Windsor students from 1954 forward. These documents are made available for personal study and research purposes only, in accordance with the Canadian Copyright Act and the Creative Commons license—CC BY-NC-ND (Attribution, Non-Commercial, No Derivative Works). Under this license, works must always be attributed to the copyright holder (original author), cannot be used for any commercial purposes, and may not be altered. Any other use would require the permission of the copyright holder. Students may inquire about withdrawing their dissertation and/or thesis from this database. For additional inquiries, please contact the repository administrator via email (scholarship@uwindsor.ca) or by telephone at 519-253-3000ext. 3208.

Demographic and Psychiatric Correlates of Performance Validity Profiles of Individuals
Assessed Subsequent to Motor Vehicle Accidents

By

Shayna Hannah Nussbaum

A Dissertation
Submitted to the Faculty of Graduate Studies
through the Department of Psychology
in Partial Fulfillment of the Requirements for
the Degree of Doctor of Philosophy
at the University of Windsor

Windsor, Ontario, Canada

2018

© 2018 Shayna Hannah Nussbaum

Demographic, Cultural and Psychiatric Correlates of Performance Validity Profiles of
Individuals Assessed Subsequent to Motor Vehicle Accidents
by
Shayna Hannah Nussbaum

APPROVED BY:

K. L. Votruba, External Examiner
University of Michigan Medical School

S. H. Yun
School of Social Work

C. A. Abeare
Department of Psychology

P. A. Timmons Fritz
Department of Psychology

L. A. Erdodi, Advisor
Department of Psychology

March 19, 2018

DECLARATION OF ORIGINALITY

I hereby certify that I am the sole author of this thesis and that no part of this thesis has been published or submitted for publication.

I certify that, to the best of my knowledge, my thesis does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained a written permission from the copyright owner(s) to include such material(s) in my thesis and have included copies of such copyright clearances to my appendix.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee and the Graduate Studies office, and that this thesis has not been submitted for a higher degree to any other University or Institution.

ABSTRACT

The current study explored the effects of cultural, demographic and psychiatric variables on Performance Validity Test (PVT) base rates of failure (BR_{FAIL}) in 325 examinees with traumatic brain injury (TBI) following motor vehicle accidents. PVTs are widely used measures of credibility in neuropsychological assessment. Gaps in the PVT literature regarding the effects of various demographic, cultural, and psychiatric factors limit the generalizability of PVTs.

Higher false-positive rates in minority groups may lead to the inaccurate characterization of members as noncredible, resulting in the denial of treatment and compensation following injuries. To address this gap in the literature, the first objective of the study explored the relationship between BR_{FAIL} , and limited English proficiency, time spent in Canada, education, age, gender, and injury severity. Results indicated that examinees with limited English proficiency had higher BR_{FAIL} on PVTs with low verbal mediation (i.e., tests that did not have verbal components beyond the instructions) compared to Anglophone Canadians. Examinees who had language interpreters had higher BR_{FAIL} on PVTs with both high and low verbal mediation compared to examinees assessed in English. Examinees who immigrated to Canada had higher BR_{FAIL} on both high and low verbal mediation PVTs compared to Canadian-born examinees. Examinees aged 40 to 49 and those with less than high school education had higher BR_{FAIL} for low verbal mediation PVTs than other groups. There were no differences for gender or TBI severity on BR_{FAIL} . These results may be explained by several cultural factors,

including cultural concepts of distress and differences in health literacy, which may contribute to PVT BR_{FAIL}. As such, neuropsychologists should consider the contribution of these cultural factors when interpreting PVT results of examinees who have immigrated to Canada.

Another important gap in the literature is in regards the relationship between PVTs and dissociative symptoms (i.e., disrupted consciousness, affect, and memory). Findings on the effects of psychiatric factors (e.g., posttraumatic, depressive, and anxious symptoms) on PVT BR_{FAIL} are mixed but generally indicate that PVTs are robust to psychiatric disorders except psychosis. However, disruptions in consciousness, memory, and affect due to dissociative pathology might be expected to interfere with test performance. The second objective of this study explored the relationship between BR_{FAIL} and dissociative, posttraumatic, anxious, and depressive symptoms. Results indicated elevated rates of PVT BR_{FAIL} for examinees with higher levels of self-reported posttraumatic, depressive, and anxious symptoms. Results also indicated that those with high self-reported dissociative symptoms had higher BR_{FAIL} for verbally mediated PVTs. The findings suggest that dissociative symptoms may interfere with verbally mediated PVTs, and highlight the need for further research into the effects of dissociative pathology on neuropsychological and PVT performance.

The current study demonstrated that previously unexplored cultural, demographic, and psychiatric factors are related to PVT performance, and may affect the interpretation of PVTs. Implications, limitations, and avenues for future research are discussed.

DEDICATION

I dedicate this dissertation to my community of graduate students,
including the spouses and partners of graduate students who remain supportive
throughout the journey.

ACKNOWLEDGEMENTS

I would like to acknowledge all of the support and guidance of my family, friends, and mentors. To my parents, thank you for your dedication to education, and for your boundless love and support. To Daniel, thank you for your belief in me and for your patience and love through my triumphs and struggles. To Katherine, graduate school would not have been the same without you. To Siqi, you somehow help me to stay humble and believe in myself at the same time. To Ying, your coding and plumbing expertise have saved me countless hours, even if you will not admit it. To David, I have no idea where I would be in life without you as a role model. To Mark, our work together was the impetus for this dissertation and beyond.

To my committee members, thank you for your help in honing this project.

To Laszlo, thank you for all of your guidance and support throughout this process.

You are an insightful and entertaining mentor.

TABLE OF CONTENTS

DECLARATION OF ORIGINALITY	iii
ABSTRACT	iv
DEDICATION	vi
ACKNOWLEDGEMENTS	vii
LIST OF TABLES	xiv
LIST OF ABBREVIATIONS	xviii
CHAPTER 1 Background.....	1
Introduction	1
Traumatic Brain Injury	2
Clinical classification: Injury severity	3
mTBI epidemiology.....	7
mTBI pathophysiology.....	8
mTBI cognitive effects.....	11
mTBI symptoms.....	12
mTBI recovery.....	13
Litigation and Health Outcomes	22
Litigation and physical health outcomes.....	22
Litigation and psychological outcomes.....	25
Litigation and neuropsychological assessment.....	26
Performance Validity	28

Definition.....	28
Research design.....	28
PVT use guidelines.....	29
PVT controversy.....	30
mTBI and performance validity.....	33
Litigation and performance validity.....	33
PVT types.....	36
CHAPTER 2 Culture, Linguistics, Demographics, and Performance Validity.....	38
Culture and Neuropsychological Assessment.....	38
Guidelines for Multiculturally Competent Assessment.....	39
Current Practice of Multicultural Neuropsychology.....	41
Cultural Factors and Performance Validity.....	43
PVTs administered in Spanish.....	45
PVTs in Asia and with Asian Americans.....	48
Other Demographic Variables and PVTs.....	54
Ontario Ethnic Demographics.....	58
Conceptualization of the Effect of Cultural, Linguistic, and Demographic Factors on PVTs in the Current Study.....	59
CHAPTER 3 Psychiatric Symptoms and Performance Validity.....	62
Depression, Anxiety, and PTSD and Performance Validity.....	62

Psychiatric disorders and mTBI.	62
Psychiatric disorders and neuropsychological functioning.	64
Psychiatric disorders and PVT performance.	66
Dissociative Symptoms and Disorders.....	70
Dissociation and symptom validity.	71
Dissociation and neuropsychological functioning.....	72
Neuroanatomical correlates of dissociation.....	73
Conceptualization of Psychiatric Symptoms in the Current Study.....	74
CHAPTER 4 General Methods.....	77
Participants.....	77
Descriptive statistics.....	77
Power Analysis.....	79
Measures.....	80
Performance validity tests.	80
Multivariate indicators of performance validity.....	89
Missing Data.....	98
CHAPTER 5 Methodological Adjustments and Descriptive Results.....	100
Methodological Adjustments.....	100
Removing older adults from analyses.	100
Cutoff adjustment.	105

General Results	121
Statistical assumptions.....	121
CHAPTER 6 Objective 1: Cultural, Linguistic, and Demographic Variables and PVTs Methods	124
Language Proficiency.....	124
Time in Canada	125
Education.....	126
TBI Severity	127
Gender	127
Age	127
Hypotheses	128
Research Questions	128
Participants	129
Measures.....	129
Demographic information.....	129
Performance validity tests.	129
CHAPTER 7 Objective 1 Results	130
Hypothesis 1a	130
Exploratory analyses.....	131
Hypothesis 1b.....	137
Hypothesis 1c	142

Hypothesis 2	147
Exploratory analyses.....	147
Question 1	152
Question 2	155
CHAPTER 8 Objective 1 Discussion	158
Limited English Proficiency and BR _{FAIL}	158
Cultural concepts of distress.....	161
Impact of interpreters.....	164
Time in Canada and BR _{FAIL}	166
Health literacy.....	167
Alternative explanations.....	169
Education and BR _{FAIL}	171
TBI severity and BR _{FAIL}	171
Gender and BR _{FAIL}	173
Age and BR _{FAIL}	173
CHAPTER 9 Objective 2: Psychiatric Symptoms and Performance Validity	
Method	174
Research Questions	174
Participants	174
Measures.....	174
Performance validity tests.	174

Symptom validity tests.	174
Clinical scales.	177
CHAPTER 10 Results Objective 2	180
Initial Analyses.....	180
Question 1	186
Question 2	189
Question 3	193
Question 4	196
CHAPTER 11 Discussion Objective 2	199
Self-reported Depression, Anxiety, and PTSD Symptoms and BR _{FAIL}	199
Dissociation and BR _{FAIL}	203
CHAPTER 12 General Discussion	205
Strengths of the Present Research	210
Limitations of the Present Research.....	212
Future Directions.....	213
Conclusion.....	214
References.....	216
Vita Auctoris.....	266

LIST OF TABLES

Table 1 <i>Glasgow Coma Scale</i>	4
Table 2 <i>Means, Standard Deviation, and Range of Demographic Characteristics</i>	79
Table 3 <i>A Priori Levels of Failure for EI-7_{VER} Components</i>	94
Table 4 <i>A Priori Levels of Failure for EI-7_{VIS} Components</i>	96
Table 5 <i>Independent Samples t Tests on Examinees ≥ 70 Years and ≤ 69 Years on TOMM, EI-7_{VER}, EI-7_{VIS}, VI-10, and their Components</i>	103
Table 6 <i>TOMM Percent Administered, Means, Standard Deviations, Cutoff Scores and Associated BR_{FAIL}</i>	107
Table 7 <i>Frequency, Percentage, Cumulative Percentage and Classification Ranges for A Priori EI-7_{VER} and EI-7_{VIS}</i>	110
Table 8 <i>Adjusted Levels of Failure for EI-7_{VER} Components</i>	115
Table 9 <i>Adjusted Levels of Failure for EI-7_{VIS} Components</i>	117
Table 10 <i>Frequency, Percentage, Cumulative Percentage and Classification Ranges for Adjusted EI-7_{VER}, EI-7_{VIS}, and VI-10</i>	119
Table 11 <i>Sample Size, Mean, Standard Deviation, Range, Median, Skewness, Kurtosis, and Shapiro-Wilk's Test for TOMM Trials, EI-7_{VER}, EI-7_{VIS}, and VI-10</i>	123
Table 12 <i>Independent t Tests on Age, TOMM Trial 1, EI-7_{VER}, EI-7_{VIS}, and VI-10 Scores in Examinees Included or Excluded from Hypothesis 1 Analyses</i>	133
Table 13 <i>Independent t Tests on TOMM Trial 1, EI-7_{VER}, EI-7_{VIS}, and VI-10 Scores for Examinees by English Proficiency</i>	134

Table 14 <i>TOMM, EI-7_{VER}, EI-7_{VIS}, and VI-10 BR_{FAIL} for Examinees by English Proficiency</i>	135
Table 15 <i>Independent t Tests on TOMM Trial 1, EI-7_{VIS}, and VI-10 Scores in Examinees Who Did or Did Not Utilize an Interpreter</i>	136
Table 16 <i>Independent t Tests on Age, TOMM Trial 1, EI-7_{VER}, EI-7_{VIS}, and VI-10 Scores in Examinees Included or Excluded from Hypothesis 2 Analyses</i>	139
Table 17 <i>Independent t Tests on TOMM Trial 1, EI-7_{VER}, EI-7_{VIS}, and VI-10 Scores for Examinees by Time in Canada</i>	140
Table 18 <i>TOMM, EI-7_{VER}, EI-7_{VIS}, and VI-10 BR_{FAIL} for Examinees by Time in Canada</i>	141
Table 19 <i>One-way ANOVAs on TOMM Trial 1, EI-7_{VER}, EI-7_{VIS}, and VI-10 Scores by Education Level</i>	144
Table 20 <i>TOMM, EI-7_{VER}, EI-7_{VIS}, and VI-10 BR_{FAIL} by Education</i>	145
Table 21 <i>TOMM, EI-7_{VER}, EI-7_{VIS}, and VI-10 BR_{FAIL} for Examinees with Lowest and Highest Education</i>	146
Table 22 <i>Independent t Tests on TOMM Trial 1, EI-7_{VER}, EI-7_{VIS}, and VI-10 Scores in Examinees with mTBI (Uncomplicated and Complicated) and Moderate-to-Severe TBI</i>	148
Table 23 <i>TOMM, EI-7_{VER}, EI-7_{VIS}, and VI-10 BR_{FAIL} by TBI Severity: mTBI (Uncomplicated and Complicated) and Moderate-to-Severe TBI</i>	149
Table 24 <i>Independent t Tests on TOMM Trial 1, EI-7_{VER}, EI-7_{VIS}, and VI-10 Scores in Examinees with Uncomplicated mTBI Compared to Complicated Mild, Moderate, and Severe TBI</i>	150

Table 25 <i>TOMM, EI-7_{VER}, EI-7_{VIS}, and VI-10 BR_{FAIL} by TBI Severity: Uncomplicated mTBI Compared to Complicated Mild, Moderate, and Severe TBI</i>	151
Table 26 <i>Independent t Tests on TOMM 1, EI-7_{VER}, EI-7_{VIS}, and VI-10 Scores by Gender</i>	153
Table 27 <i>TOMM, EI-7_{VER}, EI-7_{VIS}, and VI-10 BR_{FAIL} by Gender</i>	154
Table 28 <i>One-way ANOVAs on TOMM Trial 1, EI-7_{VER}, EI-7_{VIS}, and VI-10 Scores by Age</i>	156
Table 29 <i>TOMM, EI-7_{VER}, EI-7_{VIS}, and VI-10 BR_{FAIL} by Age</i>	157
Table 30 <i>Sample Size, Mean, Standard Deviation, Range, Median, Skewness, Kurtosis, and Shapiro-Wilk's Test for BDI-II and BAI</i>	183
Table 31 <i>One-way ANOVAs on BDI-II and BAI Scores for Examinees with Valid, Invalid, and Missing TSI-II-A Data</i>	184
Table 32 <i>PAI and TSI-II-A Symptom Validity Test Failures</i>	185
Table 33 <i>One-way ANOVAs on TOMM Trial 1, EI-7_{VER}, EI-7_{VIS}, and VI-10 Scores in Examinees with Low, Moderate, and Severe Self-Reported Depression (BDI-II)</i>	187
Table 34 <i>TOMM, EI-7_{VER}, EI-7_{VIS}, and VI-10 BR_{FAIL} for Examinees with Low, Moderate, and Severe Self-Reported Depression (BDI-II)</i>	188
Table 35 <i>One-way ANOVAs on TOMM Trail 1, EI-7_{VER}, EI-7_{VIS}, and VI-10 Scores in Examinees with Low, Moderate, and Severe Self-Reported Anxiety (BAI)</i>	191
Table 36 <i>TOMM, EI-7_{VER}, EI-7_{VIS}, and VI-10 BR_{FAIL} for Examinees with Low, Moderate, and Severe Self-Reported Anxiety (BAI)</i>	192

Table 37 <i>One-way ANOVAs on TOMM Trial 1, EI-7_{VER}, EI-7_{VIS}, and VI-10 Scores in Examinees with Low, Mild, and Moderate-to-Severe Self-Reported Posttraumatic Symptoms (PAI Anxiety Related Disorders Traumatic Stress Subscale)</i>	194
Table 38 <i>TOMM, EI-7_{VER}, EI-7_{VIS}, and VI-10 BR_{FAIL} for Examinees with Low, Mild, and Moderate-to-Severe Self-Reported Posttraumatic Symptoms (PAI Anxiety Related Disorders Traumatic Stress Subscale)</i>	195
Table 39 <i>Independent t Tests on TOMM Trial 1, EI-7_{VER}, EI-7_{VIS}, and VI-10 Scores in Examinees with Normal and Elevated Self-reported Dissociative Symptoms (TSI-II-A Dissociation)</i>	197
Table 40 <i>TOMM, EI-7_{VER}, EI-7_{VIS}, and VI-10 BR_{FAIL} for Examinees with Normal and Elevated Self-Reported Dissociative Symptoms (TSI-II-A Dissociation)</i>	198

LIST OF ABBREVIATIONS

BAI	Beck Anxiety Inventory
BDI-II	Beck Depression Inventory Second Edition
BR _{FAIL}	Base Rate of Failure
CI	Confidence Interval
CT	Computed Tomography
EI-14	Erdodi Index 14
EI-7 _{VER}	Erdodi Index 7 Verbal
EI-7 _{VIS}	Erdodi Index 7 Visual
FAS	F, A, and S trials of the Controlled Oral Word Association Test
MRI	Magnetic Resonance Imaging
mTBI	Mild Traumatic Brain Injury
PAI	Personality Assessment Inventory
PVT	Performance Validity Test
PTSD	Posttraumatic Stress Disorder
TBI	Traumatic Brain Injury
TOMM	Test of Memory Malingering
TSI-II-A	Trauma Symptom Inventory Second Edition Alternate Version
VI-10	Validity Index Ten
WAIS	Wechsler Adult Intelligence Scale
WASI	Wechsler Abbreviated Scale of Intelligence
WMS	Wechsler Memory Scale

CHAPTER 1

Background

Introduction

Performance validity tests (PVTs) are instruments used to determine the credibility of cognitive data and are considered essential to neuropsychological assessment, as they help identify a common confound in psychometric testing.

Noncredible performance is particularly common in the presence of external incentive to appear impaired, such as claiming medical and financial benefits after sustaining a mild traumatic brain injury (mTBI) in a motor vehicle accident caused by a third party. There is a complex relationship among injury parameters, external incentives, premorbid functioning, postinjury mental health, and performance on neuropsychological testing.

Although demographic, cultural, and linguistic factors can further complicate the clinical interpretation of test results, their effect on PVTs has received little attention in the scientific literature. There is a similar knowledge gap regarding the relationship between PVTs and certain psychiatric symptoms.

It is important to understand how cultural, demographic, linguistic, and psychiatric factors affect PVT failure. Determining noncredible neuropsychological performance can negatively affect the examinee, through health care and financial benefit denial and potentially resultant worsening symptoms and functional disability. These determinations may also be discriminatory if they systematically target particular groups. Practicing ethically and effectively requires knowledge of how to interpret PVTs accurately with people from nondominant cultures and with people who have prominent psychiatric symptoms.

The present study included examinees who had been assessed to determine insurance benefits following TBIs sustained in motor vehicle accidents and had two major objectives. The first addressed the contribution of cultural, linguistic, demographic, and injury factors to PVT results. The second addressed the contribution of self-reported psychiatric symptoms to PVT results.

Traumatic Brain Injury

Appreciating the complex interplay of factors contributing to PVT performance after motor vehicle accidents begins by understanding the index injury itself, namely traumatic brain injury (TBI). TBI is defined as a change in brain function or brain pathology resulting from an external force (Menon, Schwab, Wright, & Maas, 2010). Changes in brain function may include loss of consciousness, posttraumatic amnesia, changes in mental state (e.g., confusion), and neurological deficits (e.g., double vision). Evidence of brain pathology can include signs of edema, hemorrhage, or other abnormalities on scans. External forces that cause TBI include the head striking an object (e.g., steering wheel); acceleration and deceleration forces on the brain (e.g., upon an abrupt stop in a motor vehicle accident); an object entering the brain (e.g., a bullet); and explosive forces (e.g., from a bomb). TBI severity should predict symptom severity and recovery, but a web of complex and interrelated factors affects the relationship.

Despite the lack of a clear dose-response relationship in recovery from TBI, there are important differences in the pathophysiology and recovery course depending on injury severity. Mild TBI (mTBI) is by far the most common severity both in the general population (70% to 90%; Holm, Cassidy, Carroll, & Borg, 2005) and in the current sample (89.3%). It is thus particularly important to understand the pathophysiology,

typical course, and factors that interfere with recovery from mTBI. The following sections explain TBI and the factors that affect recovery.

Clinical classification: Injury severity. TBI is classified by initial level of consciousness, loss of consciousness duration, or posttraumatic amnesia duration. The Glasgow Coma Scale is the most widely used measure of consciousness level across emergency medical services and departments (American Congress of Rehabilitation Medicine Mild Traumatic Brain Injury Committee, 1993). Glasgow Coma Scale scores range from 3 to 15, where 3 is indicative of coma and 15 is oriented to verbal commands and responsive to space and time as shown in Table 1.

Table 1

Glasgow Coma Scale

Eye opening (E)		Verbal response (V)		Best motor response (M)	
1	None	1	None	1	None
2	To pressure	2	Sounds	2	Extension
3	To speech	3	Words	3	Abnormal flexion
4	Spontaneous	4	Confused	4	Normal flexion (withdrawal)
		5	Oriented	5	Localising
				6	Obeying commands

Note. This table is adapted from Teasdale et al. (2014)'s review of the Glasgow Coma Scale. Reproduced with permission.

Uncomplicated mild TBI. mTBI is defined as a head injury with loss of consciousness \leq 30 minutes; posttraumatic amnesia \leq 24 hours; and initial Glasgow Coma Scale score between 13 and 15 (American Congress of Rehabilitation Medicine Mild Traumatic Brain Injury Committee, 1993). mTBI can be diagnosed regardless of the presence of positive neuroimaging findings (e.g., the presence of edema or hemorrhage). mTBI can be further classified as uncomplicated when neuroimaging is negative or complicated when neuroimaging findings are positive (Bigler, 2013).

Complicated mild TBI. Complicated mTBI refers to injuries that meet criteria for mTBI with abnormalities on neuroimaging, such as bleeding, swelling, or skull fracture (e.g., Iverson, 2005; Ruff, Iverson, Barth, Bush, & Broshek, 2009; Williams, Levin, & Eisenberg, 1990). Of patients who present to the emergency room with mTBI, 7–20% have positive computed tomography (CT) results, indicating complicated mTBI (Iverson, 2005). Notably, those who present to the emergency room with mTBI are not representative of all individuals with mTBI, many of whom do not present for any medical evaluation or treatment. People with complicated mTBI would be expected to present for medical evaluation at higher rates compared to people with less severe mTBI. As such, the proportion of complicated mTBI cases presenting to emergency rooms is likely higher than what would be found for the mTBI population as a whole (Iverson, 2005).

Some authors argue that people with complicated mTBI and people who have moderate TBI have similar cognitive (e.g., memory, processing speed) and functional (e.g., independence in daily living) outcomes immediately following injury (Kashluba, Hanks, Casey, & Millis, 2008). However, other researchers have found similar vocational

outcomes and 6-month follow-up neuropsychological test results in those with uncomplicated and complicated mTBI (Hanlon, Demery, Martinovich, & Kelly, 1999; Hughes et al., 2004). Mild complicated TBI is thus a meaningful intermediate severity category between uncomplicated mTBI and moderate TBI, as it shares features with both classifications.

Moderate and severe TBI. Moderate and severe TBI encompass wide severity ranges. Glasgow Coma Scale scores of 9 to 12 or 1- to 24-hour posttraumatic amnesia indicate moderate TBI (Lezak, Howieson, Bigler, & Tranel, 2012; Teasdale et al., 2014). Glasgow Coma Scale scores of 8 or less, or one to seven days of posttraumatic amnesia indicate severe TBI (Lezak et al., 2012; Teasdale et al., 2014). More than four weeks posttraumatic amnesia duration indicates very severe TBI (Lezak et al., 2012). Glasgow Coma Scale scores between three and five indicate extremely severe TBI and poor prognosis, and 90% of patients with Glasgow Coma Scale scores of three die within a day of the injury (Kaufman & Milstein, 2007; Langlois Orman, Kraus, Zaloshnja, & Miller, 2011). Moderate and severe TBI are rarer than mTBI, and their pathophysiological effects are more enduring than the effects of mTBI (Green et al., 2014).

Classification limitations. Even with the addition of complicated mTBI to the classification system, symptom and injury parameters vary widely within categories (Teasdale et al., 2014). For example, uncomplicated mTBI encompasses injuries of very mild severity, such as having a Glasgow Coma Scale of 15 (confusion and disorientation) that resolves within two minutes, without any other symptoms. Uncomplicated mTBI also includes injuries with 20 minutes of loss of consciousness, double vision, vomiting,

and 40 minutes of posttraumatic amnesia when there is no evidence of pathology on neuroimaging.

Categorizing injury severity can be helpful for research (Carone, 2008) and treatment (Green et al., 2014). However, TBI severity is inherently continuous. Grouping inherently continuous data introduces error from the wide variance within each group and small differences between participants who are close to classification cutoffs. In other words, people who are just above or below a classification cutoff have more in common with each other than with others within their classification. These limitations to classification are present in any system that attempts to classify inherently continuous data.

mTBI epidemiology. The following sections will focus on mTBI, as it is both the most common and the most prognostically complex TBI category. A review of the literature that included 313 articles showed that 100 to 300 per 100,000 people are treated for mTBI annually in hospitals and that 70% to 90% of TBI cases seen are mild (Holm et al., 2005). Self-report mTBI incidence rates are higher, at around 600 per 100,000 in the general population (Holm et al., 2005). Motor vehicle accidents account for the majority of TBIs worldwide (World Health Organisation, 2004).

Motor vehicle accidents also accounted for 11.9% of emergency room visits and hospitalizations in Ontario for TBI between 2002 and 2007 (Colantonio et al., 2010). A study reviewing TBI cases in Ontario between 2004 and 2007 ($N = 11,970$) showed 2,515 new cases of TBI from motor vehicle accidents (Chen et al., 2012). In sum, mTBIs are common and represent a significant portion of acute care visits in Ontario and abroad.

mTBI pathophysiology. Complex metabolic mechanisms cause acute symptoms of mTBI, whereas permanent structural damage is minor and often only detectable with high-resolution techniques. A minority of mTBI patients show positive findings on CT, but some others with normal conventional CT have abnormalities on magnetic resonance imaging (MRI) or single-photon emission CT (Iverson, 2005). The results of recent research using diffusion tensor imaging (a high-resolution brain imaging technique), however, showed regional brain volume differences in patients following mTBI compared to healthy control participants (e.g., Zagorchev et al., 2016), indicating that there are some lasting pathophysiological changes following mTBI. Diffusion tensor imaging will be explored in more detail in a later section.

Mild TBI structural damage is both quantitatively and qualitatively less severe than moderate or severe TBI damage. Although axonal shearing is common in more severe TBI, even mild injury can sometimes cause a small number of damaged axons gradually swelling and separating (Gaetz, 2004; Iverson, 2005; Smith, 2011). Axonal separation does not necessarily cause cell death, and axonal shearing and cell death are qualitatively different from the effects of uncomplicated mTBI (Larrabee, Binder, Rohling, & Ploetz, 2013; McAllister, 2011). In a review of the extant literature, Iverson (2005) stressed that cell death is commensurate with injury severity, and that very few cells are likely to die from mTBI. Many of the effects of mTBI are transitory, whereas others may endure, as explained below.

MacFarlane and Glenn (2015) reviewed the literature on metabolic cascade following mTBI. In their description, mTBI triggers ionic fluxes causing an uncontrolled release of excitatory neurotransmitters. Potassium leaves neurons, triggering excitatory

neurotransmitter release, which causes calcium influx and absorption into mitochondria. Glucose metabolism rises as cells use glucose as fuel to initiate ion pumps to restore normal membrane potential. Concurrent mild reduction in cerebral blood flow due to compromised cerebrovascular automatic regulation exacerbates the differences in glucose availability and demand. Oxidative metabolism may be compromised, and mitochondrial function may decline. Anaerobic energy pathways are used instead, elevating lactate levels. Magnesium, which is essential for energy production, decreases significantly for several days following injury. Increased calcium levels damage axons, leading them to swell and separate. Necrosis and apoptosis can occur, although transient metabolic changes are more common than cell death in mTBI. MacFarlane and Glenn (2015) concluded that recovery from the neurochemical cascades is commensurate with injury severity and closely matches typical neurobehavioural recovery following TBI.

Bigler (2008) reviewed the literature on the mechanics of mTBI and concluded that rapid deceleration of the brain—even without collision of the head with an external object—strains the upper brainstem, pituitary-hypothalamic axis, medial temporal lobe, and basal forebrain. He also found that mTBI irritated the vasculature and meninges, and was associated with white matter degeneration in the fornix, anterior commissure, and most prominently in the corpus callosum. Some studies he reviewed showed subtle brain volume loss associated with white matter pathology. There was also evidence of increased frequency of dilated perivascular spaces in mTBI, and changes in white matter volume and composition, which may relate to persistent symptoms. The author additionally speculated that blood vessel stretching in mTBI might impair neurogenesis.

Bigler hypothesized that white matter changes and volume loss cause postconcussive symptoms. He reviewed several studies in which boxers showed pathophysiological changes on lumbar puncture and diffusion tensor imaging, even in the absence of cognitive complaints. The author also pointed to several postmortem case studies of individuals with postconcussion syndrome whose petechial hemorrhagic lesions or edema were revealed only upon autopsy. Some studies he reviewed found that those with postconcussion syndrome had normal blood flow at rest on positron emission tomography, but abnormal cerebral blood flow during cognitively demanding situations. Bigler also reviewed studies showing increased likelihood of incurring a second mTBI after a first mTBI, with greater pathology following the second TBI than would be expected following only one mTBI. He argued that evidence of greater pathology following a second mTBI demonstrates that the first injury is not benign—in other words, individuals may be adapting to, rather than healing from, injuries. Overall, Bigler posited that mTBI results in significant pathophysiological changes that may persist beyond the expected three-month recovery time for the injury.

Overall, mechanical injury and metabolic changes lead to symptoms experienced in the weeks following mTBI (Larrabee & Rohling, 2013). Neurological explanations for ongoing symptoms following mTBI, however, have been largely unverifiable until recently, with the advent of more sensitive MRI techniques such as diffusion tensor imaging.

Diffusion tensor imaging and mTBI. Diffusion tensor imaging is a recent imaging technique that primarily targets white matter, tracking the mobility of molecules in anisotropic (i.e., directionally dependent) tissue (Le Bihan et al., 2001). A detailed

discussion of diffusion tensor imaging is beyond the scope of the current work, but diffusion tensor imaging is a promising technology that can detect subtle changes in white matter integrity conventional CT and MRI might miss (Bigler, 2008).

Studerus-Germann et al. (2016) reviewed the literature on neuroimaging in mTBI and found that although mTBI was associated with white matter changes, evidence of a relationship between abnormal diffusion tensor imaging findings and poorer neuropsychological performance was mixed. Diffusion tensor imaging abnormalities are not always present in mTBI, do not consistently predict the presence of postconcussion syndrome (i.e., lingering effects of mTBI three or more months postinjury), and are present in some healthy individuals who have not sustained head injuries (Waljas et al., 2015). Psychological and social factors that contribute to postconcussion syndrome (e.g., expectation, emotion, and incentive) may obscure the relationship between postconcussion syndrome and structural damage. Mixed findings of the relationship between neuropsychological performance, symptom report, and evidence of pathology in highly sensitive neuroimaging underscores the complex development and maintenance of symptoms following mTBI.

mTBI cognitive effects. In a systematic review of the literature, Carroll et al. (2014) reported that cognitive deficits were common in mTBI patients in the first two weeks post-injury. Type and magnitude of deficits across studies were inconsistent. There was limited evidence that loss of consciousness predicts slower processing speed, and that positive CT scans (i.e., complicated mTBI) are associated with poorer cognitive functioning. There were, however, more similarities than differences between the cognitive effects of complicated and uncomplicated mTBI (e.g., attention, working

memory, executive functioning, memory, psychomotor speed). The authors concluded that there is rapid recovery over the first month post-injury but that deficits may linger. The authors found limited evidence that some deficits may persist for three to six months and limited evidence that these lingering deficits remit by one to five years post-injury. The authors point to a need for better control of confounds and well-conducted, confirmatory, longitudinal studies to gain a better understanding of the effects and course of recovery from mTBI.

mTBI symptoms. Common mTBI symptoms include headache, dizziness, and fatigue, with full recovery generally within 3 to 12 months (Holm et al., 2005). Evidence consistently shows that, in most cases, cognitive deficits attributable to mTBI completely resolve within one to three months post-injury (Karr, Areshenkoff, & Garcia-Barrera, 2014). Individuals with symptoms that linger beyond three months are diagnosed with postconcussion syndrome (Bigler, 2008; Pertab, James, & Bigler, 2009), dubbed the “miserable minority” of people who experience symptoms for months or years post-mTBI (Rohling, Larrabee, & Millis, 2012; Ruff, Camenzuli, & Mueller, 1996). The International Statistical Classification of Diseases and Health Problems – Tenth Edition (World Health Organization, 2010) defines postconcussional syndrome as a post head injury syndrome characterized by headache; fatigue; dizziness; concentration and memory issues; irritability; insomnia; and decreased stress, emotional, and alcohol tolerance. The *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.; *DSM-5*; American Psychiatric Association, 2013) cautions that although clinicians may consider diagnosing major or mild neurocognitive impairment due to traumatic brain injury, neurocognitive symptoms of mTBI resolve within days or weeks postinjury, and

clinicians should consider additional diagnoses if there is significant deterioration beyond this timeframe. In other words, clinicians and researchers should be aware that factors other than the traumatic brain injury itself contribute to deficits and symptoms in the chronic period following mTBI.

mTBI recovery. Researchers and clinicians actively debate the etiology of symptoms that individuals experience in the chronic period (Pertab et al., 2009; Rohling et al., 2012). Well-designed prospective studies examining the resolution of symptoms following mTBI are rare; however, it appears that lingering symptoms within nonsports concussion populations may be partially or wholly attributable to factors other than the injury. These factors include previous TBI, comorbid psychiatric difficulties, and having incurred the injury in a motor vehicle accident (Karr et al., 2014; Ponsford et al., 2000).

Sampling bias further obfuscates scientific understanding of the effects of mTBI. The majority of people who have incurred an mTBI do not present to emergency rooms, and may never be diagnosed or treated for the injury (McAllister, 2011). Most participants are recruited through community health care, and as such, studies exclude people with mTBI who do not present for assessment or treatment (McAllister, 2011). These self-selected samples likely represent a more severe subset of mTBI, or a subset of individuals who incur mTBI and differ in other ways (e.g., anxiety, help-seeking, compensation seeking) from those who do not seek diagnosis or treatment. Although researchers have long recognized sampling bias in mTBI studies (Ruff, Camenzuli et al., 1996), the field has only begun to attempt to overcome sampling issues via prospective studies that recruit patients in hospital emergency rooms (e.g., Isokuorti et al., 2016; Waljas et al., 2015). The prospective design circumvents bias of post-acute mTBI

recruitment, when most individuals would have recovered from mTBI, leaving only people who are a part of the “miserable minority” as potential participants. These prospective studies, however, cannot include individuals who sustain TBI but never present to treatment, or those who seek care through family physicians.

Prospective mTBI recovery findings. Recently, a group of researchers in Finland have published several prospective studies of mTBI to circumvent some of the previously explained confounds. In a large inception cohort study, Isokuortti et al. (2016) attempted to screen for emergency room patients who presented with “pure” mTBI (i.e., patients who met criteria for mTBI but did not have other injuries, illnesses, diseases, or psychiatric disorders). Only 2.5% of the 3,023 participants met these criteria. The authors emphasized that it is difficult to disentangle effects of mTBI from pre-existing conditions both clinically and in research due to the high rates of pre-existing conditions that are known to be risk factors for poor mTBI outcome or have similar symptoms or signs to mTBI. The most common pre-existing conditions were cardiovascular, neurological (including prior TBI), and psychiatric (mostly alcohol abuse and affective disorders). The authors concluded that research on subgroups of individuals with mTBI who have various pre-existing characteristics would help develop a better understanding of effective conceptualization and treatment of people with various pre-existing conditions who sustain mTBI.

In a sample of 126 participants, Waljas et al. (2015) found that depressive symptoms, pre-injury mental health difficulties, and nonhead injuries predicted postconcussive symptoms one month postinjury and found that depressive symptoms were related to postconcussive symptoms at one year. Interestingly, the authors noted

postconcussive symptom (e.g., headache, fatigue, irritability) endorsement rate did not differ between participants with mTBI (38%) and matched controls (31%) at one year postinjury. In other words, these symptoms do not appear to be specific to individuals who had incurred an mTBI, calling the etiology of “postconcussive” symptoms into question. The authors cautioned that postconcussion syndrome diagnostic criteria have a high false-positive rate, and further concluded that postconcussion syndrome is likely the result of cumulative effects of multiple variables and that the contribution of structural damage to postconcussion syndrome remained unclear, as evidence of greater structural abnormality on imaging was unrelated to greater symptom reporting.

Another prospective study by the same research group (Losoi et al., 2016) compared 74 cases with mTBI recruited consecutively from an emergency department to a control group of 40 participants who sustained only ankle injuries. The orthopaedic control and mTBI groups did not differ on postconcussive symptoms at 12-month follow-up. Self-reported life satisfaction, fatigue, insomnia, depressive symptoms, and pain did not differ significantly between the two groups at one, six, or 12-month follow-up. Quality of life was only lower for the mTBI group at six months following injury. Almost all (96%) of the mTBI group returned to school or work within the 12-month follow-up period, and 16 days was the median time to return to work. It is apparent from this recent prospective study that the vast majority of individuals with mTBI who present to emergency department make a fast and complete recovery.

Sports-related mTBI recovery. In addition to growing evidence that a full recovery is typically the normative outcome after mTBI, similar results have been found in studies of sports-related mTBI specifically. In a systematic review of the literature,

Iverson (2005) found that athletes typically fully recover symptomatically and cognitively within two to 14 days of mTBI. Patients who incurred the mTBI in nonsports related accidents were slower to recover. Ongoing symptoms in the nonsports group were related to substance use, poor overall health, nonhead injuries, pain, depression, life stress, unemployment, and litigation (Iverson, 2005). There are also monetary incentives to remain symptomatic after motor vehicle or other (e.g., workplace) accident mTBI, whereas athletes with mTBI have opposing incentives to appear asymptomatic so that they can return to play (Spenceley, 2013). These opposing incentives further complicate the understanding of recovery following mTBI.

mTBI complaint specificity. Cognitive complaints (e.g., subjective memory difficulties) following head and neck injuries are also not specific to mTBI and are reported in various other groups, including major depressive disorder, chronic pain, posttraumatic stress disorder (PTSD), and those who fail PVTs (Holm et al., 2005). Many factors other than injury severity are associated with poorer health following possible mTBI in motor vehicle accidents. For instance, in a study of motor vehicle accident insurance claimants from Saskatchewan, older age, not seeking healthcare quickly after the motor vehicle accident, poorer self-rated health one month prior to motor vehicle accident, past motor vehicle accident claim, depression, dizziness, sleep problems, restriction of daily home activities, greater neck/shoulder and low back pain, lower expectations for recovery, and initially only seeking healthcare from a medical doctor were all predictors of worse outcome (Zhang, Carroll, Cassidy, & Paniak, 2009).

Overall, research suggests that symptoms that occur immediately following the mTBI, and which last for a brief period, are attributable to pathophysiological changes

caused by the mTBI. The etiology of lingering symptoms is more complex, involving expectation, pain, emotional difficulties, and external incentives.

Postconcussion syndrome. Symptoms persisting beyond three months following an uncomplicated mTBI indicate postconcussion syndrome (Bigler, 2008), which occurs in about 10% of mTBI cases (Wood, 2004). Symptoms include fatigability, sleep difficulties, headache, dizziness or vertigo, irritability, mood dysregulation, personality changes, and apathy (Bigler, 2008). Several models, described below, have been posited to explain the persistence of postconcussion syndrome symptoms.

Expectation as etiology. Mittenberg et al. (1992) posited “expectation as etiology” as a model to explain postconcussion syndrome. In this model, many of the symptoms attributed to mTBI could be conceptualized as common everyday experiences which may account for much of the experience of postconcussion syndrome. In this empirically supported model, people with mTBI report lower symptomatology pre-incident when compared to current symptomatology, and even report lower *retrospective* symptomatology than healthy controls. In other words, the people with mTBI may be misattributing their current complaints to the mTBI when they compare the complaints to overly positive recollections of pre-injury experiences. Instead of attributing their current headache or fatigue to a long and stressful day, for example, people with postconcussion syndrome may attribute the symptoms to the injury given that they expect the mTBI to cause headaches and fatigue.

Another source of evidence for this model is that participants with depression, healthy athletes, and healthy controls can all accurately anticipate postconcussion syndrome symptoms for a hypothetical mTBI, but anticipate these symptoms at different

rates (Gunstad & Suhr, 2001). Healthy athletes anticipate fewer symptoms than other groups, perhaps because they have witnessed other athletes making a rapid recovery from mTBI. Expectation as etiology may therefore partially explain quicker recovery times following sports-related mTBI compared to mTBI incurred in other contexts. Specifically, the athletes would be less likely to attribute everyday complaints to mTBI in the post-acute period, when they expect their mTBI to have resolved.

Good old days bias. Gunstad and Suhr (2001) suggested that “expectation as etiology” may be too narrow a model, and that *any* negative event, not just mTBI, may trigger a similar “good old days” bias, with an accompanying focus of the past being better than the present. Recent research supports this explanation, revealing that the retrospective ratings of pre-accident symptoms in an mTBI sample were less severe than the retrospective ratings of controls who had not been involved in an accident (Lange, Iverson, & Rose, 2010). Unsurprisingly, people with mTBI also rated their current symptoms as more severe than their pre-accident symptoms and more severe than the current ratings of the control group. Furthermore, a recent study found that low pre-accident postconcussion symptom reporting is not specific to participants with mTBI, but is also reported by participants with orthopaedic injuries, finding that this good old days bias was most prominent in examinees who were seeking compensation following mTBI or orthopaedic injury (Silverberg et al., 2016).

In a systematic review of the literature, people who sustained mTBI, compared to control participants, reported more postconcussion symptoms such as headache, fatigue, and self-perceived cognitive deficits (Cassidy et al., 2014). These symptoms, however, are *not specific to mTBI*. Although reported at higher rates by mTBI patients up to one

year following injury compared to healthy controls, these symptoms are no more prevalent in mTBI patients than in people with orthopedic injuries (Cassidy et al., 2014). These findings converge in a reconceptualization of the etiology of postconcussion symptoms as psychological and social rather than pathophysiological.

Diagnosis threat. Diagnosis threat, a variant of stereotype threat in which the individual is cued to the presence of head injury history, has been posited as another explanation of reduced neuropsychological test results in mTBI (Suhr & Gunstad, 2002). The stereotype threat model posits that individuals from negatively stereotyped groups will experience undue pressure to avoid confirming others' biases in areas in which their group is perceived to be less capable (Steele, 1997). A woman taking a mathematical test might for example experience such pressure. The person will concurrently experience the pressures that any person would experience in the situation such as pressures to succeed and to be perceived as competent (Steele, 1997). The pressure to disconfirm others' biases undermines the individual's performance and causes social underperformance phenomena, whereby certain groups are underrepresented in certain fields—for example, stereotype threat may cause underrepresentation of women in science, technology, engineering, and mathematics fields (Spencer, Logel, & Davies, 2016).

The diagnosis threat model is an extension of the stereotype threat model. In the diagnosis threat model, being reminded of the effects of mTBI on performance cues negative expectations for performance in the individual with mTBI (Suhr & Gunstad, 2002). The cuing increases anxiety and inhibits effort leading to poor performance. Initial supporting evidence for this model comes from a study where participants with a history of mTBI who were assigned to a diagnosis threat group performed worse on intelligence

and memory testing when compared to matched controls with mTBI (Suhr & Gunstad, 2002). Participants in the diagnosis threat group rated themselves as being less confident, performing worse, and putting forth less effort than mTBI controls. Of note, neither group of college students with a history of mTBI was involved in litigation or disability claims. These findings may indicate that diagnosis threat itself, even in the absence of external incentive, can contribute to negative alterations in performance following mTBI. More recent studies evaluating the diagnosis threat model have shown more modest differences between groups that are or are not cued to the effects of mTBI (Carter-Allison, Potter, & Rimes, 2016), and indicate that diagnosis threat may affect self-reported functioning more than neuropsychological performance (Ozen & Fernandez, 2011).

Confluence of factors. In a study comparing subjective complaint of cognitive impairments with neuropsychological test results across TBI severity, Jamora, Young, and Ruff (2012) found that participants with moderate-to-severe injuries performed more poorly on memory and attention tasks than those with mTBI, but did not differ in executive functioning scores. Conversely, people with mTBI rated themselves as having significantly more attention, concentration, and executive functioning impairment than those with moderate-to-severe injuries. The groups did not differ in self-reported memory, learning, or language impairment. For the mTBI group, self-reported emotional dysfunction scores predicted self-reported cognitive impairment. Self-reported emotional impairment was unrelated to self-reported cognitive impairment for the more severe group. However, it should be noted that these patterns emerged after individuals with

noncredible profiles were dropped from the data set and that the majority of both groups were litigants.

The authors explained the discrepancy between self-reported cognitive impairment and cognitive performance in several ways. Examinees with mTBI may have misattributed difficulties to the injury or catastrophized everyday failure (Jamora et al., 2012). Alternatively, participants with mTBI may have had nonhead injuries that led to litigation and contributed to complaints. Participants with mTBI may also have had emotional or personality factors that contributed both to compensation-seeking following a mild injury and to cognitive complaints. Compared to examinees with more severe injuries, examinees with mTBI may also have had greater demands and less social support, which made their cognitive concerns more salient to them. Another possible explanation may be poor awareness of deficits in the moderate-to-severe TBI group (i.e., severe injuries impaired insight).

Overall, a large number of factors may influence poor neuropsychological performance and high symptom complaint in the chronic period following mTBI. These factors are related to, but not a direct result of, the injury itself.

TBI and pain. Another factor that complicates the mTBI picture is chronic pain. Chronic pain is paradoxically more prevalent in mTBI populations as compared to those with more severe TBIs. In a systematic review of the literature, Nampiaparampil (2008) found that 75.3% of mTBI patients reported chronic pain, compared to only 32.1% of moderate-to-severe TBI patients.

An investigation of the relationship between chronic pain and mTBI showed no association between various neuropsychological test performances between high and low

postmorbid pain group (Jamora, Schroeder, & Ruff, 2013). The high pain group, however, complained of higher levels of anger, aggression, anxiety, depression, paranoia, and suspicion compared to those in the low pain group. The high pain group also subjectively endorsed worse attention, concentration, executive functioning, somatic complaints, activities of daily living, and psychosocial integration as compared to the low pain group. This research indicates that chronic pain is a common clinical comorbidity of mTBI and that the connection between chronic pain and subjective complaints of impaired cognition may be related to emotional difficulties and catastrophization of common difficulties.

Overall, several complex factors may contribute to ongoing symptom complaint and lower neuropsychological test scores in the chronic period following mTBI. Determining the extent to which each factor may be driving slow recovery in postconcussion syndrome is difficult. Importantly, pathophysiological injury effects cannot wholly explain postconcussion syndrome. In addition to the previously explored psychological and social factors, external monetary incentive is another important variable that affects recovery from mTBI. The following section discusses the negative impacts of litigation and compensation-seeking following injuries.

Litigation and Health Outcomes

Litigation and physical health outcomes. In addition to the nature of initial injuries and comorbidities, compensation-seeking litigation significantly affects health outcomes following TBI and nonhead injuries (Spearing, Connelly, Gargett, & Sterling, 2012). The literature on compensation seeking in nonhead injuries provides some context for the direct impact of litigious factors on recovery in the absence of TBI. This

discussion illustrates litigation's significant contribution to poor prognosis regardless of the nature of the injuries.

Spearing et al. (2012) conducted a systematic review of research about compensation-related factors and health outcomes in adults following whiplash injury. Inclusion criteria for the review included longitudinal design, adult participants, and comparison of compensation and health outcome. Exclusion criteria included serious neck injuries, chronic pain, TBI, and other injuries, proxy measures of health outcome, and lack of control group. The review ultimately included eleven studies, and method variability precluded meta-analysis.

Poorer health outcomes were related to compensation seeking in seven studies with measures of compensation (i.e., having sought a lawyer, present litigation, compensation claim, and previous claim). The authors note that most studies did not address reverse causality (i.e., people who make claims may be more severely injured, driving the association between compensation seeking and health outcomes). Overall evidence from this systematic review was equivocal.

Murgatroyd et al. (2015) systematically reviewed the literature regarding the effects of litigious financial compensation seeking on health outcomes following musculoskeletal injury. This review included 29 studies with prospective designs with at least 6-month follow up, with musculoskeletal injuries of adults and an aim at determining prognostic factors. The authors excluded studies with participants with dementia, cognitive impairment, moderate-to-severe TBI, spinal cord injury, organ injury, or psychological injury such as PTSD.

Twelve studies showed that compensation seeking was associated with poorer physical functioning whereas eight did not. In contrast, all five studies that evaluated psychological function found associations between poorer psychological function and compensation seeking. Compensation seeking was linked to pain in 9 of 15 studies. All three studies that explored a relationship between having legal representation and psychological wellbeing found legal representation was associated with worse wellbeing. Five of seven studies showed that hiring legal representation was related to poorer physical function. Overall, there is some evidence that having a lawyer and compensation seeking are associated with poorer physical and psychological functioning and greater pain after injury.

An inception study was conducted to ascertain predictors of fatigue one year following mTBI (de Leon et al., 2009). Participants were identified in emergency departments as having sustained mTBI, with loss of consciousness ≤ 30 min, posttraumatic amnesia ≤ 24 hr, and Glasgow Coma Scale ≥ 13 . Presence or absence of posttraumatic amnesia and loss of consciousness did not predict fatigue. The most robust predictor of greater 12-month fatigue was baseline fatigue, followed by being in litigation, marital status (divorced, widowed, or single), having a medical disability, and having sought mental health treatment. Sustaining a TBI was unrelated to fatigue when controlling for baseline predictors.

These results suggest that within mTBI, factors other than injury severity likely account for long-term health outcomes. Overall, there is some evidence that litigation and compensation seeking are associated with poorer health outcomes for individuals with mTBI and other injuries. Young (2008) posited that iatrogenic litigation process might

cause poor health outcomes, whereby monetary and nonmonetary secondary gain and the need to demonstrate impairment trigger somatization.

Litigation and psychological outcomes. Litigation is associated with poor psychological wellbeing in addition to poor physical health. Bay and Donders (2008) studied risk factors for depressive symptoms in TBI. Perceived stress was the strongest risk factor, followed by pain and being involved in litigation. Another study showed that people involved in litigation who perceived the other driver in the accident to be at fault were more depressed and less likely to return to work than those who perceived the other driver to be partially responsible or not at fault (Thompson, O'Donnell, Stafford, Nordfjaern, & Berk, 2014). Furthermore, the presence of depressive symptoms mediated the relationship between fault attribution and return to work (Thompson et al., 2014).

A meta-analysis of mental health in compensation seeking participants showed that claimants who sought compensation had poorer mental health than participants who were not seeking compensation (Elbers, Hulst, Cuijpers, Akkermans, & Bruinvels, 2013). Baseline mental health differences (including worse self-reported anxiety, depression, and posttraumatic stress symptoms) accounted for 75 percent of the compensation seeking difference. That said, people who were compensation seeking were also slower to recover from mental health difficulties compared to people who did not seek compensation.

These studies suggest that being involved in litigation is consistently associated with poorer self-reported mental health. The exact causal relationship between litigation and psychological outcomes, however, remains unclear. Posited mechanisms include more severe injuries in compensation seeking groups, the development of a

“compensation seeking mindset,” more anger and blame in compensation seeking cases, secondary gain (i.e., financial or other incentives to remain unwell), and secondary victimization—whereby the claimant is stressed by the litigation process itself, leading to poorer mental health (Elbers et al., 2013). Importantly, poor psychological and physical health in compensation seeking examinees can affect the results of neuropsychological assessments.

Litigation and neuropsychological assessment. A meta-analysis of 29 studies investigated predictors of neuropsychological outcomes following mTBI (Belanger, Curtiss, Demery, Lebowits, & Vanderploeg, 2005) and showed that mTBI was associated with deficits in global cognitive function, attention, executive functions, fluency, acquisition memory, delayed memory, language, and visuospatial skill. Results were, however, quite heterogeneous. Time since the injury (acute [≤ 3 months] or post-acute), and litigation status accounted for many of the differences between studies. Time since injury moderated the effects of mTBI on neuropsychological test results, with reduction in all but visuospatial skills deficits. Litigation status accounted for post-acute visuospatial deficits (i.e., all acute studies that measured visuospatial function had nonlitigant samples, and all post-acute studies had litigant samples). Deficit severity was similar in the acute period for litigant and prospective samples, but the prospective samples had test results that were equivalent to controls by three months postinjury. In contrast, the differences between litigant participants and controls *increased* in the post-acute period. Furthermore, litigant participants’ deficits were similar across all cognitive domains, whereas nonlitigant participants only had fluency and delayed memory deficits. In other words, litigant participants showed decreased function across all cognitive

domains over time, whereas nonlitigant participants had specific deficits that healed within three months of the injury.

These findings reveal an expected cognitive profile for individuals following mTBI, namely deficits in verbal fluency and delayed memory that resolve quickly and completely. Individuals in litigation do not follow this expected profile of deficits or recovery (Belanger et al., 2005). Presence of performance validity tests (PVTs) in litigation studies did not change the effect sizes significantly. Interestingly, clinic-based samples (for which there were only post-acute data) fared similarly to litigant participants on neuropsychological tests. Litigious and post-acute clinic samples may be similarly unrepresentative of mTBI patients, in that they are continuing to report symptoms and are seeking treatment more than three months following the injury when symptoms have largely resolved for most people who incurred an mTBI.

Overall, involvement in litigation is associated with a pattern and chronicity of deficits that was markedly different from those found in individuals assessed in the acute period following mTBI. PVT failure did not account for the differences, and the authors offered several interpretations, including psychological factors, enduring neurological dysfunction, and poor coping (Belanger et al., 2005). In sum, the chronic and worsening symptoms of litigants are not wholly attributable to the injury.

Expectation, pain, compensation seeking, litigation, and pre-morbid factors can all contribute to the pattern of symptoms and cognitive performance following mTBI. As such, an important element of post-mTBI assessment is ascertaining the particular factors that contribute to the examinee's level of functioning. Further, it is important to examine the degree to which examinees assessed following mTBI are providing an accurate

picture of their current level of neuropsychological functioning. PVTs are an important part of determining the credibility of neuropsychological test results and will be discussed in the following section.

Performance Validity

Definition. Performance validity is the extent to which performance on a neuropsychological test reflects cognitive ability (Larrabee, 2012). It is measured with stand-alone PVTs designed specifically to detect noncredible performance, as well as atypical performance on neuropsychological tests originally designed to tests an array of neurocognitive functions (Larrabee, 2012). PVTs differ from symptom validity tests, which measure the credibility of symptom report (e.g., symptom exaggeration, socially desirable responding, random responding; Larrabee, 2012). Base rates of failure (BR_{FAIL}) of PVTs are high, occurring with 30% to 40% of individuals in litigation or compensation seeking (e.g., Howe, Anderson, Kaufman, Sachs, & Loring, 2007; Larrabee, 2012).

Multiple factors other than deliberately exaggerating deficits can result in noncredible performance, including lack of interest or poor engagement and nonmonetary incentives such as maintenance of care or avoidance of responsibility (Schutte & Axelrod, 2012). PVTs also cannot reliably distinguish somatic symptom disorders from intentional exaggeration (Boone, 2007). Due to the variety of factors that can contribute to noncredible performance, PVT failure cannot connote deliberate feigning (Boone, 2007), but effectively indicate whether examinees' neuropsychological test scores reflect their underlying cognitive abilities.

Research design. PVT research follows two primary designs. In simulation designs, a group of healthy participants is instructed to simulate a brain injury without

being detected as feigning, whereas another group is instructed to complete testing to the best of their ability (e.g., Bashem et al., 2014). Simulation design prioritizes internal validity over external validity by directly manipulating test engagement to maximize group homogeneity.

In contrast, criterion group designs define groups based on either litigation status, injury severity (i.e., mTBI compared to moderate-to-severe TBI), or failure of other PVTs (e.g., Victor, Boone, Serpa, Buehler, & Ziegler, 2009). Criterion group designs offer greater generalizability than simulation designs but have poorer experimental control. None of the criterion group methodologies can classify individuals with complete accuracy (Bigler, 2015; Larrabee, 2012). The selection of the particular criterion PVT and cutoffs may dramatically affect BR_{FAIL} . Even with more than one PVT, results will vary unless studies use the same combination of PVTs and cutoff scores. The current study used a descriptive design to explore the effects of demographic and psychiatric variables on BR_{FAIL} to avoid the controversy of classifying profiles as valid or invalid.

PVT use guidelines. Guidelines for neuropsychological assessment and consultation published by the Board of Directors of the American Academy of Clinical Neuropsychology (2007) suggest that performance validity assessment is essential in neuropsychological assessment, especially in forensic assessment and in the presence of financial incentives. Resistant behaviour, atypical patterns of performance, and PVT failure all indicate noncredible performance. The guidelines suggest using converging indicators of noncredible presentation to determine the veracity of results. Despite the consensus that determining test credibility is important, the research and clinical interpretation of PVTs remain controversial.

PVT controversy.

PVT limitations. The most pressing concern about PVTs is the circular reasoning used to judge credibility. Bigler (2012) argued that PVTs alone determine examinee, without any independent, direct measures of credibility or test engagement. Some authors have also expressed concern that the use of multiple PVTs may increase false-positive identification of noncredible performance (Odland, Lammy, Martin, Grote, & Mittenberg, 2015). It is also assumed that failed PVTs imply suboptimal validity throughout the assessment, which may or may not be true (Bigler, 2012; Boone, 2007).

Boone (2007, 2009) used case studies to illustrate that examinees may perform noncredibly at distinct times during assessment (e.g., noncredible performance early or late in the process, or only during purported transitory events like panic attacks). Thus, PVTs may miss noncredible performance, or detect a brief period of noncredible performance in an otherwise credible assessment. Bigler (2012, 2015) has also expressed concern about misclassifying individuals with “near-pass” PVT performance.

The dichotomization of an individual’s performance as either “credible” or “noncredible” is also somewhat arbitrary. Forensic standards accommodate variability in methods, instruments, and cutoffs used to determine performance validity, leading to the use of a variety of PVTs. An unintended consequence of multiple PVT use is method variance in which criterion groups can vary dramatically between studies causing variable cutoffs and variable signal detection profiles (i.e., sensitivity, specificity, positive and negative predictive power).

An example of methodology affecting cutoff score development may help to illustrate this point. Whiteside et al. (2015) found sensitivity (i.e., true positive rate) of

.23 and specificity (i.e., true negative rate) of .91 using a T -score ≤ 24 cutoff for Animal Fluency, where examinees name as many animals as possible in 60 seconds (Ruff, Light, Parker, & Levin, 1996). In contrast, Sugarman and Axelrod (2015) found sensitivity of .42 at the same specificity with the more liberal T -score ≤ 32 cutoff on the same task. The researchers' differing methodologies may account for this discrepancy.

Whiteside et al. (2015) compared people with mTBI who were compensation seeking and who failed at least two of 11 PVTs to a group of noncompensation seeking people with previously diagnosed severe TBI who passed all of the PVTs administered. Sugarman and Axelrod (2015), in contrast, compared individuals with mTBI who failed at least two of six PVTs to individuals with mTBI who passed all six PVTs. The comparison groups differed between credible performers with mTBI and credible performers with severe TBI, confounding the results. The inclusion of 11 PVTs—compared to six PVTs in the Sugarman and Axelrod (2015) study—also gave participants in the Whiteside et al. (2015) study more opportunity to fail PVTs and be included in the noncredible mTBI group or excluded from the credible severe TBI group. The criterion PVTs, except the Test of Memory Malinger (TOMM; Tombaugh, 1996) were also different in each study. PVTs have different signal detection profiles, and choosing different PVTs may have further affected the inclusion of participants into the groups across the studies.

Overall, Whiteside et al. (2015) used a comparison group that sustained more severe injuries, and would thus be more likely to have truly impaired verbal fluency, and included different criterion PVTs, both of which may have contributed to the indication of a more stringent cutoff score to maintain adequate specificity. Both of the studies used

common, accepted methodologies, therefore neither cutoff score can be easily determined as more correct. It is clear that subsequent use of these cutoff scores, however, would lead to very different signal detection and categorization of examinees. Using different cutoff scores then cyclically perpetuates instrumentation bias and divergent results in future research. Overall, current research paradigms and the implementation of PVTs are contentious, despite their utility and importance.

PVT strengths. Despite their limitations, PVTs provide important information. PVT performance is robust to many neuropathological and psychiatric conditions, excluding psychotic disorders, dementias, and intellectual disabilities (Goldberg et al., 2007; Larrabee, 2012). People can pass PVTs with experimentally induced acute pain (Etherton, Bianchini, Ciota, & Greve, 2005; Etherton, Bianchini, Greve & Ciota, 2005). PVTs also reliably differentiate credible and noncredible chronic pain patients (Greve, Bianchini, & Brewer, 2013). Removing individuals who fail PVTs from analyses uncovers expected patterns of neuropsychological results that are otherwise obscured (Larrabee, 2012). For example, California Verbal Learning Test (Delis, Kramer, Kaplan, & Ober, 1987) scores discriminated TBI patients with abnormal brain imaging from those with normal scans only once participants with failed PVTs were excluded (Green, 2007b).

These findings strengthen the argument that PVT failures are indicative of noncredible performance across neuropsychological batteries, and that “impaired” results of neuropsychological tests in the presence of PVT failures likely do not represent examinees’ true cognitive capacities. Although cutoff scores for PVTs are arbitrary, optimizing cutoff scores to maximize specificity (Larrabee, 2012), and classifying

noncredible performance by multiple PVT failures renders PVTs a highly sensitive and specific means to determine neuropsychological test result veracity (Victor et al., 2009).

mTBI and performance validity. Paradoxically, individuals with mTBI often fail PVTs at much higher rates than those with moderate-to-severe TBI (Carone, 2008; Green, Iverson, & Allen, 1999; Green, Rohling, Lees-Haley, & Allen, 2001; Mittenberg, Patton, Canyock, & Condit, 2002; Webb, Batchelor, Meares, Taylor, & Marsh, 2012; West, Curtis, Greve, & Bianchini, 2011). Bigler (2014) argued that the higher BR_{FAIL} for people with mTBI might be due to neurophysiological changes resulting from fatigue, which is a common symptom of mTBI. He argued that fatigue might impair attentional, working memory, and inhibitory systems making PVT tasks more difficult, resulting in higher BR_{FAIL} . Bigler (2014) admitted that no one has yet empirically tested this hypothesis, and it is unclear why this mechanism would operate in individuals with mTBI with greater frequency than in individuals with moderate-to-severe TBI. Other possibilities for these findings may be symptom expectation (Mittenberg et al., 1992), or diagnosis threat (Larrabee & Rohling, 2013), where the testing situation triggers expectations and fears of detrimental cognitive effects caused by the TBI in individuals with mTBI, resulting in poorer performance. Monetary incentive is another possible contributor to the differences in PVT BR_{FAIL} between examinees with mTBI and examinees with more severe injuries.

Litigation and performance validity. Concerns about performance validity are particularly germane in forensic assessments, where the examinee stands to gain from appearing impaired. These concerns have been reflected in common neuropsychological credibility classification.

Slick Criteria. Slick et al. (1999) produced widely used (Lezak et al., 2012) criteria for the diagnosis of malingered neurocognitive deficit, which mimic the *Diagnostic and Statistical Manual of Mental Disorders* (American Psychiatric Association, 2013) classification system:

- A. Presence of a substantial external incentive.
- B. Evidence from neuropsychological testing.
 - 1. Definite negative response bias: below chance performance on one or more PVTs.
 - 2. Probable response bias: performance on two or more PVTs consistent with feigning.
 - 3. Discrepancy between test data and known patterns of brain functioning.
 - 4. Discrepancy between test data and observed behaviour.
 - 5. Discrepancy between test data and reliable collateral reports.
 - 6. Discrepancy between test data and documented background history.
- C. Evidence from self-report.
 - 1. Self-reported history is discrepant with documented history.
 - 2. Self-reported symptoms are discrepant with known patterns of brain functioning.
 - 3. Self-reported symptoms are discrepant with behavioural observations.
 - 4. Self-reported symptoms are discrepant with information obtained from collateral informants.
 - 5. Evidence of exaggerated or fabricated psychological dysfunction.

D. Behaviours meeting necessary criteria from groups B or C are not fully accounted for by Psychiatric, Neurological, or Developmental Factors.

A diagnosis of definite malingered neurocognitive deficit requires criteria A, B1, and D. A diagnosis of probable malingered neurocognitive deficit requires criterion A, two or more of B2-B6 or one of B2-B6 and one of C1-C5, and criterion D. A diagnosis of possible malingered neurocognitive deficit requires criterion A, one or more of C1-C5, and D or criteria met for definite or probable malingered neurocognitive deficit except for criterion D.

The current study did not utilize the Slick Criteria for several reasons. Firstly, the presence of external incentive is a false dichotomy: it is either present or ultimately unknown. Ruling out external incentive involves the formidable task of proving a negative. Researchers too often take the absence of evidence to imply the evidence of absence. Lack of knowledge about incentives does not equate to absence. There is also no effective way to differentiate between somatic symptom and related disorders and “malingering” (Boone, 2007)—“noncredible neurocognitive function” was suggested by Boone (2007, p. 38) as better terminology for this reason. The label of “malingering” is also unnecessary, as the determination of test result credibility, and not the intent, is the purpose of PVTs in neuropsychological assessment.

Another difficulty with the use of the external incentive criterion is that legitimately impaired people may have to obtain legal representation to secure care. As an illustration that is germane to the current study’s sample, a Financial Services Commission of Ontario (2011) report showed that 99% of motor vehicle accident claimants who dispute insurer decisions in Ontario utilize legal services. Thus, criterion

A, the presence of substantial external incentive, appears unnecessary in the determination of noncredible performance as it contributes no unique information to the assessment model (Boone, 2007), and would, perhaps unfairly, pre-determine suspicion of motives in those undergoing neuropsychological assessment.

PVT types. Many PVTs are available to neuropsychologists for research and clinical practice. These PVTs can be broadly categorized as either stand-alone or embedded measures.

Stand-alone. Stand-alone PVTs are designed specifically to detect noncredible performance while appearing to be tests of cognitive function such as memory (Greve & Bianchini, 2004; Schutte & Axelrod, 2012). Many studies and several systematic reviews have compared PVT BR_{FAIL} with groups of simulators, compensation-seekers, and those with noncredible performance on other PVTs (e.g., Sollman & Berry, 2011; Vickery, Berry, Inman, Harris, & Orey, 2001). Control groups have included individuals with a wide variety of conditions, including psychiatric pathology, attention deficit hyperactivity disorder, orthopaedic injury, TBI of varying severity, learning disability, intellectual disability, and dementia.

Stand-alone PVTs remain the gold standard to determine credibility, and a thorough review of the extant literature is beyond the scope of the current text. Despite, or perhaps because of their extensive validation, stand-alone PVTs may be susceptible to coaching (Brennan et al., 2009; DiCarlo, Gfeller, & Oliveri, 2000; Suhr & Gunstad, 2000). These measures are often familiar to lawyers, who may directly coach their clients to perform well on these tests (Brennan et al., 2009). Examinees stand only to gain from performing well on stand-alone PVTs, as they exclusively measure credibility.

Tests of cognitive function with embedded PVTs, however, have a window of impaired performance, above which examinees will appear functional, and below which examinees will appear noncredible. Examinees with true impairment will spontaneously fall within the window of impaired performance, but these scores are difficult to simulate (Schutte & Axelrod, 2012).

Embedded. Embedded validity indicators measure credibility and are derived from a larger test of some cognitive function (e.g., working memory, attention, processing speed, motor functions; Schutte & Axelrod, 2012). These measures have several advantages over stand-alone PVTs: they simultaneously measure ability and credibility; may be less affected by coaching (Ashendorf, O’Bryant, & McCaffrey, 2003; Schutte & Axelrod, 2012); and allow credibility assessment in multiple cognitive domains without requiring additional time or resources (Boone 2009; Greve et al., 2013). The current study’s embedded validity indicators are described in the methods section. A thorough review of every measure is beyond the scope of the current work.

Many empirically supported cutoff scores (and associated differing BR_{FAIL}) are available for embedded and stand-alone PVTs. As previously mentioned, arbitrary cutoff scores and a priori group assignment criteria contribute to cutoff scores variability. The following chapters will review the extant literature on the contribution of demographic and psychiatric characteristics to PVT performance and provide rationales for the current study.

CHAPTER 2

Culture, Linguistics, Demographics, and Performance Validity

Culture and Neuropsychological Assessment

Culture affects neuropsychological assessment (Mitrushina, Boone, Razani, & D'Elia, 2005), and it is important to consider cultural factors in the administration and interpretation of neuropsychological tests. Culture itself is a broad and complex concept, which refers to many aspects of people's experiences, behaviour, and means of expression (Matsumoto & Juang, 2016). Culture includes but is not limited to food, language, modes of dress, religious beliefs, and traditions. The current research will focus primarily on country of origin, educational, and linguistic aspects of culture, with the caveat that many other aspects of culture are important and may affect neuropsychological assessment. In certain cases language, education, and country of origin will be referred to as demographic characteristics, as they are also simple statistical characteristics of samples.

Neuropsychological assessment assumes a native level of English language proficiency (Lezak et al., 2012). Most tests and norms are developed with a primarily or exclusively White, Anglophone sample (Mitrushina et al., 2005; Strauss, Sherman, & Spreen, 2006). A link between culture and English language proficiency has been reported with differences in neuropsychological test performance across multiple studies and multiple measures (e.g., Boone et al., 2007), but there is a paucity of research on the topic. Boone et al. (2007) found differences in scores on naming, visuoconstruction, verbal repetition/attention span, nonverbal processing, and executive functioning tasks in different linguistic and cultural groups. Despite comparable clinical diagnoses, African

American, Hispanic, and Asian patients scored lower on these tasks than Nonhispanic Caucasian patients, with differences related to years in the United States and age at which English was learned. Additionally, years of education, acculturation, and reading ability account for much of the differences in variance between African American and Caucasian American participants (Manly, Byrd, Touradji, & Stern, 2004). Differences in neuropsychological test results between cultural groups carry the risk of overpathologizing individuals from minority cultures (Mindt, Byrd, Saez, & Manly, 2010).

Neuropsychologists are beginning to recognize the importance of demographic variables in the development of tests and the interpretation of their results. As an example, researchers provide adjusted norms with cultural and educational corrections (e.g., Heaton, Miller, Taylor, & Grant, 2004; Mitrushina et al., 2005). English language proficiency specifically, however, has remained an understudied area in neuropsychological assessment (Mindt et al., 2010). In addition to normative adjustments, professional associations have also developed guidelines for competent assessment of members of minority groups.

Guidelines for Multiculturally Competent Assessment

The American Academy of Clinical Neuropsychology developed guidelines for competently assessing members of minority groups (Board of Directors, 2007) that outline expectations for competency, knowledge, and experience in neuropsychological assessment with the minority group(s) to which the examinee belongs. If neuropsychologists are unable to demonstrate their competence, they must demonstrate that they have attempted to refer the examinee to a more qualified colleague, and have

considered the benefits of assessment to outweigh possible harm. Neuropsychologists must then demonstrate that they have attempted to offset limitations of their competence through consultation with colleagues and review of research. The guidelines caution that the use of a language interpreter and translated or adapted instruments may threaten the validity of results. The guidelines further suggest that neuropsychologists incorporate nonstandardized sources of supplementary information when culturally appropriate tests and norms are not available.

The Canadian Psychological Association also provides guidelines for the competent assessment and treatment of members of minority groups (Canadian Psychological Association, 2000). The Code of Ethics and Standards suggests that psychologists be empathic and informed about cultural factors and vulnerabilities, as well as being attentive to the potential for harm and benefit from providing services. Additionally, the Code suggests choosing interventions that are empirically supported, considering the needs and characteristics of the client, and consulting with persons relevant to the client's culture.

Despite the clear expectation of referral to culturally informed practitioners and use of culturally specific tests and norms, the available evidence suggests that at present these ideals are unfeasible. Canadians speak over 200 languages (Statistics Canada, 2011), with 30.5% of Canadians reporting a language other than English or French as their first language and over 2.1 million Canadians exclusively speaking languages other than English or French at home.

Many minority groups in North America have small enough populations that it would be impractical or impossible to develop and maintain norms, translations, and tests

for each cultural or linguistic group. Neuropsychologists have the same ethical responsibility to provide equitable and competent service to members of all minority groups, regardless of size. The development of appropriate tests and norms for the largest minority linguistic groups in North America (i.e., Spanish and French speakers) would not meet the needs of individuals from smaller groups, such as Kurdish and Thai speakers. Furthermore, adapted tests and norms have been criticized as not addressing the heterogeneity of the ethnic or linguistic groups that they are purported to serve—for instance; the normative groups may differ in important ways from the examinee (Elbulok-Charcape, Rabin, Spadaccini, & Barr, 2014). As an example, a normative group composed of Mexican-Americans who have lived in California for several decades would not speak the same dialect of Spanish or have the same culture as an individual who recently immigrated from Puerto-Rico to Ottawa.

A more pragmatic solution may be to explore relevant transcultural factors that affect test results across a wide range of minority groups and levels of acculturation rather than attempting to create tests and norms for all identifiable segments of the population. Development of correction factors applied to standard neuropsychological tests based on this paradigm would aid in the equitable provision of services to multiple groups. The current research thus explored how PVT performance relates to common factors that can be ascertained across cultural groups, such as education, limited English proficiency, and immigration status.

Current Practice of Multicultural Neuropsychology

A recent survey of 512 doctoral level neuropsychologists in the United States and Canada assessed the current trends in neuropsychological assessment of ethnic minorities

(Elbulok-Charcape et al., 2014). The authors found that 91% of respondents were White, but that respondents spent 65.7% of time with White clients, 15.7% with Black or African American clients, 11.7% with Latino or Hispanic clients, 4.2% with Asian, 1.2% with Native American, and 0.5% of their time with Native Hawaiian/Pacific Islander clients.

Although being White does not preclude culturally competent assessment of clients from minority cultures, the imbalance between the examiner and examinee demographics is clear. Only 15.2% of respondents conducted neuropsychological assessments in a language other than English. Except Spanish ($n = 47$) and French ($n = 13$), there were fewer than five respondents who conducted assessments in any one language other than English, with only 18 languages represented. It is clear that the aspirational guideline of referring clients to neuropsychologists who can competently conduct assessments in clients' preferred language is at odds with reality.

Further, Elbulok-Charcape et al. (2014) found that despite the apparent dearth of neuropsychologists who are fluent in languages other than English, Spanish, or French in North America, 69% of respondents reported that they typically refer clients to neuropsychologists who are fluent in the patient's language. It is not clear exactly how respondents made these referrals, considering that there are very few neuropsychologists from diverse cultural and linguist backgrounds practicing in the United States and Canada. Similarly, respondents endorsed using culturally specific norms, less culturally biased tests, and adjustment of test scores. The specific norms, tests, and adjustments used were not specified. Despite the American Academy of Clinical Neuropsychology's

(Board of Directors, 2007) caution, 41% of respondents reported frequently employing interpreters in assessment.

Respondents identified the lack of appropriate norms and tests as impediments to culturally competent assessment (Elbulok-Charcape et al., 2014). Neuropsychologists also reported that it is difficult to find colleagues to whom they could refer or whom they could approach for consultation, and reported a lack of trained neuropsychologists, psychometrists, and training opportunities. The authors concluded that there are several issues with the current practice of neuropsychology with culturally diverse individuals, including insufficient training in culturally competent practice and a lack of neuropsychologists with linguistic proficiency in a variety of languages.

Despite clear consensus that practitioners are ethically obligated to use tests that have been validated with the cultural group to whom the examinee belongs, and to include assessments of credibility in neuropsychological testing (Board of Directors, 2007), very little research has explored the cross-cultural validity of PVTs (Boone et al., 2007). Most cross-cultural PVT research has been conducted with Spanish speaking participants in the United States (Boone et al., 2007). The following section reviews findings from multicultural PVT research.

Cultural Factors and Performance Validity

This section reviews the extant PVT literature that includes culturally and linguistically diverse participants. PVT research has only been conducted in English, Spanish, select Asian languages (described below), and Western European languages (where findings are similar to those in North America; e.g., Merten, Thies, Schneider, & Stevens, 2009; Stulemeijer, Andriessen, Brauer, Vos, & van der Werf, 2007). To this

author's knowledge, there is no information about the impact of cultural and linguistic factors on PVT performance in Africa, Western Asia, Southeast Asia, Central America or South America. Findings from studies conducted in North America with culturally diverse samples, studies in Spanish, and studies conducted with Asian participants are discussed below.

An archival study was conducted with neuropsychological data from 168 individuals assessed in English at a public hospital in Los Angeles ($N = 168$; Salazar, Lu, Wen, & Boone, 2007) to address the gap in the literature in PVTs with minority culture and limited English proficiency populations in the United States. One hundred and thirty-nine participants spoke English as their first language, and 28 spoke English as a second language. Eighty-five participants were Anglo-Caucasian, 32 were African American, 32 were Hispanic American, and 19 were Asian American. Personal injury litigants, people who met criteria for dementia or who had a Wechsler Adult Intelligence Scale-Revised or Wechsler Adult Intelligence Scale-III (WAIS-R; Wechsler, 1981; WAIS-III; Wechsler, 1997a, respectively) Full-Scale Intelligence Quotient < 70 were excluded.

The goal of the study was to propose adjusted cutoffs for the PVTs that would result in $BR_{FAIL} \leq .10$ for each group (i.e., fewer than 10% of participants failed the given cutoff). Several PVT cutoffs were examined, including Digit Span age corrected scaled score ≤ 5 , Reliable Digit Span ≤ 6 , Rey 15-IR (recall + [recognition – false positives]) < 20 (Rey, 1964), Rey Auditory Verbal Learning Test (Rey, 1941) recognition ≤ 7 , Rey Auditory Verbal Learning Test effort equation ≤ 12 , Dot Counting Test (Boone, 2002) E-score ≥ 17 , Warrington Recognition Memory Test-Words (Warrington, 1984) < 33 , Rey-Osterrieth (RO) effort equation ≤ 47 (Lu, Boone, Cozolino, & Mitchell, 2003), and

RO/AVLT discriminant function $\leq -.40$ (Sherman, Boone, Lu, & Razani, 2002). The BR_{FAIL} for the cutoffs were variable, with no clear pattern. It is possible that heterogeneity in education, English language proficiency, or other factors accounted for the variable BR_{FAIL} . In general, the BR_{FAIL} , despite being variable, were acceptable across ethnic groups for most of the cutoffs tested.

There were some limitations to the study. Despite the inclusion of members of minority groups, each group was small, and the groups were diagnostically and culturally heterogeneous. The study, however, lends some evidence for the use of these PVTs in African American, Hispanic American, and Asian American individuals, as well as individuals with limited English proficiency. The remainder of this section will review research conducted with participants from a variety of cultural and linguistic groups in their first languages.

PVTs administered in Spanish. A recent study was conducted with 82 Spanish-speaking volunteers from North Carolina (Burton, Vilar-Lopez, & Puente, 2012). Participants had emigrated from several countries in Central and South America, with 54 participants originating from Mexico. Data were collected from 28 private neuropsychological files, 28 murder defendant cases, and 25 personal injury, social security disability, or workers' compensation cases. Cutoffs were Dot Counting Test E-score < 17 , Rey 15 < 9 (Rey, 1964), and TOMM Trial 2 < 45 (Tombaugh, 1996). The TOMM and Rey 15 were able to differentiate groups (capital murder, other forensic or clinical control) from each other, whereas the Dot Counting Test was not. Interestingly, capital murder defendants performed similarly to clinical controls on all measures, whereas the other forensic group means for the Rey 15 and the TOMM both fell below

cutoffs. The PVT BR_{FAIL} of the other forensic group in this study ranged from 33% (TOMM) to 47% (Rey 15), which is similar to those reported in other studies. This study lends some evidence to the utility of PVTs with Spanish speaking examinees in the United States.

Vilar-Lopez et al. (2007) conducted a study that included 12 Spanish individuals who met postconcussion syndrome criteria and who were not involved in litigation, 14 Spanish individuals with postconcussion syndrome who were involved in postconcussion syndrome related litigation, and 25 analog university students who were coached to feign brain injury without being detected. The Victoria Symptom Validity Test (Slick, Hopp, Strauss, & Thompson, 1997) < 44, TOMM Trial 2 < 45, and b Test (Boone et al., 2002) e-score < 90 were used to differentiate between groups.

Although ANOVAs indicated that PVT scores did not differ between litigants and nonlitigants (whereas the analog group differed from both), the nonlitigant group performed close to the ceiling of all tests (i.e., clearly passed), whereas the litigant group mean scores were close to the cutoff points (passing TOMM and failing Victoria Symptom Validity Test and b Test). The authors concluded that whereas the nonlitigant group and analog group were homogenous (i.e., all nonlitigants had credible performance, and all analogue participants had noncredible performance), the litigant group showed a bimodal distribution, i.e., seven (50%) participants' performances were noncredible, and seven participants showed credible performance. The Spanish speaking participants' performance did not differ from North American norms provided with any of the included tests, lending further evidence to the utility of these PVTs for use with Spanish speakers.

Vilar-Lopez, Gomez-Rio, and Santiago-Romajo et al. (2008) sought to validate the TOMM and Dot Counting Test in a Spanish sample. The study included 54 Spanish mTBI patients who met criteria for postconcussion syndrome – at least three items with a rating ≥ 3 – on the Rivermead Post Concussion Symptoms Questionnaire (King, Crawford, Wenden, Moss, & Wade, 1995). CTs for all participants were normal. Thirty of the participants were classified as not compensation seeking. The second group were compensation seeking and passed cutoff criteria for Victoria Symptom Validity Test, b test, and Rey 15 (specific cutoff scores not reported; $n = 14$). The third group had evidence of noncredible performance (failing two of Victoria Symptom Validity Test, b Test, or Rey 15-Item Test) and were compensation seeking ($n = 10$). The final analog group was composed of 54 psychology students from the Universidad de Granada who were coached to fake impairment and avoid detection. TOMM Trial 2 < 45 showed perfect specificity and sensitivity. Dot Counting Test grouped item time $> 7s$ had 1.00 specificity and .30 sensitivity and Dot Counting Test errors > 3 displayed .85 specificity and .40 sensitivity. Dot Counting Test ratio < 1.5 showed .79 specificity with .80 sensitivity. Combination score for the Dot Counting Test ≥ 17 had 1.00 specificity and .40 sensitivity. The authors concluded that the TOMM and Dot Counting Test had comparable failure rates to those published in the TOMM and Dot Counting Test manuals in this experimental setting.

Another study by Vilar-Lopez, Gomez-Rio, and Llamas-Elvira et al. (2008) included 54 mTBI patients from Spain who met identical criteria to those previously outlined. This study sought to validate the use of the Victoria Symptom Validity Test, b test, and Rey 15 in a Spanish population. CTs for all participants were normal. Thirty of

the participants were not compensation seeking. The second group was compensation seeking with credible performance on the TOMM and Dot Counting Test (passing unspecified “U.S.” cutoff criteria on TOMM or Dot Counting Test; $n = 14$). The third group were classified as having noncredible performance and were compensation seeking ($n = 10$). A simulator group of 54 psychology students was coached to fake impairment while avoiding detection. Victoria Symptom Validity Test difficult items < 16 had 1.00 specificity and .63 sensitivity, whereas Victoria Symptom Validity Test easy items < 16 had 1.00 specificity with only .13 sensitivity. Victoria Symptom Validity Test total < 30 also had 1.00 specificity with .38 sensitivity. The b Test e-score > 90 and b Test d errors > 1 both showed .81 specificity and 1.00 sensitivity. The b Test commission errors > 3 showed 1.00 specificity and .38 sensitivity, and b Test omission errors > 50 had .81 specificity and .38 sensitivity. The b test time > 850 seconds had .81 specificity and .25 sensitivity. Rey 15 < 9 showed .82 specificity with .56 sensitivity. All tests were capable of differentiating groups, but the Rey 15-Item test did not perform as well as the Victoria Symptom Validity Test or b Test in classifying noncredible performers. Importantly, the noncompensation seekers and compensation seekers who displayed credible performance did not differ from each other in performance on the Victoria Symptom Validity Test, b Test, or Rey 15, suggesting that litigation does not completely explain the study findings. Overall, there is some promising evidence that PVTs may be effective with Spanish speakers in the United States and Spain.

PVTs in Asia and with Asian Americans. A series of studies were conducted in Hong Kong to develop a test battery to assess credibility (Chang, 2006). In the first study, 58 community participants were randomly assigned to exaggerate symptoms following a

hypothetical mTBI or to perform as well as possible. The Hong Kong List Learning Test (Chan & Kwok, 1999), a Cantonese language word list learning task similar to the California Verbal Learning Test was employed, as well as the Test of Nonverbal Intelligence – 3 (Brown, Sherbenou, & Johnson, 1997), the Cantonese Mini-Mental State Examination (Chiu, Lee, Chung, & Kwong, 1994), the Beck Depression Inventory (BDI; Beck, 1987), and the TOMM (Tombaugh, 1996). Total retrieval, recognition hits, false alarms, and difference of recall and recognition on the Hong Kong List Learning Test differentiated the two groups. All control participants scored either 49 or 50 on TOMM Trial 2, whereas the mean score for simulators was 28.4 ($SD = 13.6$).

In a second study in the same thesis, 20 patients with major depressive disorder were included and compared to the previously described true performance group and simulator group. The simulator group performed significantly more poorly on TOMM Trial 2 than both the credible control and major depressive disorder groups, and the control and major depressive disorder groups did not differ significantly from each other (major depressive disorder TOMM Trial 2 $M = 48.0$, $SD = 3.92$). Specificity at TOMM Trial 2 < 45 for the control group was 1.00 but was .80 for depressed patients. Using failure of any two PVTs resulted in .72 sensitivity to simulators but an unacceptable specificity of .75 for depressed patients. Use of any four or more failures as the failure criterion resulted in 1.00 specificity for depressed patients with .60 sensitivity to simulators. The findings of these studies suggest that PVTs are effective at differentiating simulators from controls in a Hong Kong sample, but that use of four or more PVT failures is necessary to differentiate simulators from patients with major depressive disorder in this region.

Another study was conducted with Hong Kong student simulators to compare the classification accuracy of a 48-item version of the Digit Memory Test (Hiscock & Hiscock, 1989) at two levels of difficulty (Chiu & Lee, 2002). All 38 participants participated in both the control and simulator groups using a Latin Square design. Using Digit Memory Test ≤ 29.7 (developed via a formula based on the curtailed items in the study), the authors found specificity of 1.00 for both easy and difficult versions of the task, and sensitivity of .30 for easy items and .76 for the difficult items. The authors concluded that their study provides preliminary evidence that Asian individuals perform in a similar pattern to Caucasian North American individuals on the task, but suggest more stringent cutoffs than those provided by the test publishers for easy items (i.e., cutoffs that will increase the sensitivity of the test).

A study by Yang et al. (2012) sought to validate the use of several Wechsler Adult Intelligence Scale-Third Edition (WAIS-III) Digit Span PVT scores in a sample in Taiwan. The authors compared the normative data available for the validation of the WAIS-III in the United States (Wechsler, 1997a) and the standardization sample for the Chinese version of the WAIS and Wechsler Memory Scale (WAIS-IIIC and WMS-IIIC; Hua et al., 2005). The Taiwanese sample consisted of 1,658 participants, and the American WAIS sample consisted of 2,450 participants. The Taiwanese standardization sample had significantly longer longest digits forward and significantly shorter longest digits backward than did the American standardization sample. The Taiwanese participants were also more likely to pass the Longest Digits Forward ≤ 4 cutoff, and less likely to pass the Longest Digits Backward ≤ 2 cutoff. The authors then examined Digit Span performance comparing 96 TBI patients ($n = 22$ in financial compensation

litigation, $n = 33$ nonlitigant mTBI) and 253 psychiatric outpatients ($n = 72$ schizophrenia spectrum disorders, $n = 22$ depressive disorders, $n = 16$ bipolar disorders) in Taipei. Eighteen of the psychiatric patients were in litigation for financial compensation. Participants were classified by presence or absence of litigation. The authors proposed a cutoff of ≤ 8 on Reliable Digit Span for Taiwanese populations to differentiate litigants from nonlitigants, noting that Vocabulary minus Digit Span did not provide good discrimination in the sample. This study provided some evidence that Reliable Digit Span is a useful PVT for examinees from Taiwan, and that Vocabulary minus Digit Span is a less useful PVT in this population.

Yamaguchi (2005) completed a simulation study with 52 adults in Kyoto to evaluate the effectiveness of the WAIS-R Digit Span and Rey-15 in detecting Japanese simulators. Groups included 15 normal controls, 17 participants instructed to simulate head trauma symptoms from a hypothetical motor vehicle accident, 12 healthy adults > 65 , and eight nursing home residents with dementia. The Rey-15 cutoff of ≤ 8 for items and ≤ 1 for columns was able to discriminate young controls from young simulators but did not discriminate simulators from older Japanese adults. Additionally, the nursing home residents with dementia were all classified as noncredible when the cutoff score was adjusted to ≤ 9 correct items. A score of three or fewer correct rows was the most effective cutoff when nursing home residents were dropped from the analysis. Regarding Digit Span, the cutoffs of ≤ 8 raw total, ≤ 5 or ≤ 4 forward, and ≤ 3 or ≤ 2 backward digits was adequate for the correct classification of normal young and older adult individuals from simulators. Nursing home residents performed similarly to simulators on Digit Span. Overall, this study showed that Digit Span scores are useful with younger

Japanese adults, but that clinicians should use caution when interpreting older adults' scores.

There is very little information about PVT use with South Asian examinees. One study used the Dot Counting Test and the Clinician-Administered Posttraumatic Stress Disorder Scale (Blake et al., 1995) with 105 Punjabi individuals who were engaged in a civil suit against the government of India for the killing and illegal cremation of family members by police between 1992 and 1993 (Weiss & Rosenfeld, 2010). Although this population is not directly relevant to the current study, the results are reviewed, as the available research about South Asian people and PVTs is quite limited. About half of participants were never formally educated. Four participants exceeded examiner judgment of symptom exaggeration (Clinician-Administered Posttraumatic Stress Disorder Scale cutoff > 2 ; $BR_{FAIL} = .038$). The mean Dot Counting Test E-score was 21.1 ($SD = 9.4$), a score which is significantly worse than any published group norm except for those with moderate dementia (Boone, Lu, & Herzberg, 2002). The authors reported that using the most liberal recommended cutoff, ≥ 14 for patients with depression, 78.1% of the sample would be classified as noncredible, and using the most conservative recommended cutoff, ≥ 22 for patients with mild dementia, 40.0% of the sample would be classified as noncredible.

The design of this study had several limitations. The first of the limitations was using a PVT that required counting with participants who had very little education. The use of cutoffs that were developed with a better-educated population may have inflated PVT BR_{FAIL} in the sample. Secondly, the authors were attempting to use the Dot Counting Test as a measure of symptom exaggeration, rather than as a PVT. Measuring

symptom exaggeration is distinct from noncredible neuropsychological performance, each measuring different constructs that may or may not overlap in an individual (Larrabee, 2012). What is evident from this research is that a large proportion of the participants in rural India fail the Dot Counting Test with current American norms. Further research is needed in the applicability of PVTs in South Asian examinees, both in South Asia and in North America.

Cirlugea (2014) reviewed prior research that included PVTs with Asian participants in North America and Asia. Aside from the above-noted research conducted by Chang (2006), none of the North American research on the TOMM that included any Asian participants analyzed the performance of Asian participants separately, and most included only one or two Asian participants. The only exception about sample size was a study of the effects of different styles of coaching on PVT performance (Weinborn, Woods, Nulsen, & Leighton, 2012). This study included 42 Asian participants ($N = 103$) but did not analyze the performance of Asian participants separately from Caucasian participants. Cirlugea (2014) concluded that there was minimal research on the use of the TOMM with Asian clients, and further reviewed studies that include the Medical Symptom Validity Test (Green, 2004) and Nonverbal Medical Symptom Validity Test (Green, 2007a) with some Asian participants.

The author noted that none of the research has analyzed Asian participants specifically, but mentions one study (Armistead-Jehle, 2010) of 45 veterans with TBI that found no differences in Medical Symptom Validity Test performance between ethnic groups (6.7% Asian). Cirlugea (2014) briefly reviewed other publications that included Asian clients and the Digit Memory Test, concluding that further study is necessary,

especially considering that extant research was conducted in China, and may not generalize to other Asian people or Chinese people living in the United States (or, by extension, Canada).

Overall, there is some preliminary evidence for the utility of the use of the TOMM with Hong Kong residents, various DS scores with Taiwanese residents, and Rey-15 and Reliable Digit Span with Japanese residents. Research suggests that current Dot Counting Test cutoff scores are inappropriate for rural Punjabi residents. There is, however, very little information about the validity of the use of these tests with any Asian populations in North America or the use of PVTs with examinees from other East, South, Southeast or West Asian nations. Further, there is very little information about the use of PVTs with North American residents of Asian descent who have sustained a TBI or are in litigation, as the previous studies have largely used simulation designs.

There is a great need for further research into the generalizability of PVTs to people from diverse cultural and linguistic backgrounds. Excluding PVTs—a core component of neuropsychological assessment—threatens the validity of the assessment. The use of PVTs with examinees from diverse cultures is poorly understood, though, posing threats to the validity of the measures. Denying services to examinees from diverse cultures is also unfeasible and unethical. The need for further validation of PVTs with culturally diverse examinees is clear.

Other Demographic Variables and PVTs

Less directly culturally-related demographic variables may also affect PVT performance and are therefore germane to the discussion of PVT interpretation. Demographic variables are not typically the primary focus of studies on PVTs, which

limits the availability of information about them. This section reviews the extant literature about the contribution of education, age, and gender to PVT performance.

Lower education is usually found to be related to higher BR_{FAIL} in studies. For example, Prieto de Estebecorena (2007) found high correlations between education and TOMM performance ($r = .51, .53, \text{ and } .51$ for Trials 1, 2, and Retention, respectively) in a community sample of 120 Hispanic individuals living in San Francisco. Stulemeijer et al. (2007) showed that lower education was associated with higher BR_{FAIL} on the Amsterdam Short Term Memory Test (cutoff < 86 ; Schagen, Schmand, de Sterke, & Lindeboom, 1997) in a sample of 118 mTBI referrals to an emergency department in the Netherlands.

In some studies, however, education has not been related to BR_{FAIL} . For example, Gervais et al. (2004) found no educational differences on the Word Memory Test < 86 (Green et al., 2003), TOMM < 45 on Trial 2 or Retention, or Computerized Assessment of Response Bias total score (Allen, Conder, Green, & Cox, 1997) in a sample of 519 pain patients (Gervais, Rohling, Green, & Ford, 2004). No educational differences in scores were reported in the TOMM test manual (Tombaugh, 1996). Overall, however, the preponderance of evidence indicates higher BR_{FAIL} for examinees with lower educational attainment.

Findings about the relationship between age and PVT failure in adults is inconsistent (Strauss et al., 2006; Webb et al., 2012), although there is a growing body of evidence that younger age is associated with higher BR_{FAIL} in children and adolescents (e.g., Brooks & Ploetz, 2015; Lichtenstein, Erorid, Rai, Mazur-Mosiewicz, & Flaro, 2016). Some studies (e.g., Lange, Iverson, Brooks, & Rennison, 2010; Stulemeijer et al.,

2007) showed no significant age differences in BR_{FAIL} . Others showed that older individuals have higher BR_{FAIL} than younger examinees (Donders & Boonstra, 2007; Grote et al., 2000; Webb et al., 2012). For example, Grote et al. (2000) compared 30 noncompensation-seeking and 53 compensation-seeking examinees in a Victoria Symptom Validity Test validation study. The compensation seeking group was significantly older. Older examinees had fewer correct responses and longer response times for easy and difficult items. When compensation and noncompensation groups were analyzed separately, there were no age-related differences. Thus, it appears that compensation-seeking status was a more important predictor than age in this sample.

Donders and Boonstra (2007) investigated correlates of PVT failure in a sample of 87 participants with TBI, using California Verbal Learning Test Second Edition Forced Choice Recognition ≤ 14 , and the published cutoffs for the Word Memory Test ($\leq 83\%$ on Immediate Recognition, Delayed Recognition, or Consistency; Green et al., 2003). They found that older age was associated with greater rates of noncredible performance.

Babikian et al. (2006) found some differences for both education and age on Digit Span embedded validity indicators in a sample with a control group of 32 healthy women, a nonlitigant mixed clinical group of 56 participants, and a suspect effort group of 66 examinees with noncredible performance. Noncredible performance was defined as: Dot Counting Test ≥ 17 ; Rey Word Recognition ≤ 6 or \leq Rey Auditory Verbal Learning Test Trial 1 (Lezak, 1983); b test > 2 commission errors, > 0 “d” commission errors, > 40 omission errors, > 12 minutes completion time; Warrington Recognition Memory Test-Words < 33 ; Rey 15 < 9 or Rey 15 + Recognition Combination Score < 20 ;

and Rey Auditory Verbal Learning Test 30 minute recognition trial ≤ 7 or ≤ 30 -minute free recall.

In the clinical group, younger age correlated with better Longest Digits Forward scores. There were no age-related differences for Digit Span embedded validity indicators in the suspect or control groups. More education was related to higher age-corrected scaled score, Reliable Digit Span and Longest Digits Forward in the noncredible and control groups, but not in the clinical group. The authors offered no interpretation of these findings, which were not the primary focus of the study.

Webb et al. (2012) developed a model to predict PVT failure that included both age and education in 555 private practice patients with TBI in New Zealand. PVT failure was defined as below chance performance ($<18/50$) on TOMM Trial 2 or Retention or failure of two or more of TOMM Trial 2 < 47 , Reliable Digit Span < 8 , or Rey-15 < 9 . PVT failure was related to milder TBI severity, less education, older age, having immigrated to New Zealand, having a workplace accident, compensation-seeking, self-reported diagnosis of mood or psychotic disorders, and exhibiting florid behaviours during the examination (e.g., lying on the floor and complaining of fatigue following interview). In a logistic regression model, self-reported mood or psychotic disorder, florid behaviours, compensation seeking, having immigrated to New Zealand, and lower education remained significant predictors of PVT failure, but age was not a significant predictor in the model.

Whenever reported, gender has not been associated with PVT BR_{FAIL} (Constantinou & McCaffrey, 2003; Donders, 2005; Rees et al., 1998; Webb et al., 2012). Research typically does not focus on gender differences, which are rarely reported.

Overall, research generally indicates that lower education is related to BR_{FAIL} across several PVTs. The evidence for an association between age and BR_{FAIL} is mixed, and there is no association between gender and BR_{FAIL} . These characteristics are of interest in the current research, as they are fundamental demographic characteristics that differentiate examinees and may be associated with or even partially explain PVT failure.

Ontario Ethnic Demographics

Given that cultural and linguistic factors are important in the accurate interpretation of PVT results, the diversity in Ontario's population and lack of multicultural norms may result in inaccurately designating examinees from minority groups as noncredible in this province. Although the paucity of information is not limited to Ontario, the current study will address these limitations with a sample of Ontarian motor vehicle accident litigants.

The demographic composition of Ontario renders it an emblematic location to explore cultural common factors that affect neuropsychological test results. The Ontario Ministry of Finance (Ministry of Finance, 2014) reports that in the 2011 Household Survey 25.9% of individuals in Ontario identified themselves as being a member of a visible minority. This segment of the population is growing nearly five times faster than the general population. Notably, Ontario First Nations people were not considered members of a visible minority in this census.

Sixty-nine percent of self-identified visible minority Ontarians were born outside of Canada. South Asian people make up 29.5% of visible minority individuals in Ontario; 19.2% are Chinese, 16.4% identify as Black, 8.4% are Filipino, 5.3% are Latin American, 4.6% identify as Arab, 4.2% identify as Southeast Asian, 3.7% are West Asian, 2.4% are

Korean, and 0.9% are Japanese. People who identify themselves as belonging to multiple visible minorities account for 2.9% of members of a visible minority, and people who identify as a visible minority not otherwise identified account for 2.5% of visible minority individuals in Ontario. This demographic picture differs markedly from that of the United States (United States Census Bureau, 2010), where Nonhispanic White people account for 63.7% of the population, Black or African American people account for 13.2%, Hispanic or Latino people account for 16.3% and other groups account for the remaining 6.8% of the population.

Given the greater ethnic diversity of Ontarians, and the fact that most multicultural neuropsychological assessment research in North America was completed with Hispanic Americans and African Americans, neuropsychologists practicing in Ontario do not have adequate measures or norms to assess examinees in their practices.

Conceptualization of the Effect of Cultural, Linguistic, and Demographic Factors on PVTs in the Current Study

The first objective in the present research addressed the cultural diversity of Ontarians and the limitations of previous multicultural PVT research by identifying common factors across cultural groups that may affect PVT performance. These common factors are necessarily demographic due to the constraints of the secondary data used in this study. In other words, it was impossible to explore the effects of more nuanced cultural variables such as acculturation or acculturative stress due to the lack of pertinent measures in the data set.

Consistent with the American Academy of Clinical Neuropsychology guidelines on multiculturally competent assessment, it was expected that cultural, demographic, and

linguistic factors might affect PVT performance, leading to inaccurate and even prejudicial classification of individuals from minority cultural groups as noncredible (Board of Directors, 2007). This outcome would be stigmatizing, lead to denial of services, and may alienate the examinee from the mental health system (Paniagua, 2005). It was predicted that BR_{FAIL} on PVTs with high verbal mediation would be affected by these common factors, whereas PVTs with low verbal mediation would not be affected, as they may be less culturally biased (Boone et al., 2007).

Limited English proficiency is a factor shown in previous literature to affect performance on cognitive tests (Boone et al., 2007), yet it is poorly defined in neuropsychological assessment and research and is often based on subjective examiner report (Erdodi, Jongsma, & Issa, 2017). Standardized measures of English language proficiency such as the Test of English as a Foreign Language (Education Testing Service, 2016) exist, but are extensive and not well suited for neuropsychological testing. Using available direct tests of language included in the battery (e.g., Wide Range Achievement Test 4 Word Reading Subtest; Wilkinson & Robertson, 2006) as the basis of limited English proficiency designation might have provided an objective index of English proficiency in the current study. Their use, however, rests on the assumption of credible performance, which is questionable in a sample of compensation-seeking litigants. Instead, native-level English language proficiency was defined as having English as a first language. Limited English proficiency status was defined as having English as a second language and having immigrated to Canada after age 17. Intermediate groups whose first language is not English and who immigrated to Canada as children (age ≤ 9 years), or adolescents (age 10 to 17 years), would also be analyzed.

These classifications were chosen to be consistent with previous research on language acquisition (Archila-Suerte, Zevin, & Hernandez, 2015; Dekeyser et al., 2010).

Longer time in Canada was also expected to be related to better PVT performance. Time in Canada was divided according to Canadian immigration policy and previous health and immigration research (Citizenship Act, 1985; Vang, Sigouin, Flenon, & Gagnon, 2015): Canadian-born, people who have lived in Canada for four or fewer years, five to nine years, and 10 years or more.

Other factors, including TBI severity, education, age, and gender were also explored. Consistent with previous research, it was expected that examinees with lower education would have higher PVT BR_{FAIL} . It was also expected that examinees with mTBI would have higher PVT BR_{FAIL} than those with more severe injuries. The direction of the relationship between gender and PVT performance and age and PVT performance was not predicted due to null or equivocal prior research findings.

CHAPTER 3

Psychiatric Symptoms and Performance Validity

Another consideration in the accurate interpretation of PVT BR_{FAIL} is the impact of psychiatric factors on performance, which was the focus of the second objective of the current research. Research on the association between various psychiatric disorders and PVT BR_{FAIL} have shown mixed results. Most research on PVT failure following mTBI or motor vehicle accident has examined their association with concurrent self-reported PTSD and depression. The link between dissociative symptoms and PVT BR_{FAIL} has never been tested. The following sections review research into the relationship between psychiatric symptoms and performance validity.

Depression, Anxiety, and PTSD and Performance Validity

Psychiatric disorders and mTBI. Several researchers have explored the presence of psychiatric disorders following TBI and motor vehicle accidents. Moore, Terryberry, and Hope (2006) reviewed previous literature examining anxiety following mTBI. The authors reported that anxiety disorders are common following TBI (23% to 29%). These disorders included generalized anxiety disorder (2% to 28%), panic disorder (4% to 17%) and PTSD (3% to 27%). The anxiety disorders often predated the injury.

The authors reported that evidence has been inconsistent regarding increased anxiety symptoms in examinees with mTBI compared to matched control participants without mTBI. Research into obsessive-compulsive disorder and panic disorder following mTBI was sparse and inconsistent, and research into generalized anxiety disorder following mTBI was sparse but indicated twice the rate of generalized anxiety disorder following mTBI compared to the general population. PTSD was the most thoroughly

researched disorder, yet the authors found inconsistencies in the literature about the association between mTBI and PTSD. They also commented on the controversy over the possibility of developing PTSD in the presence of loss of consciousness (i.e., when there is no memory of the traumatic event in which the head injury occurred).

In another literature review, Hesdorffer, Rauch, and Tamminga (2009) found that TBI was consistently associated with the development of depression, even in the absence of a prior history of depression. The risk of developing a depressive episode was higher in those with TBI and a prior history of depression. TBI was also associated with the development of anxiety, and anxiety was more common following mTBI compared to orthopaedic injuries. Panic disorder was also more common following TBI than in the general population in the literature review.

In contrast, PTSD was not associated with mTBI following motor vehicle accident when compared to individuals who had been in a motor vehicle accident without head injury. mTBI was also not associated with PTSD in participants recruited from emergency departments. However, participants with comorbid mTBI and PTSD reported greater postconcussion syndrome symptoms than individuals with mTBI who did not have PTSD. Individuals with PTSD also had higher postconcussion syndrome symptoms compared to no-PTSD controls.

Overall, previous research suggests that mTBI is associated with higher rates of depression, whereas evidence for a link between mTBI and various anxiety disorders and PTSD is inconsistent. The causal link between mTBI and psychiatric disorders is also unclear.

Psychiatric disorders and neuropsychological functioning.

Neuropsychological functioning can be negatively affected by mental health disorders that are common following motor vehicle accident. Scott et al. (2015) conducted a meta-analysis of neurocognitive functioning in PTSD that included 60 studies and 4,108 participants. The authors found that PTSD was associated with moderate deficits in verbal learning and memory, working memory and processing speed, and small deficits in executive functions, language, visual learning, memory, and visuospatial abilities. Being involved in treatment was associated with greater neurocognitive deficits, which the authors posited may be due to greater symptom severity in those who seek treatment. Excluding participants with TBI did not alter the effect sizes, indicating that comorbid TBI was not the reason for neuropsychological deficits in the sample with PTSD.

Castaneda et al. (2008) conducted a review of cognitive impairment in young adults with major depressive disorder and anxiety disorders. The review included nine studies with major depressive disorder, two with panic disorder, 15 with obsessive-compulsive disorder, two with generalized anxiety disorder, and five with PTSD. The authors found that major depressive disorder was associated with deficits in executive function; attention; verbal, visual short-term and working memory; and psychomotor tasks. The authors found inconsistent evidence of visual memory and learning deficits for examinees with panic disorder, but some evidence of impairments in short-term and long-term verbal memory, executive functioning, and concentration for those with the disorder. The authors further found some evidence of deficits in attention, executive and visuospatial function, and short-term verbal memory and learning in examinees with social phobia. No cognitive deficits were associated with generalized anxiety disorder.

Young adults with PTSD showed deficits in attention, short-term and long-term verbal and visual memory, and executive functioning. The much larger body of literature on neuropsychological functioning in examinees with obsessive-compulsive disorder indicated deficits in executive functioning, visual memory, attention and processing speed in examinees with that disorder. Overall, there is evidence that anxiety disorders and PTSD are associated with some cognitive impairment.

McClintock et al. (2010) conducted a review of 35 studies exploring the association between depression severity and neurocognitive function. The authors found that depression was associated with deficits in attention, learning, memory, and executive function, with increased symptom severity at examination associated with worse neurocognitive dysfunction. Recurrent depression was also associated with more severe deficits when compared to single episodes of depression. The authors noted highly variable research designs and definitions for both cognitive functioning and depression severity that contribute to inconsistent findings across studies.

Rock et al. (2014) conducted a meta-analysis of cognitive deficits in individuals with major depressive disorder during depressive episodes and remission. The meta-analysis included 24 studies with 784 participants who were in a current episode compared to 727 control participants and 168 participants who were in remission compared to 178 control participants. The authors found that participants who were currently depressed had moderate deficits in executive function, memory, and attention. Participants in remission had moderate deficits in executive function and attention, and small-to-moderate deficits in memory. This meta-analysis indicated that the effects of neurocognitive effects of depression persist even during periods of remission.

Overall, a large body of evidence suggests that PTSD is associated with neuropsychological deficits, and evidence suggests that major depressive disorder is also associated with neuropsychological deficits. Limited evidence for anxiety disorders other than obsessive-compulsive disorder suggests that they may also be associated with neuropsychological deficits.

Psychiatric disorders and PVT performance. Despite their association with decreased neuropsychological test results, depression and anxiety are generally not associated with worse PVT performance. Ashendorf, Constantinou, and McCaffrey (2004), for example, explored TOMM score differences based on self-reported state and trait anxiety and depressive symptoms in a sample of 197 adults between ages 55 and 75. They found no differences in TOMM scores between groups with high or low depression or state or trait anxiety. Considine et al. (2011) also did not find differences between 45 patients with major depressive disorder and 32 healthy control participants on the TOMM. Yanez et al. (2006) likewise found no difference in TOMM scores between a group of 20 participants with severe major depressive disorder who were assessed for Social Security Disability and a control group of 20 nondepressed family members of the major depressive disorder group. O'Bryant, Finlay, and O'Jile (2007) found that self-reported depression and anxiety were not related to TOMM scores in a sample of 67 patients referred for outpatient neuropsychological assessment.

Schroeder and Marshall (2011) evaluated PVT BR_{FAIL} in 104 patients with psychosis and 178 patients with nonpsychotic disorders (91.5% major depressive disorder, 3.4% generalized anxiety disorder, 2.2% PTSD, 0.6% adjustment disorder, 1.1% impulse control disorder, 0.6% obsessive-compulsive disorder, and 0.6% Social

Anxiety Disorder). BR_{FAIL} for the nonpsychotic group were: Reliable Digit Span (≤ 7) 22%; Reliable Digit Span (≤ 6) 4%; California Verbal Learning Test Second Edition Forced Choice Recognition (≤ 14) 2%; Logical Memory Rarely Missed Items (≤ 136) 8%; Finger Tapping Test (≤ 35 for males, ≤ 28 for females) 6%; and Rey Complex Figure Test (Recognition True Positive ≤ 3 or False Positive > 4) 1%. The authors concluded that PVT failure was relatively rare in examinees with psychiatric illness and that even in the case of psychotic disorders, only 7% of examinees failed more than one PVT, indicating that PVTs are robust to psychiatric illnesses.

Most research on PTSD and PVTs focuses on identifying noncredible PTSD, rather than examining the effects of credible PTSD on PVT performance as is the case with major depressive disorder (e.g., Rubenzer, 2009; Young, 2015a). Merten et al. (2009), for example, used Reliable Digit Span and Word Memory Test to determine credibility in 77 examinees with self-reported PTSD of various origin in Germany. Causes included motor vehicle accident ($n = 35$), industrial accident ($n = 28$), assault and robbery ($n = 7$), witnessing death or serious illness ($n = 3$), other violence ($n = 2$), witnessing violence against another person ($n = 1$), and medical malpractice ($n = 1$). Eleven participants also reported having mTBI. PVT BR_{FAIL} was high, with 23% of participants failing Reliable Digit Span ≤ 7 and 51% of participants failing Word Memory Test ($\leq 83\%$ on Immediate Recognition, Delayed Recognition, or Consistency). The authors concluded that noncredible performance is common in people who are seeking compensation for PTSD.

Wisdom et al. (2014) examined the cognitive performance of 166 American veterans with mTBI. Examinees were categorized as controls if they had no self-reported

PTSD ($n=36$), PTSD-pass if they self-reported PTSD and passed the Word Memory Test at published cutoffs ($n=30$), and PTSD-fail if they self-reported PTSD and failed the Word Memory Test ($n=68$). The authors found that the PTSD-pass and control groups did not differ on any neuropsychological tests, but that the PTSD-fail group had lower scores than the control group on 14 of 19 tests. The authors concluded that this lent support to the hypothesis that previously documented cognitive deficits in people with PTSD might be at least partially attributable to failure to control for PVT performance.

Recent research has challenged the previous assumption that non-psychotic psychiatric symptoms are unrelated to PVTs. While previous studies generally included patients with a single psychiatric diagnosis such as major depressive disorder, new studies have explored the performance of more heterogeneous samples with psychiatric comorbidities, histories of trauma, and cognitive complaints in the absence of evidence of a neurological cause. The data for newer studies were extracted from consecutive neuropsychological referrals to medical centres, suggesting more generalizable samples when compared to older research that used strict exclusion criteria. Older, restrictive exclusion criteria may have led to higher internal validity at the expense of external validity. The recent studies are more representative of clients who are referred for clinical neuropsychological assessment, and results of these studies reveal a relationship between PVT performance and self-reported psychiatric symptoms. Erdodi, Tyson, et al. (2016) conducted a study that included a sample of 106 patients referred for neuropsychological assessment for epilepsy, postconcussive disorder, psychogenic nonepileptic seizure, or cognitive deficits due to psychiatric pathology. Thirty-six of the patients had self-reported cognitive deficits that were judged to result from emotional distress. The authors found a

60.7% BR_{FAIL} on EI-5 ≥ 4 (composite measure of reliable Digit Span, Digit Span age-corrected scaled score, Logical Memory recognition, California Verbal Learning Test Second Edition recognition hits, and California Verbal Learning Test Second Edition Forced Choice Recognition), and 73.3% on TOMM (Trial 1 ≤ 39 or Trial 2 ≤ 48) for patients whose deficits were psychiatric. These BR_{FAILS} were far higher than is typical for psychiatric populations. The authors suggested that their results may diverge from previous research because the psychiatric group in this study may have been experiencing more severe mental illness, complex emotional trauma, or somatic complaints than previous studies that focused on participants whose primary concern was major depressive disorder.

Erdodi, Seke, et al. (2017) examined the relationship between Grooved Pegboard performance, established PVTs, and self-reported psychiatric symptoms in a sample of 190 examinees referred for neuropsychological assessment with an established neurological or psychiatric diagnosis based on previous medical records. Examinees who failed the dominant hand Grooved Pegboard EVI (T -score ≤ 29) had higher self-reported scores on the Beck Depression Inventory (BDI-II), Beck Anxiety Inventory (BAI), and Personality Assessment Inventory (PAI) Somatic Complaints, Borderline Features, Antisocial Features, Alcohol, and Drug Problems subscales. Examinees who failed the non-dominant hand Grooved Pegboard EVI failure (T -score ≤ 29) also had higher BAI, Antisocial Features, Alcohol, and Drug Problems scores. The authors posited that examinees with noncredible presentations may be more likely to overreport symptoms on face-valid self-report measures such as the BDI-II and BAI as compared to less transparent measures such as the PAI. The authors proposed a psychogenic interference

hypothesis to explain the findings, whereby emotional distress interferes with test performance, leading to atypical score patterns—such as internal inconsistencies and PVT failure—which do not correspond to physiological patterns of deficits.

Overall, despite evidence that depression and anxiety are associated with cognitive deficits, these psychiatric disorders are not usually associated with elevated PVT BR_{FAIL} , which is likely attributable to the nature of PVTs, which are designed to capture only the tail of the neuropsychological performance distribution where score credibility is questionable. In other words, mild-to-moderate cognitive deficits associated with emotional distress should not be—and are not typically—detected by PVT failure. Recent research suggests that emotional distress may be related to PVT failure in complex cases with more severe mental illness, developmental trauma, and somatic symptom presentations. PVTs can detect noncredible PTSD, and PVT failure rates may be high in circumstances where examinees are seeking compensation for PTSD related disability (e.g., veterans; Young, 2015c).

Dissociative Symptoms and Disorders

Links between dissociation, neuropsychological function, and PVTs have not received as much attention as other psychiatric disorders. Dissociative disorders involve disrupted integration of awareness, emotion, perception, action, and memory (American Psychiatric Association, 2013). Dissociative symptoms are common in several diagnostic groups, including schizophrenia, borderline personality disorder, major depressive disorder, attention deficit hyperactivity disorder, obsessive-compulsive disorder, anxiety disorders, eating disorders, somatic symptom disorders, substance use disorders, and bipolar disorders (Soffer-Dudek, 2014). Dissociative symptoms also predict the

development of PTSD (Murray, Ehlers, & Mayou, 2002), and the presence of prominent dissociative symptoms constitute a subtype of PTSD (American Psychiatric Association, 2013). Peritraumatic dissociative symptoms are a risk factor for developing PTSD (Lensvelt-Mulders et al., 2008), and predict the development of PTSD following motor vehicle accidents (Berna, Vaiva, Ducrocq, Duhem, & Nandrino, 2012; Naim et al., 2014). Despite the prevalence of dissociative symptoms and their effects on cognitive functioning, very little research has explored the relationship between dissociative symptoms and mTBI following motor vehicle accident.

Dissociation is often portrayed as a defense mechanism to protect the self from aversive events and is conceptualized as originating in severe psychological trauma (Giesbrecht, Lynn, Lilienfeld, & Merckelbach, 2008). Although some researchers argue that dissociative symptoms are quite common (Soffer-Dudek, 2014) and are relevant to the development of PTSD following motor vehicle accidents, the causal relationship between trauma and dissociative symptoms has been questioned (Giesbrech et al., 2008; Merckelbach, Horselenberg, & Schmidt, 2002).

Dissociation and symptom validity. According to Merten and Merckelbach (2013), although psychologists may explain poor symptom validity test and PVT performance as resulting from clients' reported dissociative symptoms, this approach is logically flawed. Specifically, the clinician assumes an explanation for any symptom validity test or PVT outcome as resulting from the dissociative symptoms: either the examinee passed because they are honest, or they failed because of their symptoms. Either way, the antecedent self-reported dissociative symptoms are affirmed.

More recently, the same group of researchers (Merckelbach et al., 2015) explored the relationship between dissociative symptoms and symptom validity test failure in 269 undergraduate students and 22 psychiatric trauma inpatients in the Netherlands. The inpatients were diagnosed with PTSD ($n = 10$), Dissociative Disorders ($n = 7$), Mood Disorders ($n = 7$), and/or Borderline Personality Disorder ($n = 5$). The participants completed the Dissociative Experiences Scale (Bernstein & Putnam, 1986), the Cambridge Depersonalisation Scale (Sierra & Berrios, 2000), the Symptom Over-reporting Index (Merckelbach, Langeland, de Vries, & Draijer, 2014), and the Structured Inventory of Malingered Symptomatology (Smith & Burger, 1997). Dissociative symptoms were related to symptom overreporting in the student sample, but not in the inpatient sample. The authors recommended that researchers be cautious about student self-report while concluding that their study lent evidence to credible reporting of dissociative symptoms for psychological trauma patients.

Dissociation and neuropsychological functioning. To this author's knowledge, the relationship between PVTs and dissociative symptoms has yet to be examined. There is some information, however, about neuropsychological functioning in individuals with dissociative symptoms. Haaland and Landro (2009) conducted a study comparing the neuropsychological functioning of 30 healthy controls, 10 individuals with Borderline Personality Disorder and dissociative symptoms, and 20 individuals with Borderline Personality Disorder without dissociative symptoms. Individuals with high dissociation performed significantly worse than healthy controls on all domains, including attention, working memory, executive function, verbal long-term memory, nonverbal long-term memory, and general cognitive functioning. They also performed worse than individuals

with Borderline Personality Disorder without dissociative symptoms on working memory, executive functioning, verbal long-term memory, and general cognitive functioning tasks. Individuals with Borderline Personality Disorder without dissociative symptoms scored worse than healthy controls only on executive functioning (Haaland & Landro, 2009).

Parlar et al. (2016) conducted a study of the neuropsychological performance of 23 participants with major depressive disorder and 20 healthy controls. In participants with major depressive disorder, dissociative symptoms were correlated with worse performance on verbal and visuospatial memory, processing speed, and sustained attention. In contrast, depressive symptom severity was unrelated to neuropsychological performance. These results provide preliminary evidence that dissociative symptomatology has a negative impact on neuropsychological functioning.

Neuroanatomical correlates of dissociation. Research is also sparse regarding the neuroanatomy of dissociation. A recent study compared fMRI results between 36 participants with nondissociative PTSD, 13 participants with Dissociative Subtype PTSD, and 40 healthy control participants (Nicholson et al., 2015). When comparing the dissociative and nondissociative PTSD groups, the dissociative group had greater connectivity between the amygdala and multiple areas of the brain, including the superior parietal lobe, culmen of the cerebellum, posterior cingulate, precuneus, and medial frontal gyrus. There was no evidence of greater connectivity between any brain regions in the nondissociative group compared to the dissociative group. The authors interpreted these findings to be consistent with the hypothesis that individuals with dissociative

subtype PTSD have increased connectivity between areas involved in emotion regulation and consciousness.

These connections were implicated in earlier research in the downregulation of emotion by the prefrontal cortex, resulting in depersonalization and derealization symptoms (Lanius et al., 2010). The dissociative symptoms, in turn, affect cognitive functioning (Haaland & Landro, 2009; Parlar et al., 2016), and may affect PVT BR_{FAIL} . Dissociative symptoms are particularly related to poor frontotemporal functions including attention (Parlar et al., 2016), which has been previously implicated in higher PVT BR_{FAIL} in patients with psychosis (Hunt, Root, & Bascetta, 2014).

In sum, despite the lack of research into the association between dissociative symptoms and PVT performance, there is some evidence that dissociative symptoms negatively affect neuropsychological performance more generally. Dissociative symptoms, which involve alterations in consciousness, might be expected to interfere with PVT performance. PVT failure resulting from dissociative pathology might be attributed to noncredible performance. It is important to explore this avenue of research to reduce possibly inaccurate designation of noncredible performance in people with dissociative pathology.

Conceptualization of Psychiatric Symptoms in the Current Study

The current study explored the relationship between PVT BR_{FAIL} and self-reported depression, anxiety, and PTSD symptoms. Given mixed findings in the previous research regarding self-reported mood and anxiety symptoms and PVT performance, the relationships between self-reported symptoms PVT BR_{FAIL} were explored as research questions.

The current study also addressed the gap in the literature on the effects of dissociative symptoms on PVT BR_{FAIL}. Dissociative symptoms were suspected to be associated with higher PVT BR_{FAIL}. These PVT failures may lead to inaccurate noncredible performance designation, which would result in denial of access to psychological services. Examinees with active dissociation during testing may be experiencing an alteration in normal cognitive capacity that interferes with their ability to pass PVTs – for reasons that are fundamentally different from poor effort or outright malingering.

It was important, however, to control for noncredible self-reported psychiatric symptomatology. People who seek compensation following mTBI often engage in symptom over-reporting (Greiffenstein & Baker, 2008). In those cases, self-reported symptomatology provides an inaccurate estimate of the examinee's level of psychiatric pathology. To account for the effects of noncredible symptom reporting, individuals who failed symptom validity tests were categorized as a separate group in analyses of the relationship between psychiatric symptoms and PVT performance.

Findings that dissociative symptoms are related to higher PVT failure rates may indicate the need to develop adjusted cutoff scores for examinees with high dissociative symptoms to control for false positive errors, in line with previous research with other diagnoses that affect PVT performance, such as schizophrenia and dementia (Goldberg et al., 2007). It could also provide a novel pathway to cognitive rehabilitation through the treatment of psychiatric symptoms.

Possible interactions of dissociation and cultural factors were not examined in the current research. Dissociation has been conceptualized as a reaction to extreme

psychological stress (Giesbrecht et al., 2008). It has been explored in qualitative analyses with victims of extreme trauma such as the Rwandan genocide, which are rarely experienced by the Canadian majority group (Sandole & Auerbach, 2013). To this author's knowledge, however, there is no information about differences in the experience or expression of dissociation across cultures, which may confound the interpretation of results in the current study. The scale used in the current study may not accurately capture dissociation expression across cultures. Interactions between cultural variables and self-reported dissociation in this data set may then be due to limited cross-cultural construct validity rather than true differences in dissociative symptoms across cultures.

Additionally, the secondary data set that was used for the current study did not reliably include information about experiences of extreme stress or child abuse, which made controlling for experiences of extreme stress unfeasible. Using past experiences of extreme stress and child abuse to examine the convergent validity of the dissociation measure across cultures was also not possible.

Additionally, both the exploration of cultural common factors in the first study objective, and dissociation in the second study objective, are the first examinations of the association of these variables to PVT failure. Foundational information about the association of these constructs with PVT failure is necessary before further exploration into potential interactive effects of dissociation and common cultural factors on PVT failure is warranted.

CHAPTER 4

General Methods

Objectives 1 and 2 used the same archival data set. Participants were examinees who were involved in a motor vehicle accident and were assessed by a clinical neuropsychologist in Ontario between January 01, 2013 and August 15, 2015. The assessments were completed as part of independent medical examinations on behalf of the examinee's auto insurer to provide recommendations for accident benefits. The data consisted of the neuropsychological reports, as well as neuropsychological test data and self-report measures. This research received approval from the University of Windsor Research Ethics Board on June 08, 2015.

Participants

The study included a sample of 325 adults who underwent neuropsychological assessments while seeking compensation following motor vehicle accidents in Southern Ontario.

Descriptive statistics. Demographic and injury characteristics of the sample are displayed in Table 2, indicating that the majority of examinees had an mTBI, were Canadian born, and were right-handed. The time between the motor vehicle accident and the assessment had a wide range (two months to over 16 years), as did time living in Canada for immigrant examinees (zero to 57 years), and the age at which immigrant examinees moved to Canada (one year to 53 years). The proportion of males and females was roughly equal.

Examinees were born in 47 different countries including Canada. The most common countries of origin were Jamaica ($n = 8$), Afghanistan ($n = 7$), Sri Lanka ($n = 7$),

Poland ($n = 6$), China ($n = 5$), Philippines ($n = 5$), and Guyana ($n = 4$). Three or fewer examinees immigrated to Canada from any other country. The small sample size from each non-Canadian country precluded country-based statistical analyses.

The diversity in demographic variables reinforces concerns raised by Elbulok-Charcape et al. (2014; see pp. 41-43 for a more thorough discussion). Namely, despite the imperative to provide a culturally nuanced assessment to examinees, neuropsychologists do not have adequate tests, norms, or culturally competent colleagues to adequately assess the diverse examinees they encounter. Furthermore, this demographic picture supports the assertion that developing tests and norms that are appropriate for use with the largest North American linguistic minorities (i.e., Spanish and French speakers) would not fully address the broad diversity of examinees referred to neuropsychologists.

Table 2

Means, Standard Deviation, and Range of Demographic Characteristics

	<i>N</i>	<i>M</i>	<i>SD</i>	Range
Age	304	43.58	13.59	18 - 69
Education	304	13.00	2.67	3 – 20
Time since accident (mo)	292	30.40	28.30	2 - 196
Time in Canada (yr)	88	22.97	13.15	0 – 57
Age at kmmigration	88	23.70	12.60	1 – 53
Interpreter utilized	282	9.2%		
Handedness	282	Right 88.3%		
Gender	304	Female 48.7%		
TBI severity	280	Unc. mTBI 75.0%	mTBI 89.3%	
Country of origin	304	Canada 62.5%		

Note. TBI Severity = Traumatic brain injury severity; Unc. mTBI = Uncomplicated mild traumatic brain injury; mTBI = mild traumatic brain injury.

Power Analysis

An a priori power analysis using G*Power 3 was conducted for a one-way ANOVA to calculate adequate sample sizes for the studies (Faul, Erdfelder, Lang, & Buchner, 2007). The power analysis used seven predictor groups (the seven education level groups), 0.05 alpha level, 0.8 power level, and 0.40 estimated effect size, which is conservative given previous research using similar methodology (Erdodi, Roth, et al., 2014; Erdodi et al., 2016). The power analysis estimated a sample size of 140 would be

required to find significant results. A larger sample size of 325 was used in the current study to account for multiple analyses (Cohen, 1988).

Measures

Measures included in the current research are copyrighted. According to the College of Psychologists of Ontario Standards of Professional Conduct (2017), except when required by law, test materials should never be released. Furthermore, many of the measures used in the current research involve apparatuses that cannot be attached to documents. As such, measures are not provided in appendices. Measures that were included in both objectives are described below, and measures that were included in only the first or second objective are described in their respective measures sections.

Performance validity tests. This section describes the PVTs included in the current study and the cutoff scores that were selected a priori for each PVT based on the cutoffs suggested in prior research studies. However, not all of the preselected cutoff scores fit well with the data in the current study. Adjustments to the methodology were necessary, and are described in detail in Chapter 5.

Stand-alone PVT. Test of Memory Malingering (TOMM; Tombaugh, 1996).

The TOMM is a stand-alone, two-alternative forced-choice PVT that was designed to identify noncredible memory impairment. The TOMM consists of two learning trials and one retention trial. In the learning trial a series of 50 line drawings of common objects are presented, followed by a series of 50 panels with two line drawings – a target and foil – and the examinee chooses the one that had been previously presented. Feedback is given after each response. The retention trial, consisting only of the recognition portion, is administered about 15 minutes after Trial 2. One point is awarded for each correct

response on each trial, and total scores for each trial are compared to cutoffs. Trial 2 and Retention were not administered to individuals who scored ≥ 49 on Trial 1, as they were assumed to be credible based on this performance. Likewise, Retention was not administered to individuals who scored ≥ 45 on Trial 2 in this dataset. Internal consistencies were reported in the manual ($N = 40$) as Trial 1 $r = .94$, Trial 2 $r = .95$, and Retention $r = .94$ (Tombaugh, 1996). The standard cutoff is ≤ 44 on Trial 2 (Tombaugh, 1996). Recent studies, however, introduced more liberal cutoffs of ≤ 41 on Trial 1 (Greve et al., 2006), ≤ 47 on Trial 2 (Greve et al., 2006), and ≤ 48 on Retention (Greve et al., 2006). For this study, failure was planned to be defined as Trial 1 ≤ 41 , Trial 2 ≤ 47 , or Retention ≤ 48 .

Embedded PVTs.

California Verbal Learning Test Second Edition; (Delis et al., 2000). The California Verbal Learning Test Second Edition is a list-learning task. The primary purpose of the test is to assess semantic strategy use in auditory verbal learning. The test administrator reads a list of 16 words to the participant five times (List A, Trials 1 - 5), and the examinee recalls as many words as possible after each trial. These words belong to four semantic categories: furniture, vegetables, means of transportation, and animals. A second list of 16 words (List B) is then administered. The examinee then recalls words from List A (short delay free recall), and is then asked to produce words from each semantic category (short delay cued recall). The procedure is repeated after a 20-minute delay (long delay free recall and long delay cued recall trials). A 48-item yes/no recognition trial is then administered, which includes words from List A, List B, and novel words. A 16-item two-alternative forced-choice recognition trial is administered 20

minutes later, with completely unrelated foils. Verbatim responses are entered into a computer scoring software, which provides age-, gender-, and education-corrected z -scores for the number of correct responses for each trial, repetition errors, intrusion errors, false positive errors, and discrimination between true positive and false positive responses on recognition. A T -score is provided for the total correct responses across the 5 learning trials. The reader is directed to the test manual for descriptions of other parameters provided by the software that are not included in the current analysis (Delis et al., 2000). Internal consistency was reported at $r = .82$ (Delis et al., 2000). Test-retest reliability has been reported as $r = .75$ for Trials 1 – 5, and $r = .75$ for long delay free recall (Calamia, Markon, & Tranel, 2013).

Several PVTs were derived from this test. Recognition Hits refers to the number of correctly identified items from List A during the yes/no recognition trial, and the selected cutoff rate was ≤ 10 (Greve et al., 2009; Wolfe et al., 2010). A score ≤ 15 on Forced Choice Recognition was considered a failure (Root, Robbins, Chang, & VanGorp, 2006; D. Delis, personal communication, May 10, 2012). A logistic regression equation based on long delay free recall, total recall discriminability and d prime developed by Wolfe et al. (2010) was also used to calculate the probability of noncredible performance. A Logistic Regression Equation value ≥ 0.625 was considered a failure.

Wechsler Adult Intelligence Scale-Fourth Edition (WAIS-IV; Wechsler, 2008)
Coding. This subtest is a digit-symbol substitution task that measures graphomotor processing speed and procedural memory. The examinee is given a sheet of paper with a coding key showing pairs of numbers and their corresponding symbols, as well as a series of symbols without their matching numbers. The examinee is requested to write the

number that corresponds to each symbol. The total number of correct responses in the 120 second time interval is recorded and converted to an age-corrected scaled score. Cronbach's alpha in the standardization sample ($N = 2,200$) was .86 (Wechsler, 2008). Test-retest reliability has been reported as 0.85 (Calamia et al., 2013). Invalid responding was to be defined as age-corrected scaled score ≤ 5 (Erdodi, Abeare, et al., 2017; Etherton et al., 2006).

WAIS-IV Symbol Search. This subtest measures visuomotor processing speed. The examinee's task is to indicate whether one of the two target symbols has a match in an array. The examinee is asked to complete as many items as possible in 120 seconds. The total score is the number of correct responses less the number of errors, which is converted to an age-corrected scaled score. Cronbach's alpha in the standardization sample ($N = 2,200$) was .81 (Wechsler, 2008). Test-retest reliability has been reported as $r = .74$ (Calamia et al., 2013). An age-corrected scaled score ≤ 5 was considered failure (Erdodi, Abeare, et al., 2017; Etherton et al., 2006).

Wechsler Memory Scale-Third Edition (WMS-III; Wechsler, 1997b) Digit Span. Digit Span is a test of auditory working memory, encoding, and attention. First, the examiner reads a list of digits to the examinee, and the examinee recalls the list. The second component is administered similarly, except that the examinee is asked to recite the digits in backward order. The third component is also administered similarly, but the examinee is asked to sequence the digits in numerical order. Cronbach's alpha for Digit Span is reported as .90 (Wechsler, 1997b). The total number of correct responses for each component is converted into an age-corrected scaled score. The base rate of the longest span of correct responses for each component is then calculated.

The utility of the various derived PVTs for this test have been compared between test versions (WAIS-III, WAIS-R, WMS-III, and WMS-R; Jasinski, Berry, Shandera, & Clark, 2011) and were found to be comparable. The following validity cutoffs were selected a priori: Reliable Digit Span ≤ 7 (Babikian et al., 2006; Jasinski et al., 2011; Schroeder, Twumasi-Ankrah, Baade, & Marshall, 2012); age corrected scaled score ≤ 7 (Axelrod et al., 2006; Jasinski et al., 2011); Longest Digits Forward ≤ 4 (Babikian et al., 2006; Heinly, Greve, Bianchini, Love, & Brennan, 2005); and Longest Digits Backward ≤ 3 (Heinly et al., 2005; Yang et al., 2012).

WMS-III Spatial Span. Spatial Span is a test of visuospatial attention and working memory following the Corsi blocks paradigm. A board with 10 blocks adhered in a nongrid pattern is used. In the first component, the examiner taps some blocks, and the examinee copies the examiner's pattern – similar to Digit Span, with increasing numbers of blocks included over time. In the second component, the examinee taps the blocks in the reverse order to the assessor. The total number of correct responses for each component is converted to an age-corrected scaled score. Although considerably less research has been completed on the use of Spatial Span scores as measures of performance validity, this subtest may be particularly useful in the current study because it was designed to be a less verbally mediated analogue to Digit Span. Split-half reliability for Spatial Span has been reported at $r = .77$ (Wechsler, 1997b). Reliable Spatial Span ≤ 7 was considered a failure (Ylioja, Baird, & Podell, 2009).

WMS-III Logical Memory. Logical memory tests of episodic verbal memory and consists of two short stories that are read to the examinee. The first story is read to the examinee aloud once, and the examinee is asked to recall the story. The second story is

read twice, and the examinee recalls the story after each recitation. After a 20-30 minute delay, the examinee is asked to recall the stories, and to answer 30 yes/no questions about the stories. The total number of correctly recalled story elements for the short delay and long delay recall are converted to age-corrected scaled scores. Raw total number of correct recognition responses are recorded without normative correction. Split-half reliability has been reported as $r = .88$ for Logical Memory I, and $r = .79$ for Logical Memory II (Wechsler, 1997b). Several validity cutoffs were selected to be derived from this measure: Logical Memory I age-corrected scaled score ≤ 3 (Bortnik et al., 2010); Logical Memory II age-corrected scaled score ≤ 4 (Bortnik et al., 2010); Logical Memory Recognition ≤ 20 (Pearson, 2009); and Weighted Combination Index developed by Bortnik et al. (2010) (Logical Memory II raw + [1.5 x Logical memory Delayed Recognition raw] ≤ 39.5 (Smith et al., 2014).

Wechsler Abbreviated Scale of Intelligence – Second Edition (WASI-II; Wechsler, 2011) Vocabulary. Vocabulary is a test of semantic knowledge and expression, in which the examinee is asked to define a series of words of increasing difficulty. Each response is graded as 0 (*incorrect*), 1 (*partially correct or concrete*), or 2 (*completely correct*), and the sum of item-level scores is converted to an age-corrected scaled score. Split-half coefficients ($N = 2,300$) for Vocabulary were reported at $r = .92$ and test-retest reliability ($N = 215$, mean interval 12 days) was $r = .93$ (Wechsler, 2011). A relational validity cutoff was derived based on the difference score between Vocabulary and Digit Span. A large positive difference (age-corrected scaled score ≥ 4 ; i.e., much better performance on the Vocabulary subtest compared to the Digit Span subtest) indicates noncredible performance. A discrepancy may indicate that Digit Span performance is being

suppressed relative to Vocabulary, whereas Digit Span performance is robust to mTBI (Greve, Bianchini, Mathias, Houston, & Crouch, 2003; Iverson & Tulsky, 2003; Mittenberg, Theroux-Fischera, Zielinski, & Heilbronner, 1995).

Trail Making Test (Reitan, 1992). Part A of this sequencing test is a measure of visual attention and processing speed. The examinee connects a series of circles labelled with increasing numbers as quickly as possible. In addition to visual sequencing, Part B also requires the simultaneous processing and switching between two classes of stimuli (numbers and letters). As such, it is a more complex task that measures executive functions (working memory, mental flexibility). Seconds to complete each trial are converted to gender-, age-, and education -corrected *T*-scores, and the number of errors is tallied. Norms are provided for Caucasian-Americans and African-Americans.

Test-retest reliability has been reported as $r = .77$ for Part B (Calamia et al., 2013), and $r = .77$ for Part A (Nuechterlein et al., 2008). Part A ≥ 62 seconds, Part B ≥ 200 seconds, as well as the ratio $B/A \leq 1.49$, were selected as cutoffs (Egeland & Langfjaeran, 2007; Iverson, Lange, Green, & Frenzen, 2002). Part A was considered an embedded validity indicator with low verbal mediation, as it only requires familiarity with numerals, and Part B and B/A were classified as embedded validity indicators with high verbal mediation, as they require familiarity with the 26-letter Latin alphabet.

Finger Tapping Test (Reitan, 1969). Finger Tapping Test measures psychomotor speed. The examinees tap a device as quickly as possible with their index finger for 10-second intervals, beginning with their dominant hand, and switching to their nondominant hand after five consecutive trials within a 5-tap range, or, if that does not occur, after 10 trials. The average number of taps per trial for each hand is calculated, and

converted to a *T*-score. Test-retest reliability coefficients between $r = .58$ and $r = .93$ have been reported (Strauss et al., 2006). The dominant raw score, nondominant raw score, and difference scores were derived as PVTs. Arnold et al. (2005) developed separate cutoff scores for men and women that are provided in Table 3 alongside cutoffs for the other PVTs with low verbal mediation.

Judgement of Line Orientation (Benton, Hamsher, Varney, & Spreen, 1983).

Judgment of Line Orientation tests spatial ability. The examinee is shown a series of five practice cards, in which there are 11 lines radiating as a semicircle, each labeled with the numbers 1 to 11 on each card. A corresponding card in the booklet has two lines, each corresponding to a line on the labelled drawing, and the examinee is asked to provide the correct numbers that would correspond to those lines. The number of completely correct item responses are tallied. Two points are then added for female examinees, one point is added for examinees age 50-64 years, and three points are added for examinees age 65-74 years. The total scores are then converted to percentiles. Cronbach's alpha has been reported as .90 (Qualls, Bliwise, & Stringer, 2000). Test-retest reliability was reported at $r = .90$ (Montese, Pere, Carme, Francese, & Eduardo, 2001). A raw score ≤ 21 was to be considered a failure (Whiteside, Wald, & Busse, 2011).

Rey-Osterrieth Complex Figure Test (Rey, 1941). In this visual memory test, the examinee is first asked to copy a complex figure, then draw it from memory three minutes and 30 minutes after the initial copy task, followed by a 24-item yes/no recognition task. Total correct drawing components are converted to age-corrected *T*-scores for immediate memory and delayed memory trials. The number of correct recognition responses (true positives + true negatives) is converted to an age-corrected *T*-

score. Total correctly drawn copy components, copy time, and true and false positive and negative scores are converted to age-corrected percentile ranges. Test-retest reliability has been reported as $r = .50$ for copy (Calamia et al., 2013), $r = .96$ for immediate recall, $r = .89$ for delayed recall and $r = .87$ recognition total correct ($N = 12$; Meyers & Meyers, 1995). Median interrater reliability (Pearson product-moment correlations) for raw scores for 15 randomly selected protocols for three independent raters was $r = .94$ (Meyers & Meyers, 1995). Several PVTs were derived, including copy raw score ≤ 25.0 (Whiteside et al., 2011), immediate recall raw score ≤ 10.0 (Sugarman, Holcomb, Axelrod, Meyers, & Liethen, 2015), true positive score ≤ 6 (Sugarman et al., 2015), atypical recognition errors ≥ 1 (Lu et al., 2003), and Weighted Combination Score (copy score + [(true positive recognition – atypical recognition errors) x 3]) ≤ 45 (Lu et al., 2003).

Controlled Oral Word Association (Ruff et al., 1996). In this test of oral letter and semantic fluency, the examinee is asked to generate as many words as possible in one minute. The first three trials consist of words beginning with F, A, and S. Then the examinee is asked to generate as many animals as possible in one minute. Total number of correct, novel words across F, A, and S trials is converted to an age-corrected T -score. Total number of correct animals generated is converted to an age-corrected T -score. Test-retest reliability has been reported as $r = .79$ for FAS, and $r = .74$ for Animals (Calamia et al., 2013; Nuechterlein et al., 2008). Internal consistency for FAS was reported as $r = .83$ (Ruff, Light et al., 1996). T -scores for combined FAS trials ≤ 31 (Curtis et al., 2008; Sugarman & Axelrod, 2015) and Animals trial T -score ≤ 31 (Sugarman & Axelrod, 2015) were selected as derived PVT cutoffs.

Stroop (Benton & Hamsher, 1978; Golden & Freshwater, 2002). In the Stroop task, the examinee is first asked to read a list of colour words as quickly as possible for 45 seconds. Following this, the examinee names the ink colours of a presented page of “X”s printed in blue, green, or red ink. In the final trial, the examinee must name the ink colour in a list where the words are printed in ink colours that do not match the word. Total correctly read or named colours for each trial are converted to *T*-scores via residual scores (i.e., differences between the number of correct responses and predicted correct responses based on age and education). Internal consistency for the subtests has been reported as ($N = 450$) Color $r = .84$, Word $r = .89$, and Color-Word $r = .73$ (Strauss et al., 2006). Residuals (difference between actual and anticipated performance) for the Word trial ≤ -40 , Colour trial ≤ -30 , and Colour-Word trail ≤ -20 were to be used as cutoff scores (Guise, Thompson, Bianchini, & West, 2014).

Multivariate indicators of performance validity. Performance validity has historically been conceptualized as dichotomous, with examinees grouped as either credible or noncredible (Sollman & Berry, 2011). Passing a single PVT cannot be used to infer the credibility of an entire profile, as performance validity likely fluctuates across a long battery (Boone, 2009). Additionally, some individuals are selective about the type of PVT they fail (Cottingham, Victor, Boone, Ziegler, & Zeller, 2014). Therefore, a limited number of PVTs may not detect the full range of invalid responding. As explained previously, signal detection profiles of PVTs and cutoffs are also highly variable, leading to variable clinical interpretation (Green, 2013). There is also a loss of nuance in categorizing people as having either credible or noncredible performance based on a few cutoffs or PVTs.

One solution to overcome the limitations of using dichotomous groupings of credibility is a methodology developed by Erdodi et al. (2014, 2016), which aggregates validity indicators into a continuous measure. Using such a composite measure allows the test-taking behaviour to be monitored over the course of the assessment and a wide range of tests.

A unique feature of this method is the recapturing of the underlying continuity of noncredible performance. This methodology allows the incorporation of multiple cutoffs and multiple PVTs into a continuous measure of performance validity, increasing the sensitivity to noncredible performance while maintaining specificity (Erdodi, Abeare, et al., 2017), and allows the assessor to evaluate the contribution of various PVTs to the ultimate determination of credibility.

An additional strength of this methodology is the ability to compare validity composites nested within different cognitive domains. In a recent study, for example, Erdodi, Abeare, et al. (2017) compared a composite measure of five processing speed based indicators (EI-5_{PSP}) to a composite measure of five embedded validity indicators based on a forced-choice recognition paradigm (EI-5_{FCR}). They found that the EI-5_{PSP} outperformed any of its components. In other words, the EI-5_{PSP} more effectively differentiated credible from non-credible performance than any individual component. They also found that the EI-5_{PSP} had higher BR_{FAIL} for examinees with moderate-to-severe TBI compared to examinees with mTBI, indicating that this particular measure may be most useful with individuals with milder pathology, given the higher risk of false-positive errors exist with individuals with more severe pathology.

The EI paradigm was especially useful in the current study, which aimed to explore the relationship between cultural common factors, psychiatric factors, and PVT failure. The mismatch between the examinee's dominant language and the language of test administration is particularly prominent in tests with high verbal mediation compared to visuo-perceptual tests in bilingual examinees (Gasquoine, Croyle, Cavazos-Gonzalez, & Sandoval, 2007). The effects of verbal mediation were expected to persist in PVT performance in the current sample. Thus, a validity composite was created for the seven embedded validity indicators with high verbal mediation (EI-7_{VER}), and another composite was calculated for the seven embedded validity indicators with lower verbal mediation (EI-7_{VIS}). A third, a domain-neutral composite that included all 14 of the embedded validity indicators was also calculated (EI-14).

Instead of using a single cutoff for each PVT in these calculations, a value of zero, one, two, or three was assigned to each measure representing increasing degrees of failure. The calculation of the EI-7 or EI-14 value was straightforward for measures that had only one embedded validity indicator. In these cases, a value of zero was assigned to scores that pass even the most liberal cutoff available in previous research, indicating a very low probability of noncredible performance on the task. A value of one was to be assigned to scores that failed the most liberal cutoff available in the research. The values two and three were to be assigned to progressively more conservative failure cutoff levels available in the literature. If no alternative cutoffs were present in the literature for a particular measure, values two and three were to be calculated according to top 10% and 5% of the distribution of failure scores in the current data, respectively. Thus, increasing

values assigned to each component represent increasing evidence of noncredible performance.

The calculation of scores for those measures that have multiple embedded validity indicators, such as Digit Span, was more complex. Cutoffs for each embedded validity indicator were assigned a value as described above. Embedded validity indicators within each test were combined such that the maximum score for any single test was three to avoid multiple embedded validity indicators in any test artificially inflating of EI-7 and EI-14 scores. Where an individual obtained scores of zero for all embedded validity indicators within a given test, they were assigned a score of zero for the test. As an example, if all Digit Span embedded validity indicators were passed, a value of zero would be assigned for Digit Span. A failure at any particular level for one embedded validity indicator within a test resulted in that value being assigned for the test overall. In other words, if an examinee had a level two failure for longest Digit Span backwards, but passed all other Digit Span embedded validity indicators, a level two failure was coded for Digit Span. When there were multiple failures of the same magnitude within one test, the value of the failures was added to a maximum score of three. For example, failure level two on longest Digit Span backwards and Reliable Digit Span while passing all other Digit Span embedded validity indicators would result in a value of three being assigned for Digit Span overall. If there were two embedded validity indicator failures of differing magnitudes within a test, the higher magnitude score was used. For example, failure levels of one for Digit Span age-corrected scaled score and two for Reliable Digit Span were coded as level two failure for Digit Span overall. Failing three or more embedded validity indicators for any test resulted in a value of three being assigned for

the test overall. This system captured the unique contribution of each embedded validity indicator (such as Reliable Digit Span or Digit Span age-corrected scaled score) within each test (e.g., Digit Span) without any particular test unduly influencing the overall composite measure (in this case EI-7_{VER} and EI-14).

The total scores of the EI-7_{VER}, EI-7_{VIS}, and EI-14 were calculated as the sum of the scores of their components, which were calculated as described above. Higher scores on the EI composites represent incrementally increasing likelihood of noncredible performance. EI scores ≤ 1 were considered credible, as they represent at most one failure at the most liberal cutoff on a single embedded validity indicator. EI scores of two or three were classified as borderline, as they represent either one failure at a conservative cutoff or multiple failures at liberal cutoff scores. EI composite values ≥ 4 indicate unambiguously noncredible performance, as these scores would require either failure of at least two PVTs at conservative cutoffs or four PVT failures at liberal cutoffs. In this way, each EI captures information about the number and level of embedded validity indicator failures. The continuity in performance validity is preserved, and nuanced information is captured (e.g., an EI-7_{VER} score of 12 represents even less credible performance than an EI-7_{VER} score of four, although both can be classified as noncredible).

Table 3 and Table 4 display a priori cutoffs for the EI-7_{VER} and EI-7_{VIS}. Cutoff scores used to calculate the EI-14 were identical to those used for the EI-7_{VER} and EI-7_{VIS}.

Table 3

A Priori Levels of Failure for EI-7_{VER} Components

EI-7 _{VER} Component		Pass	LIB	INT	CON
Test	EVI	0	1	2	3
DS	ACSS	> 6	6	5	≤ 4
	RDS	> 7	7	6	≤ 5
	LDF	> 4	4	3	- ^a
	LDB	> 3	3	2	- ^a
VC-DS		< 4	4	5	> 5
LM	LMISS	> 3	3	2	1
	LMISS	> 4	4	3	≤ 2
	LMDR	> 21	19-21	16-18	≤ 15
	WCI	> 39.5	27.2-39.5	21.7-27.1 ^a	≤ 21.6 ^b
CVLT-II	FCR	16	15	14	≤ 13
	RecHITS	> 11	10	9	< 9
	LRE	< .625	.625	.70	> .80
COWA	FAS T	> 31	28-31	25-28	≤ 24
	Animals T	> 31	25-31	21-24	≤ 20
Stroop	Word Res.	> -40	-40 to -43	-44 to -48	≤ -49
	Color Res.	> -30	-30 to -33	-34 to -40	≤ -41
	C-W Res.	> -20	-20 to -24	-25 to -29	≤ -30
TMT	Part B.	< 199	200-624.87	624.88-644.81 ^a	>644.82- ^b
	B/A	> 1.49	.7880-1.49	0.7879-0.5231 ^a	<0.5230- ^b

Note. EI-7_{VER} = Erdodi Index Seven – Verbal; LIB = Liberal cutoff; INT = Intermediate cutoff; CON = Conservative cutoff; EVI = Embedded Validity Indicator; DS = Wechsler Memory Scale – Third Edition Digit Span; ACSS = Age Corrected Scaled Score (Jasinski et al., 2011; Axelrod et al., 2006; Etherton et al., 2006); RDS = Reliable Digit Span (Jasinski et al., 2011; Babikian et al., 2006; Etherton, Bianchini, Ciota et al., 2005; Larrabee, 2003); LDF = Longest Digits Forward (Heinly et al., 2005; Babikian et al., 2006); LDB = Longest Digits Backward (Heinly et al., 2005; Yang et al., 2012); VC-DS = Wechsler Abbreviated Scale of Intelligence – Second Edition/Wechsler Memory Scale – Third Edition Vocabulary minus Digit Span (Greve et al., 2003; Iverson & Tulsky, 2003); LM = Wechsler Memory Scale – Third Edition Logical Memory; LM = Logical Memory; LMISS = Logical Memory I Scaled Score (Bortnik et al., 2010); LMISS = Logical Memory II Scaled Score (Bortnik et al., 2010); LMDR = Logical Memory Delayed Recognition Raw Score (Pearson, 2009); WCI = Logical Memory Weighted Combination Index (LM II raw + [1.5 x LMDR raw]; Bortnik et al., 2010; Smith et al., 2014); CVLT-II = California Verbal Learning Test-II; FCR = Forced Choice Recognition Raw Score (Bauer et al., 2005; Root et al., 2006); RecHITS = Recognition Hits (Greve et al., 2008); LRE: = Logistic regression developed by Wolfe et al. (2010), cutoff suggested by Donders & Strong (2011); COWA = Controlled Oral Word Association Test; FAS = Letter fluency test *T*-score (Curtis et al., 2008; Whiteside et al., 2015); Animals = Category fluency test *T*-score (Sugarman & Axelrod, 2015; Whiteside et al., 2015); Word = Word Residual Score (Guise et al., 2014); Color = Color Residual Score (Guise et al., 2014); C-W Res. = Color-Word Residual Score (Guise et al., 2014); TMT = Trail Making Test; Part B. = Time Trial B in seconds (Iverson et al, 2002); B/A = Trial B time/Trial A time (Iverson et al., 2002; Egeland & Langfjaeran, 2007).

^a Value was calculated based on the top 10% failed scores in the sample.

^b Value was calculated based on the top 5% failed scores in the sample.

Table 4

A Priori Levels of Failure for EI-7_{VIS} Components

EI-7 _{VIS} Component		Pass	LIB	INT	CON
Test	EVI	0	1	2	3
SpatSp	RSS	> 7	7	6	≤ 5
JLO	Raw Score	> 21	19-21	16-18	≤ 15
RCFT	Copy Raw	> 25	25	24	≤ 23
	Imm. Rec. Raw	> 10	9.5-10	8-9	≤ 7
	True Pos.	> 6	6	5	≤ 4
	Atypical	0	1	2	3
	WCI	> 48	46-48	43-45	≤ 42
CD	ACSS	> 5	5	4	≤ 3
SS	ACSS	> 5	5	4	≤ 3
FTT	Dominant (F)	> 34	29-34	16-28	≤ 15
	Dominant (M)	> 39	36-39	22-35	≤ 21
	Nondominant (F)	> 30	26-30	15-25	≤ 14
	Nondominant (M)	> 35	31-35	26-30	≤ 25
	Combined (F)	> 63	59-63	46-58	≤ 45
	Combined (M)	> 73	67-73	59-66	≤ 58
	Difference (F)	> -2	-2 or -3	-4	≤ -5
	Difference (M)	> -1	-2 or -1	-5 to -2	≤ -6
TMT	Part A.	< 62	-	61-187	≥ 188 ^a

Note. EI-7_{VIS} = Erdodi Index-7 Visuomotor; LIB = Liberal cutoff; INT = Intermediate cutoff; CON = Conservative cutoff; EVI = Embedded Validity Indicator; SpatSp = Wechsler Memory Scale – Third Edition Spatial Span; RSS = Reliable Spatial Span (Yllogia et al., 2009); JLO = Judgement of Line Orientation (Whiteside et al., 2011); RCFT = Rey-Osterrieth Complex Figure Test; Copy Raw = Copy Trial Raw Score (Lu et al., 2003; Whiteside et al., 2011); Imm. Rec. Raw = Immediate Recall Raw Score (Lu et al., 2003; Reedy et al., 2013); True Pos. = Recognition True Positive (Lu et al., 2003; Reedy et al., 2013); Atypical = Atypical Recognition Errors (Lu et al., 2003); WCI = Weighted Combination Index (copy score + [(true positive recognition – atypical recognition errors) x 3]; Lu et al. 2003); CD = Wechsler Adult Intelligence Scale - Fourth Edition Coding (Etherton et al., 2006; Erdodi, Abeare, et al., 2017); ACSS = Age Corrected Scaled Score; SS = Wechsler Adult Intelligence Scale - Fourth Edition Symbol Search (Etherton et al., 2006; Erdodi, Abeare, et al., 2017); FTT = Finger Tapping Test; (F) = Female; (M) = Male; Dominant = Dominant hand mean taps (Arnold et al., 2005; Axelrod et al., 2014); Nondominant = Nondominant hand mean taps (Arnold et al., 2005); Combined = Combined Dominant + Nondominant mean taps (Arnold et al., 2005); Difference = Dominant minus Nondominant mean taps (Arnold et al., 2005); TMT = Trail Making Test; Part A. = Time Trial A in seconds (Egeland & Langfjaeran, 2007; Iverson et al, 2002).

^a Value was calculated based on top 5% failed scores in sample.

Missing Data

Data were not found to be missing completely at random given that Little's MCAR was significant $\chi^2(5046, N = 303) = 5473.50, p < .001$. Examinees had the right to refuse testing, and some examinees refused to complete most of the battery. These individuals often failed the PVTs that they did complete. The absence of PVT data in these cases cannot be considered a "Pass." At the same time, treating missing data as "Fail" is equally unacceptable. To handle the missing data for the calculation of the EI-7_{VER}, EI-7_{VIS}, and EI-14, cases were excluded if three or more components were missing, or if data were missing and the examinee failed the TOMM. If the case had missing tests ≤ 2 and the individual passed the TOMM, missing scores were counted as a "Pass" (i.e., score of 0).

Self-reported psychiatric symptom severity was also not missing completely at random. As described above, some of the examinees refused to complete large portions of the battery. Additionally, some examinees had poor attendance and pain behaviour (e.g., taking frequent and extended breaks, lying down with complaints of pain and fatigue after very brief testing) that interfered with the completion of the battery. When pain behaviour and poor attendance interfered with assessment completion, longer self-report inventories were the least likely to be completed. Additionally, some examinees did not complain of psychiatric symptoms during the interview, and the neuropsychologist would often remove some self-report inventories from the battery for these examinees. Given that the self-report measures were not missing completely at random, data imputation is inadvisable (Field, 2009). Missing data from EI and VI scales were treated as described on pages 85-88, and pairwise deleted when there was

insufficient data to calculate the EI or VI variable for the case. Missing TOMM and self-report inventory scores were pairwise deleted from analyses.

CHAPTER 5

Methodological Adjustments and Descriptive Results

Methodological Adjustments

Several modifications to the a priori methodology were necessary after examination of the data in the current study. The modifications are explained and presented in the following section.

Removing older adults from analyses. Older adults have increasingly higher rates of neurological deficits as they age, including but not limited to stroke (Grysiewicz, Thomas, & Pandey, 2008) and dementia (James & Schneider, 2010). Neuropathology may be undiagnosed in the cases of early-stage dementia (Nogueras, Postma, & Van Son, 2016) and silent stroke (Vermeer, Longstreth, & Koudstaal, 2007). These conditions may also be subclinical or unreported in clinical files. Even older adults who score in the normal range on screening tests are likely to demonstrate impairment on more thorough neuropsychological testing (Votruba, Persad, & Giordani, 2016). Additionally, older adults with TBI are at higher risk for both hemorrhagic and ischemic stroke even in the post-acute period following TBI (Albrecht et al., 2015), which increases the risk of misclassification of injury severity in the current study.

The risk of cognitive decline, unrelated neuropathology, and the risk of more serious secondary pathology following the TBI acquired in the motor vehicle accident may confound credibility classification. Specifically, in these cases, credible low performance resulting from unrelated decline or more severe neuropathology may be misclassified as noncredible performance (i.e., PVT failure). PVTs are not robust to dementia (Strauss et al., 2006; Tombaugh, 1996), and many of the embedded validity

indicators in the current research would likely be affected by undetected stroke (e.g., lower Finger Tapping Test scores in a person with subtle hemiparesis). The parameters of these confounds are undefinable in the data set and therefore cannot be reliably controlled statistically. Previous research convention (Pearson, 2009) suggests a cutoff of ≥ 70 years to mitigate the risks of these confounds in PVT interpretation. The Advanced Clinical Solutions for the WAIS-IV and WMS-IV does not provide cutoffs or interpretive guidelines for PVTs for these tests for examinees age ≥ 70 years because of the multiple confounds to accurate PVT interpretation in older adults (Pearson, 2009).

Independent samples t tests and χ^2 analyses were conducted to explore possible differences between the older adult group and the remaining examinees. Examinees ≥ 70 years old did not differ from younger examinees in terms of gender $\chi^2(1, N = 325) = .56, p = .453$, but were more than twice as likely to have been born outside of Canada $\chi^2(1, N = 325) = 12.98, p < .001$, relative risk = 2.30. Examinees ≥ 70 years had lower levels of education ($M = 10.95, SD = 3.76$) than younger examinees ($M = 13.00, SD = 2.67$) $t(21.42) = 2.45, p = .023, g = .75, 95\% \text{ CI } [.31, 3.78]$. These differences are likely demographic artifacts (i.e., in the general Canadian population older adults have lower average levels of education and are more likely to have immigrated to Canada than their younger counterparts; Turcotte & Schellenberg, 2007).

Examinees ≥ 70 years old did not differ from younger examinees on TOMM Trial 1 failure (raw score ≤ 38) $\chi^2(1, N = 319) = 1.26, p = .262$, TOMM Trial 2 failure (raw score ≤ 44) $\chi^2(1, N = 218) = 1.55, p = .213$, or TOMM Retention (raw score ≤ 44) $\chi^2(1, N = 99) = 1.14, p = .286$. Independent t tests were conducted comparing examinees ≥ 70 years to examinees ≤ 69 years on TOMM raw scores and the EI-7_{VER}, EI-7_{VIS}, Validity

Indicator – Ten (VI-10; described below), and their component measures. TOMM, EI-7_{VER}, EI-7_{VIS}, VI-10, and significant component results are presented in Table 5.

Results indicated that older examinees had higher scores on the EI-7_{VIS} and VI-10, and several of the embedded validity indicators, including Rey Complex Figure Test Copy, Judgment of Line Orientation, Reliable Spatial Span, Reliable Digit Span, and Logical Memory Recognition. It should be noted that each of the components with significant age differences were based on raw scores, whereas many of the nonsignificant component scales were corrected for age. Age correction accounts for the typical variations in cognitive functioning across the lifespan, and the lack of age correction in the above-noted embedded validity indicators is one possible reason for the significant findings. Furthermore, the score differences are likely a result of declining cognitive and sensory function, rather than noncredible responding. The differences between the older adult group and examinees ≤ 69 years supports the removal of the older adults from the analyses as differences between the groups may be the result of undetected neuropathology which would confound the interpretation of results. Individuals ≥ 70 years were therefore removed from the analyses ($n = 22$), per previous research convention (Pearson, 2009), to mitigate the risk of these confounds.

Table 5

Independent Samples t Tests on Examinees ≥ 70 Years and ≤ 69 Years on TOMM, EI-

7_{VER}, EI-7_{VIS}, VI-10, and their Components

Test	Group (years)	<i>n</i>	<i>M</i>	<i>SD</i>	<i>df</i>	<i>t</i>	<i>p</i>	<i>g</i>	95% CI
TOMM 1	≤ 69	300	40.22	8.94	317	.77	.443	.18	-2.56, 5.83
	≥ 70	19	38.58	10.16					
TOMM 2	≤ 69	207	41.52	10.43	216	1.32	.190	.41	-2.12, 10.62
	≥ 70	11	37.27	10.70					
TOMM Rec	≤ 69	92	33.14	10.93	97	.68	.498	.27	-5.47, 11.18
	≥ 70	7	30.29	6.26					
EI-7 _{VER}	≤ 69	261	2.69	3.06	270	.81	.420	.25	-1.09, 2.62
	≥ 70	11	3.45	3.05					
EI-7 _{VIS}	≤ 69	268	2.71	3.49	281	2.63	.009	.70	.61, 4.24
	≥ 70	15	5.13	3.14					
VI-10	≤ 69	268	3.68	4.44	280	2.50	.013	.68	.65, 5.43
	≥ 70	14	6.71	4.34					
FTTDH	≤ 69	293	33.19	14.79	310	4.31	<.001	1.02	8.11, 21.71
	≥ 70	19	18.28	11.14					

Test	Group (years)	<i>n</i>	<i>M</i>	<i>SD</i>	<i>df</i>	<i>t</i>	<i>p</i>	<i>g</i>	95% CI
RCFT	≤ 69	293	29.09	5.89	18.87	3.11 ^a	.006	1.14	2.29, 11.73
Copy Raw	≥ 70	19	22.08	9.70					
JLO	≤ 69	262	21.87	6.01	275	3.24	.001	.86	2.01, 8.26
	≥ 70	15	16.73	5.30					
RSS	≤ 69	289	7.76	2.03	308	5.01	<.001	1.18	1.39, 3.18
	≥ 70	21	5.48	1.83					
RDS	≤ 69	301	8.28	2.17	320	3.59	<.001	.81	.79, 2.71
	≥ 70	21	6.52	2.09					
LMR	≤ 69	291	22.91	4.18	309	4.56	<.001	1.09	2.66, 6.47
	≥ 70	20	18.35	4.44					

Note. ^a = Levene's test for equality of variance significant, *t* test with equal variances not assumed reported. *g* = Hedge's *g*; 95% CI = 95% Confidence interval; TOMM 1 = Test of Memory Malinger Trial 1 Raw Score (Tombaugh, 1996); TOMM 2 = Test of Memory Malinger Trial 2 Raw Score (Tombaugh, 1996); TOMM Rec = Test of Memory Malinger Recognition Trial Raw Score (Tombaugh, 1996); EI-7_{VER} = Erdodi Index Seven – Verbal; EI-7_{VIS} = Erdodi Index Seven – Visuomotor; VI-10 = Validity Index Ten; FTT DH = Finger Tapping Test Dominant Hand Raw Score (Arnold et al., 2005; Axelrod et al., 2014); RCFT Copy Raw = Rey-Osterrieth Complex Figure Test Copy Raw Score (Lu et al., 2003; Whiteside et al., 2011); JLO = Judgment of Line Orientation Raw Score (Whiteside et al., 2011); RSS = Wechsler Memory Scale – Third Edition Reliable Spatial Span (Yliogia et al., 2009); RDS = Wechsler Memory Scale – Third Edition Reliable Digit Span (Jasinski et al., 2011); LMR = Wechsler Memory Scale – Third Edition Logical Memory Recognition (Pearson, 2009).

Cutoff adjustment. The data set BR_{FAIL} is the most pressing concern for analysis and interpretation. BR is the driving force behind classification accuracy, as these BRs strongly influence the sensitivity, specificity, positive predictive power, and negative predictive power of any instrument or combination of instruments (Larrabee, Millis & Meyers, 2009; Young, 2015a). Higher BR_{FAIL} of a measure relates to higher sensitivity and lower specificity, whereas lower BR_{FAIL} is associated with lower sensitivity and higher specificity compared to the classification criterion (e.g., failure of another PVT, simulation group membership). The previous discussion of instrumentation bias within the *PVT Limitations* section provides a detailed discussion of some of the ways that these differences in the data result in highly variable cutoffs across studies (pp. 30-32). These effects of BR_{FAIL} and the inverse relationship between sensitivity and specificity exist in any signal detection model—i.e., the easier it is to detect a signal, the higher the risk for detecting both true positives and false positives (Labarge, McCaffrey, & Brown, 2003). These differences in BR result in some PVTs with higher BR_{FAIL} , such as the Word Memory Test (Green, 2003), having high sensitivity to invalid performance, but comparatively low specificity (Eglit, Lynch, & McCaffrey, 2016; Greve, Ord, Curtis, Bianchini, & Brennan, 2008). Others, such as the TOMM at standard cutoffs (Tombaugh, 1996), have very high specificity but are relatively insensitive to noncredible performance (Greiffenstein, Greve, Bianchini, & Baker, 2008). Previous research has shown $60\% \pm 10\%$ credible performance with examinees in litigation (Larrabee et al., 2009), and $15\% \pm 15\%$ of litigation samples are classifiable as “definite malingered neurocognitive deficit” using the Slick Criteria (Slick et al., 1999; Young, 2015b).

Managing false positive rates, then, must include the use of cutoffs with moderate BR_{FAILS} (Larrabee, 2012).

In a cursory overview of the current data, it was apparent that the BR_{FAIL} for the TOMM as well as many of the embedded validity indicators exceeded the highest estimates of BR_{FAIL} in the research literature. Table 6 presents the TOMM BR_{FAIL} for liberal and conservative cutoff scores in the current data. These results indicate that the BR_{FAILS} for all trials of the TOMM are approximately twice as high as those found in previous literature from which the cutoffs were derived (e.g., Greve et al., 2006; Haber & Fichtenberg, 2006; Jones, 2013; Tombaugh, 1996). It should be reiterated that TOMM Trial 2 was only administered with individuals who scored ≤ 48 on Trial 1, and Retention was only administered with examinees who scored ≤ 45 on Trial 2, leading to the comparatively low administration rates of these two trials. This likely contributed to the very high BR_{FAIL} of the subset of examinees who completed the latter trials, as they were poorer performers on the former trials. Lack of Trial 2 and Retention data for many examinees may have also contributed to higher false negative rates *in the overall sample*, as some of the examinees who were missing later trial data may have failed the later trials had they had the opportunity to complete them.

Table 6

TOMM Percent Administered, Means, Standard Deviations, Cutoff Scores and

Associated BR_{FAIL}

Test	% _{ADM}	VI	M	SD	Conservative cutoff		Liberal cutoff	
					CS	BR _{FAIL}	CS	BR _{FAIL}
TOMM	98.7	Trial 1	40.22	8.94	≤ 38	34.7	≤ 42	51.7
	68.1	Trial 2	41.52	10.43	≤ 44	44.4	≤ 49	76.2
	30.3	Ret.	33.14	10.93	≤ 44	85.9	≤ 48	100

Note. %_{ADM} = Percent administered; VI = Validity Indicator; CS = Cutoff score; BR_{FAIL} = Base Rate of Failure; TOMM = Test of Memory Malinger (Tombaugh, 1996); Trial 1 = Trial 1 raw score (Greve et al., 2006); Trial 2 = Trial 2 raw score (Erdodi & Rai, 2017; Tombaugh, 1996); Ret. = Retention Raw Score (Greve et al., 2006; Tombaugh, 1996).

Further, BR_{FAIL} of the EI-7_{VER} and EI-7_{VIS} were extremely high when compared to previous literature using the model. The EI model has evolved since the inception of the current research. Since the initial development of the present study, the model has become explicitly yoked to each study's sample BR_{FAIL} (Erdodi, Abeare, et al., 2017). Recently, Erdodi and colleagues have calibrated the intermediate and conservative cutoffs to correspond to the 10th and 5th percentiles in the data for PVTs with wider ranges that do not have intuitive a priori segmentation (e.g., Erdodi, Abeare, et al., 2017; Erdodi, Tyson, et al., 2017). For example, PVTs based on *T*-scores, such as FAS, would be calibrated according to BR, whereas PVTs with narrow ranges like Reliable Digit Span would not. Previously the cutoffs had been calculated using cutoff scores found in the extant literature (Erdodi et al., 2016; Erdodi, Roth, et al., 2014). These corrections to cutoff scores accounting for BR_{FAIL} help to offset the considerable limitations to analysis and interpretation that arise with very high or low BR_{FAIL} , and thus improve the internal validity of the design. It may also help to mitigate the likelihood of false positive errors, the risk of which increases as more PVTs are used without adjustment (Berthelson, Mulchan, Odland, Miller, & Mittenberg, 2013).

BR_{FAIL} of the EI-7_{VER} and EI-7_{VIS} in this data set far exceeded those of previous EI research. In previous EI research, 40-65% of each sample had a score of 0 or 1 (i.e., Pass) on the EI, regardless of sample characteristics and which PVTs were included in the EI (Erdodi et al., 2016; Erdodi, Abeare, et al., 2017; Erdodi, Kirsch, et al., 2014; Erdodi & Roth, 2017; Erdodi, Roth, et al., 2014). Table 7 presents the BR_{FAIL} for the EI-7_{VER} and EI-7_{VIS} in the current sample using a priori cutoff scores. It is readily apparent that the BR_{FAIL} using these cutoff scores far exceeded those of previous studies, with only 30.7%

and 20.3% of the sample having a score of 0 or 1 (i.e., Passing) the EI-7_{VER} and EI-7_{VIS}, respectively. Further to this, 16.1% and 25.9% of the sample had scores ≥ 8 on the EI-7_{VER} and EI-7_{VIS}, respectively, whereas previous research has consistently shown $\leq 10\%$ of samples have EI scores ≥ 8 (e.g., Erdodi, Abeare, et al., 2017). The immediate implication of these findings is that this sample performed much more poorly across tests compared to previous samples, which is problematic both from a conceptual i.e., what caused this sample to perform so poorly? — and practical point of view —i.e., statistical analysis and interpretation become challenging and ecological validity is limited.

Table 7

Frequency, Percentage, Cumulative Percentage and Classification Ranges for A Priori

EI-7_{VER} and EI-7_{VIS}

N	EI-7 _{VER}			EI-7 _{VIS}			Classification
	<i>f</i>	%	% <i>Cumulative</i>	<i>f</i>	%	% <i>Cumulative</i>	
	49	18.8	18.8	31	11.7	11.7	PASS
1	31	11.9	30.7	23	8.6	20.3	Pass
2	28	10.7	41.4	21	7.9	28.2	Borderline
3	38	14.6	55.9	29	10.9	39.1	Borderline
4	23	8.8	64.8	23	8.6	47.7	Fail
5	17	6.5	71.3	23	8.6	56.4	Fail
6	18	6.9	78.2	25	9.4	65.8	FAIL
7	15	5.7	83.9	22	8.3	74.1	FAIL

Note. N = 261. Examinees ≥ 70 years old, examinees with ≥ 3 missing EI variables, and examinees with 1 or 2 missing variables and TOMM Trial 1 ≤ 40 or TOMM Trial 2 ≤ 44 removed from analyses; EI-7_{VER} = Erdodi Index – Seven Verbal; EI-7_{VIS} = Erdodi Index – Seven Visuomotor; *f* = Frequency; % *Cumulative* = Cumulative percent.

These concerns are particularly important given that the EIs in the current research rely on embedded validity indicators rather than stand-alone PVTs. Traditional stand-alone PVTs do not follow a normal distribution but have strong negative skew even in neurologically impaired samples. This skew increases confidence that low scores on these measures are indicative of noncredible performance (e.g., Tombaugh, 1996). Conversely, embedded validity indicators are derived from tests that measure cognitive abilities, which are typically normally distributed in the population, and many of which are sensitive to neurological impairment (Erdodi & Lichtenstein, 2017). This creates a challenging situation in which it is difficult to distinguish genuine neurological impairment from noncredible performance for some embedded validity indicators (Erdodi & Lichtenstein, 2017).

Compounding the difficulty, even healthy individuals will often obtain one or more abnormal scores on neuropsychological batteries (Binder, Iverson, & Brooks, 2009). A related issue that further complicates the interpretation of scores is that about one-third of examinees in this data set are immigrants to Canada. As previously reviewed, there are few studies on the generalizability of the embedded validity indicators in most cultural groups, and previously developed cutoff scores have not been validated with people who were educated outside of North America or in languages other than English. In sum, the high preponderance of low scores in this data set—that could be naively interpreted as wholly attributable to noncredible performance—may represent false positive errors that are at least partially a result of factors including neurological impairment, normal variance, and cultural factors.

One solution to these challenges—both in a conceptual sense with embedded validity indicators and in a practical sense with the score distribution in the current data set—is to adjust the cutoffs within the EIs. In an effort to align the current study with recent EI research, the cutoffs were adjusted to match the 25th percentile for the most liberal cutoff, in line with the Advanced Clinical Solutions methodology for the first level of failure (Pearson, 2009), to the 10th percentile for the intermediate cutoff and the 5th percentile for the most conservative cutoff, in accordance with recent EI research methodology (Erdodi, Abeare, et al., 2017; Erdodi, Tyson, et al., 2017). The TOMM cutoffs were also adjusted to reflect the originally published cutoffs (Tombaugh, 1996). TOMM failure was defined as ≤ 44 on Trial 2 and/or Retention Trial.

The EI re-scaling convention typically results in roughly 75% of the sample achieving a score of 0 on any given component of the EI, 10% having a score of 1, and each of scores 2 and 3 being assigned to approximately 5% of the sample. Thus, a score of 1 is a relatively weak indicator of noncredible performance on an EVI, whereas a score of 3 is a very strong indicator of noncredible performance on the same EVI. These embedded validity indicator scores are then added together to create the EI composite scores. The re-scaling convention re-establishes the credibility gradient discussed in the *Multivariate Indicators of Performance Validity* section of this work (pp. 89-93). Thus the likelihood of the correct designation of scores as noncredible rises, and the likelihood of false positive errors reduces.

In a related attempt to contain BR_{FAIL} in the current research, embedded validity indicators that were less well established and have been found to be sensitive to the effects of TBI were excluded from analysis. These included Longest Digits Forward,

Longest Digits Backward, Logical Memory I Scaled Score, Logical Memory II Scaled Score, Logical Memory Weighted Combination Index, Stroop Color-Word Trial, Rey Complex Figure Test Immediate Recall, Rey Complex Figure Test Unusual Recognition Errors, Rey Complex Figure Test Combination Score, Finger Tapping Test Nondominant hand, Combined, and Difference Scores, and Trail Making Test-B and B/A ratio (Arnold et al., 2005; Bortnik et al., 2010; Egeland & Langfjaeran, 2007; Guise et al., 2014; Heinly et al., 2005; Lu et al., 2003; Reedy et al., 2013; Smith et al., 2014).

The Trail Making Test-A embedded validity indicator was also changed from being raw score based (Egeland & Langfjaeran, 2007; Iverson et al., 2002) to *T*-score based (Ashendorf, Clark, & Sugarman, 2017). This change was made on conceptual and practical grounds, as performance on Trail Making Test declines significantly with age (Heaton et al., 2004). Thus, with raw-score based validity cutoffs the likelihood of failure increases linearly with age. Use of a *T*-score based embedded validity indicator ensures that examinees are being compared to a normative score that accounts for age-related changes. These adjustments further reduce the likelihood of false-positive errors on the EIs by excluding lower quality PVTs from the analyses.

A truncated version of the EI-14, the Validity Index – 10 (VI-10) was also calculated. The two least frequently administered tests from the EI-7_{VER} (California Verbal Learning Test Second Edition and Stroop), and EI-7_{VS} (Judgment of Line Orientation and Trail Making Test-A) were dropped to create the abbreviated measure. This strategy was used to be consistent with previous research into the simultaneous interpretation of multiple validity tests. Odland et al. (2015) provide sensitivity and specificity values for a range of PVT failures as a function of the number of PVTs

administered. Their calculations are provided for a maximum of 10 PVTs. No other studies have explored sensitivity and specificity for more than 10 PVTs administered in a single battery. Therefore, the multivariate model in the present project was limited to 10 independent PVTs.

Adjusting the cutoff scores, removing the relatively weaker indices, and limiting the number of indices in the model to align them with previous research all serve to improve classification accuracy. The EIs were thus made more conservative to preserve the internal logic of the EI model—that of a gradient of credibility with a high likelihood of correct noncredible designation.

Table 8 and Table 9 display the adjusted cutoff scores and BR_{FAIL} for each embedded validity indicator of the EI-7_{VER} and EI-7_{VIS}, respectively. Table 10 presents the BR_{FAIL} for the EI-7_{VER} and EI-7_{VIS} in the current sample with the adjustments above.

Table 8

Adjusted Levels of Failure for EI-7_{VER} Components

EI-7 _{VER} Component				EI-7 Values				
				0	1	2	3	
Test	% _{ADM}	EVI	Label	Pass 75%le	LIB 25%le	INT 10%le	CON 5%le	
DS	99.0	ACSS	Cutoff	>6	6	5	≤ 4	
			BR	70.5	11.3	10.3	7.9	
			RDS	Cutoff	>6	6	5	≤ 4
			BR	80.9	9.9	6.6	2.6	
VC-DS	92.5	ACSS	Cutoff	<3	3	4	≥ 5	
			BR	89.4	5.7	2.5	2.5	
LM	95.7	Recog	Cutoff	>20	18-20	16-17	≤ 15	
			BR	74.8	14.8	5.8	4.6	
CVLTIII	87.5	FCR	Cutoff	>14	13-14	11-12	≤ 10	
			BR	80.2	10.1	4.8	4.9	
		RechITs	Cutoff	>10	10	8-9	≤ 7	
			BR	81.5	5.6	8.2	4.7	
		LRE	Cutoff	<.625	.625-.724	.725-.814	≥ .815	
			BR	78.4	11.6	4.8	5.2	
FAS	91.1	T-score	Cutoff	>31	28-31	21-27	≤ 20	

				EI-7 Values			
EI-7 _{VER} Component				0	1	2	3
Test	% _{ADM}	EVI	Label	Pass 75%le	LIB 25%le	INT 10%le	CON 5%le
FAS			BR	86.1	4.3	4.7	4.9
Animals	89.5	T-score	Cutoff	>31	24-31	16-23	≤ 15
			BR	78.1	11.4	6.3	4.2
Stroop	86.9	Word	Cutoff	> -47	-52 to -47	-62 to -53	≤ -63
			BR	85.4	5.2	4.3	5.1
		Color	Cutoff	> -40	-	-52 to -40	≤ -51
			BR	90.1		5.0	4.9

Note. EI-7_{VER} = Erdodi Index Seven – Verbal; %_{ADM} : Percent of the sample to which a given test was administered; LIB = Liberal cutoff; INT = Intermediate cutoff; CON = Conservative cutoff; %le = Percentile; EVI = Embedded Validity Indicator; DS = Wechsler Memory Scale – Third Edition Digit Span; ACSS = Age Corrected Scaled Score (Jasinski et al., 2011; Axelrod et al., 2006; Etherton et al., 2006); RDS = Wechsler Memory Scale – Third Edition Reliable Digit Span (Jasinski et al., 2011; Babikian et al., 2006; Etherton, Bianchini, Ciota et al., 2005; Larrabee, 2003); VC-DS = Wechsler Abbreviated Scale of Intelligence – Second Edition/Wechsler Memory Scale – Third Edition Vocabulary minus Digit Span (Greve et al., 2003; Iverson & Tulsky, 2003); LM = Wechsler Memory Scale – Third Edition Logical Memory; Recog = Recognition raw score (Pearson, 2009); CVLTII = California Verbal Learning Test – Second Edition; FCR = Forced Choice Recognition raw score (Bauer et al., 2005; Root et al., 2006); RecHITs = Recognition Hits (Greve et al., 2008); LRE: = Logistic regression equation (Donders & Strong, 2011; Wolfe et al., 2010); FAS = Letter fluency test (Curtis et al., 2008; Whiteside et al., 2015); Animals = Category fluency test (Sugarman & Axelrod, 2015; Whiteside et al., 2015); Word = Word Residual Score (Guise et al., 2014); Color = Color Residual Score (Guise et al., 2014); BR = Base rate (%).

Table 9

Adjusted Levels of Failure for EI-7_{VIS} Components

EI-7 _{VIS} Component				EI-7 Values			
				0	1	2	3
Test	% _{ADM}	EVI	Label	Pass 75%le	LIB 25%le	INT 10%le	CON 5%le
Spatial	95.1	RSS	Cutoff	>6	6	5	≤4
			<i>BR</i>	74.8	13.8	5.9	5.5
JLO	86.2	Raw Score	Cutoff	>19	14-19	10-13	≤9
			<i>BR</i>	71.8	18.6	5	4.6
RCFT	96.4	Copy	Cutoff	>26.5	22.5-26.5	17.5-22	≤17
			<i>BR</i>	78.8	11.6	4.2	5.4
		Recog TP	Cutoff	>6	5-6	4	≤3
			<i>BR</i>	72.3	17.2	5.1	5.4
CD	94.1	ACSS	Cutoff	>4	3-4	2	1
			<i>BR</i>	76.0	15.0	4.9	4.2
SS	94.4	ACSS	Cutoff	>4	3-4	2	1
			<i>BR</i>	76.0	14.6	1.7	7.6

				EI-7 Values			
EI-7 _{VIS} Component				0	1	2	3
Test	% _{ADM}	EVI	Label	Pass	LIB	INT	CON
				75%le	25%le	10%le	5%le
FTT	95.2	Dom (F)	Cutoff	>28.0	11.4-28.0	8.7-11.3	≤8.6
		Raw Score	BR	56.1	34.0	4.9	5.0
	97.3	Dom (M)	Cutoff	>35.0	13.6-35.0	9.5-13.5	≤9.4
		Raw Score	BR	54.5	35.7	5.6	4.2
TMTA	91.8	T-score	Cutoff	>34	21-34	14-20	≤13
			BR	68.6	21.4	5.0	5.0

Note. EI-7_{VIS} = Erdodi Index Seven – Visuomotor; %_{ADM} : Percent of the sample to which a given test was administered; LIB = Liberal cutoff; INT = Intermediate cutoff; CON = Conservative cutoff; %le = Percentile; EVI = Embedded Validity Indicator; RSS = Wechsler Memory Scale – Third Edition Reliable Spatial Span (Yliogia et al., 2009); JLO = Judgement of Line Orientation (Whiteside et al., 2011); RCFT = Rey-Osterrieth Complex Figure Test; Copy = Raw score for the copy trial (Lu et al., 2003; Whiteside et al., 2011); Recog TP = Raw score for recognition true positives (Lu et al., 2003; Reedy et al., 2013); CD = Wechsler Adult Intelligence Scale - Fourth Edition Coding (Etherton et al., 2006; Erdodi, Abeare, et al., 2017); ACSS = Age Corrected Scaled Score; SS = Wechsler Adult Intelligence Scale - Fourth Edition Symbol Search (Etherton et al., 2006; Erdodi, Abeare, et al., 2017); FTT = Finger Tapping Test; (F) = Female; (M) = Male; Dominant = Dominant hand mean taps (Arnold et al., 2005; Axelrod et al., 2014); TMTA = Trail Making Test Trial A (Ashendorf et al., 2017); BR = Base rate (%).

Table 10

Frequency, Percentage, Cumulative Percentage and Classification Ranges for Adjusted

EI-7_{VER}, EI-7_{VIS}, and VI-10

N	EI-7 _{VER}			EI-7 _{VIS}			VI-10			Classification
	f	%	% _{Cumul}	f	%	% _{Cumul}	f	%	% _{Cumul}	
0	79	30.3	30.3	72	26.9	26.9	57	21.3	21.3	PASS
1	38	14.6	44.8	63	23.5	50.4	55	20.5	41.8	Pass
2	34	13	57.9	39	14.6	64.9	29	10.8	52.6	Borderline
3	38	14.6	72.4	28	10.4	75.4	31	11.6	64.2	Borderline
4	18	6.9	79.3	12	4.5	79.9	20	7.5	71.6	Fail
5	15	5.7	85.1	16	6	85.8	21	7.8	79.5	Fail
6	10	3.8	88.9	9	3.4	89.2	9	3.4	82.8	FAIL
7	4	1.5	90.4	7	2.6	91.8	9	3.4	86.2	FAIL
8	10	3.8	94.3	3	1.1	92.9	5	1.9	88.1	FAIL
9	6	2.3	96.6	4	1.5	94.4	3	1.1	89.2	FAIL
10	3	1.1	97.7	3	1.1	95.5	4	1.5	90.7	FAIL

Note. Examinees ≥ 70 years old, examinees with ≥ 3 missing EI variables, and examinees with 1 or 2 missing variables and TOMM Trial 2 ≤ 44 or TOMM Retention Trial ≤ 44 removed from analyses. EI-7_{VER} = Erdodi Index – Seven Verbal; EI-7_{VIS} = Erdodi Index – Seven Visuomotor; VI-10 = Validity Index – Ten; %_{Cumul} = Cumulative percentage.

An obvious limitation to these adjustments is the concern about absolute and somewhat arbitrary suppression of BR_{FAIL} . In other words, it is likely that some of the individuals who were reclassified as credible with the adjustment of cutoffs would be classified as noncredible based on widely accepted cutoffs, leading to a higher false negative error. This trade-off between risk of false positive and false negative error is ubiquitous in signal detection models, as explained previously. In the field of PVTs, where false positive errors have a higher potential for harm to the examinee than false negative errors, control of false positive error is given higher priority at the expense of false negative errors (Boone, 2007), as is the case here. The comparatively large adjustments that were necessary for this research, however, limit the generalizability of findings, at least regarding the cutoffs themselves.

What remains when the high BR_{FAIL} is artificially suppressed are the patterns of performance across groups, which can inform both future research and clinical practice. Specifically, the cutoffs used in the current research cannot be generalized to other research or clinical practice. These cutoffs are very conservative, and would likely lead to high false negative errors in other samples. Nonetheless, the use of these adjusted cutoffs in the current research suppressed BR_{FAIL} systematically across examinees. This method preserved inter-individual and between group performance patterns while reducing the overall BR_{FAIL} enough that data analysis and interpretation were feasible. Thus, the comparison of performance across groups remains valid, even though the specific cutoffs are not generalizable outside of the current research.

General Results

Statistical assumptions.

The assumptions of t tests and ANOVAs are independence of observation, normal distribution, equality of variance, and approximately equal sample size (Stevens, 2009). The assumptions of χ^2 tests are independence of observation and cell sizes being greater than five (Field, 2009). Cell size and equality of variance will be addressed on an analysis-by-analysis basis.

The assumption of independence of observation is not tested statistically, and is instead embedded in the design of the study. Observations in this research were considered independent, as examinees were tested separately and had not interacted with one another (Stevens, 2009).

The assumption of normality was tested for all three trials of the TOMM, and the EI-7_{VER}, EI-7_{VIS}, and VI-10, which are the outcome variables for major analyses throughout both objectives. These results are presented in Table 11. As can be seen in Table 11, the EI-7_{VIS} and VI-10 were negatively skewed and leptokurtotic, and Shapiro-Wilk's Test was significant for each variable. Visual inspection of the data confirmed that none of the EI variables nor trials of the TOMM were normally distributed, and all were negatively skewed. Further, these data were not expected to be normally distributed. Were the data for PVTs normally distributed, it would be problematic because of the BR_{FAIL}. In samples used in previous literature, the majority of the sample is deemed credible (i.e., did not fail PVTs). PVT failure should be capturing the tail of the normal distribution, i.e., the lowest performance. When examined through stand-alone measures this tail would not be normally distributed, and would instead have ceiling effects. When

credibility is examined via embedded validity indicators derived from otherwise normally distributed measures, the embedded validity indicators themselves also create a ceiling effect whereby approximately 75% of examinees pass, as previously explained. However, ANOVAs and *t* tests are robust to violations of normality when sample sizes remain equal (Stevens, 2009).

The assumption of equality of variance was tested on an analysis-by-analysis basis, and nonparametric equivalents of parametric tests were reported in cases where assumptions of normality, equality of variance, and equal sample size were violated, as these cases are particularly prone to spurious findings (Skidmore & Thompson, 2013; Stevens, 2009).

Table 11

Sample Size, Mean, Standard Deviation, Range, Median, Skewness, Kurtosis, and

Shapiro-Wilk's Test for TOMM Trials, EI-7_{VER}, EI-7_{VIS}, and VI-10

	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Range</i>	<i>Mdn</i>	<i>Skewness</i>	<i>Kurtosis</i>	<i>Shapiro-Wilk's Test</i>		
								<i>W</i>	<i>df</i>	<i>p</i>
TOMM1	300	40.22	8.94	12 - 50	31	-1.038	.407	.89	300	<.001
TOMM2	207	41.52	10.43	5 - 50	36	-1.534	1.705	.79	207	<.001
TOMMR	92	33.14	10.93	6 - 48	35	-.711	-.449	.93	92	<.001
EI-7 _{VER}	261	2.69	3.06	0 - 15	2	1.593	2.704	.82	261	<.001
EI-7 _{VIS}	268	2.71	3.49	0 - 18	1	2.101	4.575	.74	268	<.001
VI-10	268	3.68	4.44	0 - 23	2	2.008	4.419	.77	268	<.001

Note. TOMM 1 = Test of Memory Malingering Trial 1 raw score (Tombaugh, 1996); TOMM 2 = Test of Memory Malingering Trial 2 raw score (Tombaugh, 1996); TOMM R = Test of Memory Malingering Recognition Trial raw score (Tombaugh, 1996); EI-7_{VER} = Erdodi Index – Seven Verbal; EI-7_{VIS} = Erdodi Index – Seven Visuomotor; VI-10 = Validity Indicator – Ten.

CHAPTER 6

Objective 1: Cultural, Linguistic, and Demographic Variables and PVTs

Methods

The current objective aimed to elucidate the contribution of cultural, demographic, and injury factors to PVT failure in a sample of examinees who underwent neuropsychological assessment following motor vehicle accidents in Southern Ontario. Common factors included English language proficiency, time in Canada, and level of education. These variables were chosen because correction for common factors that affect PVT performance across multiple cultures may be a pragmatic and ethical solution to the lack of culture-specific norms to match a diverse nation such as Canada.

The particular cultural factors explored were limited to the information available in the archival dataset. The litigious nature of the assessments limited confidence in examinee self-report and performance. Therefore, the use of objective demographic data, such as level of education, first language spoken, age at immigration, and time in Canada represented the most reliable measures of cultural and demographic factors in this data set.

Language Proficiency

Age of acquisition of second language is consistently associated with levels of second language proficiency (Hernandez & Li, 2007). Second language proficiency declines across age of acquisition throughout childhood, and plateaus by age 18 (Dekeyser, Alfi-Shatay, & Ravid, 2010). Although language proficiency declines throughout childhood, childhood age of acquisition can be categorized as early childhood acquisition (≤ 9 years) and late childhood acquisition (≥ 10 years; Archila-Suerte et al.,

2015). Adult acquisition (≥ 18 years) represents a stable lower proficiency group (Dekeyser et al., 2010).

English language proficiency was therefore categorized into four groups for this study: a limited English proficiency group, defined as examinees who immigrated to Canada at the age ≥ 18 and whose dominant language was not English; an intermediate group (immigrated at age 10-17) whose dominant language was not English; a near native-level English speaker group defined as those who immigrated to Canada at age ≤ 9 ; and a native-level English speaker group defined as individuals who were born in Canada and whose first language was English (Archila-Suerte et al., 2015).

Francophone Canadians and Anglophones who immigrated to Canada were excluded from these analyses, as determining their Canadian English dialect proficiency would be potentially unreliable. It was expected that more limited English proficiency would be associated with higher BR_{FAIL} of PVTs with high verbal mediation, but that there would be no relationship with BR_{FAIL} of PVTs with low verbal mediation.

Time in Canada

Time in the host country has been explored as a relevant factor to acculturation (Cheung, Chudek, & Heine, 2010; Fitzpatrick et al., 2015; Uskul & Greenglass, 2005; Vang et al., 2015). In a study of 232 participants who emigrated from Hong Kong to Vancouver, a greater length of time in Canada was associated with greater acculturation for those who immigrated at age ≤ 14 , but not participants who immigrated when older (Cheung et al., 2010). A follow-up study of 569 global immigrants in the United States showed greater acculturation with greater time living in the United States regardless of age at immigration (Chudek, Cheung, & Heine, 2015).

Time in the host country has rarely been explored in research about neuropsychological assessment. One epidemiological study of immigrants with atherosclerosis who had lived in the United States for ≥ 30 years ($N = 544$) showed immigrants having lower Cognitive Abilities Screening Instrument (Teng et al., 1994) and Digit Symbol Coding (Wechsler, 1997a) scores than US-born individuals ($N = 3362$; Fitzpatrick et al., 2015). These differences were attributed to acculturation (Fitzpatrick et al., 2015; Vang et al., 2015).

The existing literature has no consistent cutoffs for time in host country, although 10 or more years in the host country is often used as a cutoff for long-term immigrant status (La Parra-Cassado, Stornes, & Solheim, 2017; Vang et al., 2015). For the current study, time in Canada was calculated as age at assessment minus age at immigration, and was divided according to previous research on the length of stay in host country and Canadian immigration law as Canadian born; lived in Canada ≤ 4 years, 5-9 years, and ≥ 10 years (Citizenship Act, 1985; La Parra-Cassado et al., 2017; Vang et al., 2015). It was expected that shorter time in Canada would be associated with higher BR_{FAIL} with PVTs with high verbal mediation, but would not be associated with BR_{FAIL} of PVTs with low verbal mediation.

Education

As previously explained, lower levels of education have been inconsistently associated with lower PVT performance (Gervais et al., 2004; Stulemeijer et al., 2007). This variable was included in the study to explore the degree to which education affected PVT BR_{FAIL} in the current sample. Level of education was stratified by typical neuropsychological normative groups. The education groups were: ≤ 8 years; 9 to 11

years; 12 years; 13 to 15 years; 16 years; ≥ 17 years (Heaton et al., 2004; Strauss et al., 2006). It was expected that lower educational attainment would be associated with higher BR_{FAIL} for all PVTs.

TBI Severity

In addition to cultural factors, TBI severity has often been associated with PVT BR_{FAIL} . Consistent with the majority of research (e.g., Carone, 2008; Green et al., 1999; Green et al., 2001; Mittenberg et al., 2002; Webb et al., 2012; West et al., 2011), it was expected that individuals with mTBI would have higher BR_{FAIL} than examinees with moderate-to-severe TBI. This finding would indicate that the current sample characteristics are similar to most previous research, and would increase the confidence in the generalizability of the findings.

Gender

Research has not shown associations between gender and BR_{FAIL} (Constantinou & McCaffrey, 2003; Donders, 2005; Rees, Tombaugh, Gansler, & Moczynski, 1998; Webb et al., 2012). As such, the current study explored whether there was any association between BR_{FAIL} and gender as open research questions to assess the consistency of the current study with previous research findings.

Age

Findings about the associated of age and BR_{FAIL} have been mixed (Donders & Boonstra, 2007; Lange et al., 2010; Strauss et al., 2006; Stulemeijer et al., 2007; Webb et al., 2012). As such, the association of age and BR_{FAIL} was examined in this study as an open question.

Hypotheses

1. Cultural common factors would be associated with higher BR_{FAIL} for PVTs with high verbal mediation, but not PVTs with low verbal mediation.
 - a. Limited English proficiency (immigrated to Canada at age ≥ 18 and whose dominant language is not English; immigrated at age 10-17 and whose dominant language is not English; immigrated to Canada at age ≤ 9 and whose dominant language is not English) would be associated with higher BR_{FAIL} for PVTs with high verbal mediation (i.e., $EI-7_{VER}$), but would not be associated with BR_{FAIL} for PVTs with low verbal mediation (i.e., $EI-7_{VIS}$ and TOMM).
 - b. Shorter time in Canada (lived in Canada ≤ 4 years, five to nine years, and ≥ 10 years, or born in Canada) would be associated with higher BR_{FAIL} for PVTs with high verbal mediation, but would not be associated with BR_{FAIL} for PVTs with low verbal mediation.
 - c. Lower educational attainment (defined as education: ≤ 8 years; 9 to 11 years; 12 years; 13 to 15 years; 16 years; and ≥ 17 years) would be associated with higher BR_{FAIL} for PVTs.
2. Consistent with previous research (Mittenberg et al., 2002), moderate-to-severe TBI would be associated with lower BR_{FAIL} for PVTs compared to mTBI.

Research Questions

1. Would gender be associated with differences in BR_{FAIL} ?
2. Would age be associated with differences in BR_{FAIL} ?

Participants

For specific information about examinees included in this study, see the General Methods section of this document.

Measures

Demographic information. Demographic information included gender, age, country of birth, age at immigration (if applicable), first language, and whether an interpreter was utilized to aid in the assessment. The date of accident and date of assessment were coded only by month and year to ensure data de-identification.

Neuroimaging data and Glasgow Coma Scale scores were gathered from the reports and were used to determine injury severity. When the examinee reported a loss of consciousness ≤ 30 minutes and PTA ≤ 24 hours, he or she was classified as having mTBI (Lezak et al., 2012; Teasdale et al., 2014). When there were positive objective neuroimaging findings and Glasgow Coma Scale ≥ 13 the examinee was considered to have mild complicated TBI. When there were positive objective neuroimaging findings and Glasgow Coma Scale < 13 the examinee was considered to have moderate-to-severe TBI. For examinees who reported injury parameters in the moderate-to-severe range in the absence of objective data (imaging or Glasgow Coma Scale), the status of TBI was considered indeterminate.

Performance validity tests. Specific information about PVTs can be found in the General Methods section of this document.

CHAPTER 7

Objective 1 Results

Hypothesis 1a

Limited English proficiency (immigrated to Canada at the age ≥ 18 and whose dominant language is not English; immigrated at age 10-17 and whose dominant language is not English) would be associated with higher BR_{FAIL} for PVTs with high verbal mediation, but would not be associated with BR_{FAIL} for PVTs with low verbal mediation.

French Canadians ($n = 7$) and immigrants whose first language was English ($n = 21$) were excluded from analyses for this hypothesis. It was also necessary to exclude examinees who immigrated to Canada at age ≤ 9 years ($n = 11$) and age 10 to 17 years ($n = 10$) due to small sample size. Comparison of examinees included and excluded from these analyses are depicted in Table 12, indicating that examinees included in analyses were significantly younger than examinees who were excluded, but did not differ significantly on outcome measures, suggesting that excluding this subsample did not introduce a bias in the measurement model. The proportion of genders did not differ across groups $\chi^2(1, N = 281) = .00, p = .98$.

Instead of the originally planned one-way ANOVAs, independent t tests were conducted to compare TOMM Trial 1, EI-7_{VER}, EI-7_{VIS}, and VI-10 scores across examinees born in Canada and those who immigrated to Canada as adults. Hedge's g is reported as the effect size estimate to control for unequal sample sizes. Results of the t tests are displayed in Table 13.

Assumptions of normality and equality of variance were violated for the t tests with TOMM Trial 1, EI-7_{VIS} and VI-10 as the outcome variables. Mann-Whitney U Tests were thus conducted to confirm the significant results, as violations of assumptions normality and equality of variance can affect the results in the case of t tests with unequal group sizes, as was true with these analyses (Skidmore & Thompson, 2013; Stevens, 2009). The tests confirmed that the group of examinees who immigrated to Canada as adults scored higher (i.e., stronger evidence of invalid performance) on the EI-7_{VIS} and VI-10 as compared to the Canadian born group. Examinees who immigrated to Canada also had lower scores (i.e., stronger evidence of invalid performance) on TOMM Trial 1. Contrary to the hypothesis, there was no significant difference in performance between those who immigrated to Canada as adults as compared to Canadian born examinees on the EI-7_{VER}.

Chi-square analyses were conducted with the two aforementioned immigration age groups instead of the four planned groups. Table 14 displays χ^2 results comparing TOMM failure (≤ 44 on Trial 2 or Retention), EI-7_{VER}, EI-7_{VIS}, and VI-10 failure (≥ 4) across examinees born in Canada and those who immigrated to Canada at age ≥ 18 years. Contrary to the hypothesis, and as can be seen in Table 13, examinees who immigrated to Canada as adults were more than twice as likely to fail the TOMM and the EI-7_{VIS} as compared to Canadian born examinees, and were more likely to fail the VI-10 as Canadian born examinees. There was no significant difference in EI-7_{VER} failure rates between Canadian born examinees and those who immigrated as adults.

Exploratory analyses. Follow-up exploratory t tests were conducted to compare TOMM Trial 1, EI-7_{VIS}, and VI-10 scores between examinees who had an interpreter and

those who did not. These tests were not conducted with the EI-7_{VER} due to the small number of examinees who had an interpreter and data for the measures ($n = 6$). Levene's test was significant for each t test, indicating that the assumption of equality of variance was violated. As a result, t tests with equal variances not assumed are reported, along with Mann-Whitney U tests. Results presented in Table 15 indicate that examinees who utilized interpreters had lower TOMM Trial 1 scores, and higher EI-7_{VIS}, and VI-10 scores.

Table 12

Independent t Tests on Age, TOMM Trial 1, EI-7_{VER}, EI-7_{VIS}, and VI-10 Scores in

Examinees Included or Excluded from Hypothesis 1 Analyses

	Age			TOMM 1			EI-7 _{VER}			EI-7 _{VIS}			VI-10		
Inc	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>
Y	253	42.84	13.44	251	39.86	9.00	214	2.64	3.02	222	2.68	3.47	220	3.63	4.42
N	28	51.25	12.73	28	41.64	8.81	25	2.76	3.05	27	3.63	3.88	27	4.30	4.36
<i>df</i>			279			227			237			247			245
<i>t</i>			3.16			1.00			.18			1.32			.74
<i>p</i>			.002			.319			.857			.189			.461
<i>g</i>			.63			.20			.04			.27			.15
CI			3.17, 13.65			-1.74, 5.31			-1.14, 1.37			-47, 2.36			-1.11, 2.44

Note. Inc = Included in analyses; Y = Yes; N = No; *g* = Hedge's *g*; CI = 95% confidence interval; TOMM 1 = Test of Memory Malinger Trial 1 Raw Score (Tombaugh, 1996); EI-7_{VER} = Erdodi Index Seven – Verbal; EI-7_{VIS} = Erdodi Index Seven – Visuomotor; VI-10 = Validity Index Ten.

Table 13

Independent t Tests on TOMM Trial 1, EI-7_{VER}, EI-7_{VIS}, and VI-10 Scores for Examinees by English Proficiency

English Proficiency	TOMM 1			EI-7 _{VER}			EI-7 _{VIS}			VI-10		
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>
High	184	41.85	7.34	175	2.42	2.83	174	2.00	2.53	175	2.94	3.46
Low	47	34.15	10.85	24	2.88	3.53	35	5.51	5.47	30	5.90	6.29
<i>df</i>			57.21			197			36.94			32.08
<i>t</i>			4.61 ^a			.72			3.72 ^a			2.52 ^a
<i>p</i>			<.001			.473			.001			.017
<i>g</i>			.94			.16			1.10			.74
95% CI			4.35, 11.05			-.80, 1.71			1.60, 5.43			.57, 5.36
Mann-Whitney <i>U</i>			2498.00						2726.00			2823.50
<i>z</i>			4.47						4.71			3.65
<i>p</i>			<.001						<.001			<.001

Note. ^a = Levene's test for equality of variance significant, *t* test with equal variances not assumed reported, along with Mann-Whitney *U* Test; *g* = Hedge's *g*; 95% CI = 95% confidence interval; TOMM 1 = Test of Memory Malinger Trial 1 Raw Score (Tombaugh, 1996); EI-7_{VER} = Erdodi Index Seven – Verbal; EI-7_{VIS} = Erdodi Index Seven – Visuomotor; VI-10 = Validity Index Ten; High = Anglophonic and born in Canada; Low = Immigrated to Canada at age ≥ 18 years and first language not English.

Table 14

TOMM, EI-7_{VER}, EI-7_{VIS}, and VI-10 BR_{FAIL} for Examinees by English Proficiency

		TOMM	EI-7 _{VER}	EI-7 _{VIS}	VI-10
English Proficiency	<i>N</i>	232	143	155	154
High	<i>n</i>	185	128	123	131
	BR _{FAIL}	23.8	34.4	23.6	39.7
Low	<i>n</i>	47	15	32	23
	BR _{FAIL}	55.3	33.3	56.3	65.2
χ^2		17.69	.01	12.83	5.19
<i>p</i>		<.001	.94	<.001	.023
Φ^2		.28	.01	.29	.18
RR (95% CI)		2.33	.97	2.39	1.64
		(1.62, 3.35)	(.46, 2.06)	(1.54, 3.71)	(1.14, 2.37)

Note. TOMM = Test of Memory Malingering (Tombaugh, 1996); EI-7_{VER} = Erdodi Index Seven – Verbal; EI-7_{VIS} = Erdodi Index Seven – Visuomotor; VI-10 = Validity Index Ten; High = Anglophone and born in Canada; Low = Immigrated to Canada at age ≥ 18 years and first language is not English; BR_{FAIL} = Base rate of failure (≤ 44 TOMM Trial 2 or Retention; ≥ 4 on EI-7_{VER}, EI-7_{VIS}, and VI-10); RR = Relative risk ratio; 95% CI = 95% confidence interval.

Table 15

Independent t Tests on TOMM Trial 1, EI-7_{VIS}, and VI-10 Scores in Examinees Who Did or Did Not Utilize an Interpreter

Interpreter	TOMM 1			EI-7 _{VIS}			VI-10		
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>
Yes	26	33.73	12.11	18	8.56	5.53	12	10.00	7.30
No	253	40.70	8.36	231	2.35	2.89	235	3.41	4.02
<i>df</i>			27.51			17.73			11.34
<i>t</i>			2.87 ^a			8.05 ^a			3.10 ^a
<i>p</i>			<.001			<.001			<.001
<i>g</i>			.79			1.98			1.56
95% CI			1.98, 11.96			3.43, 8.97			1.93, 11.24
Mann-Whitney <i>U</i>			2203.50			686.00			526.00
<i>z</i>			2.78			4.82			3.70
<i>p</i>			.006			<.001			<.001

Note. ^a = Levene's test for equality of variance significant, *t* test with equal variances not assumed reported, along with Mann-Whitney *U* Test; *g* = Hedge's *g*; 95% CI = 95% confidence interval; TOMM 1 = Test of Memory Malinger Trial 1 Raw Score (Tombaugh, 1996); EI-7_{VIS} = Erdodi Index Seven – Visuomotor.

Hypothesis 1b

Shorter time in Canada (lived in Canada ≤ 4 years, five to nine years, and ≥ 10 years, born in Canada) would be associated with higher BR_{FAIL} for PVTs with high verbal mediation, but would not be associated with BR_{FAIL} for PVTs with low verbal mediation.

It was necessary to exclude examinees who immigrated to Canada ≤ 4 years ($n = 5$) and between 5 and 9 years prior to assessment ($n = 11$) due to small sample size. Comparison of examinees included and excluded from these analyses are depicted in Table 16, indicating that examinees that were excluded from these analyses performed more poorly on the TOMM, the EI-7_{VER}, and the VI-10. There was also a trend toward a higher proportion of men being excluded from the analyses $\chi^2 (1, N = 303) = 3.44, p = .06$.

Instead of the originally planned one-way ANOVAs, independent t tests were conducted to compare TOMM Trial 1, EI-7_{VER}, EI-7_{VIS}, and VI-10 scores across examinees born in Canada and those who immigrated to Canada ≥ 10 years prior to assessment. Hedge's g is reported as the effect size estimate to control for unequal sample sizes. As is the case with Hypothesis 1a, 2x2 χ^2 analyses were conducted rather than the originally planned 4x2 analyses. Results are presented in Tables 17 and 18.

Assumptions of normality and equality of variance were violated for the t tests with TOMM Trial 1, EI-7_{VIS} and VI-10 as the outcome variables. Mann-Whitney U Tests were thus conducted to confirm the significant results, as violations of assumptions of normality and equality of variance can affect the results in the case of t tests with unequal group sizes, as is the case for the t tests involving the EI-7_{VIS} and VI-10 (Skidmore &

Thompson, 2013; Stevens, 2009). As presented in Table 15, examinees who immigrated to Canada 10 or more years prior to assessment had higher scores on the EI-7_{VER}, EI-7_{VIS}, and VI-10, and lower scores on TOMM Trial 1 as compared to Canadian born examinees. Further, as displayed in Table 16, examinees who immigrated to Canada 10 or more years before assessment were significantly more likely to fail the TOMM, EI-7_{VER}, EI-7_{VIS}, and VI-10 when compared to Canadian born examinees.

Table 16

Independent t Tests on Age, TOMM Trial 1, EI-7_{VER}, EI-7_{VIS}, and VI-10 Scores in Examinees Included or Excluded from Hypothesis 2 Analyses

	Age			TOMM 1			EI-7 _{VER}			EI-7 _{VIS}			VI-10		
Inc	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>
Y	275	44.01	13.64	273	40.68	8.65	238	2.58	3.02	247	2.64	3.51	245	3.49	4.24
N	28	39.96	12.52	27	35.56	10.51	22	3.95	3.33	20	3.65	3.13	22	5.91	5.95
<i>df</i>			301			29.59			258			265			265
<i>t</i>			1.51			2.45 ^a			2.02			1.24			2.47
<i>p</i>			.133			.020			.044			.216			.014
<i>g</i>			.29			.58			.45			.29			.55
CI			-1.24, 9.34			.85, 9.39			.03, 2.71			-.59, 2.60			.49, 4.35

Note. ^a = Levene's test for equality of variance significant, *t* test with equal variances not assumed

reported, Inc = Included in analyses; Y = Yes; N = No; *g* = Hedge's *g*; CI = 95% confidence interval;

TOMM 1 = Test of Memory Malinger Trial 1 Raw Score (Tombaugh, 1996); EI-7_{VER} = Erdodi Index

Seven – Verbal; EI-7_{VIS} = Erdodi Index Seven – Visuomotor; VI-10 = Validity Index Ten.

Table 17

Independent t Tests on TOMM Trial 1, EI-7_{VER}, EI-7_{VIS}, and VI-10 Scores for Examinees by Time in Canada

	TOMM 1			EI-7 _{VER}			EI-7 _{VIS}			VI-10		
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>
B. in Canada	213	42.15	7.36	203	2.49	2.98	199	1.97	2.55	202	2.98	3.63
Imm. ≥ 10 yr.	71	35.72	10.17	50	3.76	3.36	56	4.68	4.20	54	5.74	5.45
<i>df</i>			95.55			251			66.82			67.17
<i>t</i>			4.92 ^a			2.64			4.59 ^a			3.65 ^a
<i>p</i>			<.001			.009			<.001			.001
<i>g</i>			.79			.40			.91			.61
95% CI			3.84, 9.03			.32, 2.22			1.81, 3.59			1.25, 4.28
Mann-Whitney <i>U</i>			4705.50						2981.00			3279.00
<i>z</i>			4.77						5.42			4.55
<i>p</i>			<.001						<.001			<.001

Note. ^a = Levene's test for equality of variance significant, *t* test with equal variances not assumed reported, along with Mann-Whitney *U* Test; Imm. ≥ 10 years = Immigrated to Canada ≥ 10 years prior to assessment; *g* = Hedge's *g*; 95% CI = 95% confidence interval; TOMM 1 = Test of Memory Malinger Trial 1 Raw Score (Tombaugh, 1996); EI-7_{VER} = Erdodi Index Seven – Verbal; EI-7_{VIS} = Erdodi Index Seven – Visuomotor; VI-10 = Validity Index Ten; B. in Canada = Born in Canada; Imm. ≥ 10 yr. = Immigrated to Canada ≥ 10 years before assessment.

Table 18

TOMM, EI-7_{VER}, EI-7_{VIS}, and VI-10 BR_{FAIL} for Examinees by Time in Canada

		TOMM	EI-7 _{VER}	EI-7 _{VIS}	VI-10
	<i>N</i>	286	184	188	198
B. in Canada	<i>n</i>	215	150	147	155
	BR _{FAIL}	23.3	34.0	23.1	38.1
Imm. ≥ 10	<i>n</i>	71	34	41	43
yr.	BR _{FAIL}	50.7	61.8	65.9	76.7
χ^2		19.12	8.97	26.70	20.25
<i>p</i>		<.001	.003	<.001	<.001
Φ^2		.26	.22	.38	.32
RR (95% CI)		2.18 (1.56, 3.05)	1.82 (1.29, 2.57)	2.85 (1.97, 4.11)	2.02 (1.56, 2.61)

Note. TOMM = Test of Memory Malingering (Tombaugh, 1996); EI-7_{VER} = Erdodi Index – Seven Verbal; EI-7_{VIS} = Erdodi Index – Seven Visuomotor; VI-10 = Validity Index – Ten; B. in Canada = Born in Canada; Imm. ≥ 10 yr. = Immigrated to Canada ≥ 10 years prior to assessment; BR_{FAIL} = Base rate of failure (≤ 44 TOMM Trial 2 or Retention, or ≥ 4 on EI-7_{VER}, EI-7_{VIS}, and VI-10); RR = Relative risk ratio; 95% CI = 95% confidence interval.

Hypothesis 1c

Lower educational attainment (defined as education: \leq Grade 8; Grade 8 – 12; graduated high school; completed college; completed an undergraduate degree; completed master's degree; and completed doctoral degree) would be associated with higher BR_{FAIL} for PVTs.

Due to small sample sizes in several of the planned groups, the data were reclassified as <12 years of education, 12 years of education, 13—15 years of education, and ≥ 16 years of education for these analyses. Three one-way ANOVAs were conducted to test for differences in TOMM Trial 1, EI-7_{VER}, EI-7_{VIS}, and VI-10 scores across educational groups. Levene's test was not significant for any of the one-way ANOVAs for this hypothesis, indicating that the assumption equality of variance was met. Four 2x4 χ^2 analyses were conducted comparing TOMM, EI-7_{VER}, EI-7_{VIS}, and VI-10 failure rates across educational groups.

Results of one-way ANOVAs and χ^2 analyses are displayed in Tables 19 and 20, respectively. Results of the one-way ANOVAs indicate that examinees with 11 or fewer years of education had higher scores than those with 16 or more years of education on the EI-7_{VIS} and VI-10, and had higher scores than those with 13 to 15 years of education on the EI-7_{VIS}. These results were mirrored in χ^2 comparisons of BR_{FAIL} . Examination of standardized residuals indicated that examinees with 11 or fewer years of education were significantly more likely to fail the EI-7_{VIS} than other groups. No differences were found for the EI-7_{VER} or TOMM across education groups.

Follow-up 2x2 χ^2 analyses comparing examinees with 11 or fewer years of education to examinees with ≥ 16 years of education on TOMM, EI-7_{VER}, EI-7_{VIS}, and VI-

10 failure rates were conducted, as multiple groups attenuates power of χ^2 analyses (Field, 2009). Results are displayed in Table 21, and indicate that examinees with ≤ 11 years of education were significantly more likely to fail the EI-7_{VER}, EI-7_{VIS}, and TOMM compared to examinees with ≥ 16 years of education but did not differ in TOMM BR_{FAIL}.

Table 19

One-way ANOVAs on TOMM Trial 1, EI-7_{VER}, EI-7_{VIS}, and VI-10 Scores by Education

Level

Education (years)	TOMM 1			EI-7 _{VER}			EI-7 _{VIS}			VI-10		
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>
≤11	70	39.03	10.77	52	3.62	3.65	61	3.90 ^{ab}	3.96	56	4.82 ^c	4.76
12	67	41.31	7.91	59	2.63	3.16	58	2.74	3.26	63	4.56	5.10
13 to 15	108	40.62	7.58	101	2.41	2.53	102	2.31 ^a	3.53	102	3.05	4.07
≥16	55	39.60	9.98	48	2.42	3.18	46	2.02 ^b	2.63	46	2.54 ^c	3.31
<i>df</i>			3			3			3			3
<i>F</i>			.91			2.03			3.50			3.87
<i>p</i>			.437			.110			.016			.010
Partial η^2			.01			.02			.04			.04

Note. Post hoc analyses used Tukey's HSD for control of Type 1 error. TOMM 1 = Test of Memory

Malingering Trial 1 Raw Score (Tombaugh, 1996); EI-7_{VER} = Erdodi Index Seven – Verbal; EI-7_{VIS} =

Erdodi Index Seven – Visuomotor; VI-10 = Validity Index Ten; ^{a,b,c} = Significant post hoc comparisons.

Table 20

TOMM, EI-7_{VER}, EI-7_{VIS}, and VI-10 BR_{FAIL} by Education

		TOMM	EI-7 _{VER}	EI-7 _{VIS}	VI-10
Education (years)	<i>N</i>	302	188	200	207
≤ 11	<i>n</i>	70	36	40	41
	BR _{FAIL}	35.7	55.6	52.5	63.4
	<i>z</i>	.7	1.7	2.1*	1.6
12	<i>n</i>	67	44	46	48
	BR _{FAIL}	20.9	36.4	34.8	54.2
	<i>z</i>	-1.5	-.2	.2	.8
13-15	<i>n</i>	109	72	76	79
	BR _{FAIL}	33.0	36.1	25.0	36.7
	<i>z</i>	.4	-.3	-1.2	-1.3
≥ 16	<i>n</i>	56	36	38	39
	BR _{FAIL}	33.9	27.8	26.3	38.5
	<i>z</i>	.4	-1.0	-.7	-.7
χ^2		4.35	6.44	9.91	9.91
<i>p</i>		.226	.092	.019	.019
Φ^2		.12	.19	.22	.22

Note. TOMM = Test of Memory Malingering (Tombaugh, 1996); EI-7_{VER} = Erdodi Index – Seven Verbal; EI-7_{VIS} = Erdodi Index – Seven Visuomotor; VI-10 = Validity Index – Ten; BR_{FAIL} = Base rate of failure (≤ 44 TOMM Trial 2 or Retention, or ≥ 4 on EI-7_{VER}, EI-7_{VIS}, and VI-10); *z* = standardized residual; * = Significant at *p* < .05.

Table 21

TOMM, EI-7_{VER}, EI-7_{VIS}, and VI-10 BR_{FAIL} for Examinees with Lowest and Highest

Education

		TOMM	EI-7 _{VER}	EI-7 _{VIS}	VI-10
Education (years)	<i>N</i>	126	72	78	80
≤ 11	<i>n</i>	70	36	40	41
	BR _{FAIL}	35.7	55.6	52.5	63.4
≥ 16	<i>n</i>	56	36	38	39
	BR _{FAIL}	33.9	27.8	26.3	38.5
χ^2		.04	5.71	5.58	4.98
<i>p</i>		.835	.017	.018	.026
Φ^2		.02	.28	.27	.25
RR (95% CI)		1.05	2.00	2.00	1.65
		(.65, 1.71)	(1.10, 3.65)	(1.01, 3.67)	(1.04, 2.61)

Note. TOMM = Test of Memory Malingering (Tombaugh, 1996); EI-7_{VER} = Erdodi Index – Seven Verbal; EI-7_{VIS} = Erdodi Index – Seven Visuomotor; VI-10 = Validity Index – Ten; BR_{FAIL} = Base rate of failure (\leq 44 TOMM Trial 2 or Retention, or \geq 4 on EI-7_{VER}, EI-7_{VIS}, and VI-10); RR = Relative risk ratio; 95% CI = 95% confidence interval.

Hypothesis 2

Consistent with previous research (Mittenberg et al., 2012), moderate-to-severe TBI would be associated with lower BR_{FAIL} for PVTs compared to mTBI.

Three examinees were removed from this analysis due to indeterminate TBI severity (i.e., examinee report indicated moderate or severe TBI, but there was no corroborating objective medical documentation). Examinees were initially classified as having mTBI (including both uncomplicated and complicated mTBI) and moderate-to-severe TBI. Three independent samples *t* tests were conducted to compare TOMM Trial 1, EI-7_{VER}, EI-7_{VIS}, and VI-10 scores across examinees with mTBI and moderate-to-severe TBI. Levene's Test was not significant for any of the independent *t* tests involved in this hypothesis, indicating that the assumption of homogeneity of variance was met. Four 2x2 χ^2 analyses were conducted to compare TOMM, EI-7_{VER}, EI-7_{VIS}, and VI-10 BR_{FAIL} between examinees with mTBI and those with moderate-to-severe TBI.

Results of *t* tests are presented in Table 22, and results of χ^2 analyses are presented in Table 23. Contrary to the hypothesis, no significant differences were found between the mTBI and the moderate-to-severe TBI groups on the TOMM, EI-7_{VER}, EI-7_{VIS}, or VI-10.

Exploratory analyses. Examinees were reclassified comparing uncomplicated mTBI to complicated mTBI, moderate TBI, and severe TBI to explore whether differences would be found if complicated mTBI were reclassified with the more severe group. The previously explained analyses were re-run. Results are presented in Tables 24 and 25 and indicated no significant differences between groups on the TOMM, EI-7_{VER}, EI-7_{VIS}, or VI-10.

Table 22

Independent t Tests on TOMM Trial 1, EI-7_{VER}, EI-7_{VIS}, and VI-10 Scores in Examinees with mTBI (Uncomplicated and Complicated) and Moderate-to-Severe TBI

Index	mTBI			M-S TBI			<i>t</i>	<i>df</i>	<i>p</i>	<i>g</i>	95% CI
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>					
TOMM 1	248	39.72	9.11	29	42.69	7.42	1.69	275	.092	.33	-.49, 6.43
EI-7 _{VER}	211	2.62	2.91	26	2.62	3.35	.01	235	.993	.00	-1.21, 1.22
EI-7 _{VIS}	220	2.85	3.52	27	2.67	3.72	.25	245	.800	.05	-1.24, 1.61
VI-10	218	3.77	4.39	27	3.22	4.93	.60	243	.548	.12	-1.24, 2.34

Note. mTBI = Mild traumatic brain injury; M-S TBI = Moderate and severe traumatic brain injury; *g* = Hedge's *g*; TOMM 1 = Test of Memory Malinger Trial 1 Raw Score (Tombaugh, 1996); EI-7_{VER} = Erdodi Index – Seven Verbal; EI-7_{VIS} = Erdodi Index – Seven Visuomotor; VI-10 = Validity Index – Ten.

Table 23

TOMM, EI-7_{VER}, EI-7_{VIS}, and VI-10 BR_{FAIL} by TBI Severity: mTBI (Uncomplicated and Complicated) and Moderate-to-Severe TBI

Injury Severity		TOMM	EI-7 _{VER}	EI-7 _{VIS}	VI-10
	<i>N</i>	278	172	181	189
mTBI	<i>n</i>	249	153	161	164
	BR	32.5	37.9	35.4	49.4
M-S TBI	<i>n</i>	29	19	20	8
	BR	20.7	36.8	30.0	32.0
χ^2		1.69	.01	.23	2.63
<i>p</i>		.193	.928	.632	.105
Φ^2		.08	.01	.04	.12
RR (95% CI)		1.57 (.75, 3.28)	1.03 (.55, 1.92)	1.18 (.59, 2.38)	1.54 (.85, 2.79)

Note. mTBI = Mild traumatic brain injury; M-S TBI = Moderate and severe traumatic brain injury; TOMM = Test of Memory Malingering (Tombaugh, 1996); EI-7_{VER} = Erdodi Index – Seven Verbal; EI-7_{VIS} = Erdodi Index – Seven Visuomotor; VI-10 = Validity Index – Ten; BR = Base rate of failure (≤ 44 TOMM Trial 2 or Retention, or ≥ 4 on EI-7_{VER}, EI-7_{VIS}, and VI-10); RR = Relative risk ratio; 95% CI = 95% Confidence interval.

Table 24

Independent t Tests on TOMM Trial 1, EI-7_{VER}, EI-7_{VIS}, and VI-10 Scores in Examinees with Uncomplicated mTBI Compared to Complicated Mild, Moderate, and Severe TBI

Index	Unc. mTBI			Com. m, M, & S TBI			<i>t</i>	<i>df</i>	<i>p</i>	<i>g</i>	95% CI
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>					
TOMM 1	208	39.85	8.80	69	40.58	9.54	.58	275	.560	.08	-1.73, 3.19
EI-7 _{VER}	179	2.48	2.78	58	3.05	3.44	1.28	235	.201	.19	-.31, 1.45
EI-7 _{VIS}	186	2.76	3.38	61	3.03	4.00	.52	245	.606	.08	-.76, 1.30
VI-10	185	3.71	4.31	60	3.72	4.86	.01	243	.990	.00	-1.29, 1.31

Note. Unc. mTBI = uncomplicated mild traumatic brain injury; Com. m, M, & S TBI = complicated mild, moderate, and severe traumatic brain injury; *g* = Hedge's *g*; TOMM 1 = Test of Memory Malingered Trial 1 Raw Score (Tombaugh, 1996); EI-7_{VER} = Erdodi Index – Seven Verbal; EI-7_{VIS} = Erdodi Index – Seven Visuomotor; VI-10 = Validity Index – Ten.

Table 25

TOMM, EI-7_{VER}, EI-7_{VIS}, and VI-10 BR_{FAIL} by TBI Severity: Uncomplicated mTBI

Compared to Complicated Mild, Moderate, and Severe TBI

Injury Severity		TOMM	EI-7 _{VER}	EI-7 _{VIS}	VI-10
	<i>N</i>	278	172	181	189
Unc. mTBI	<i>n</i>	209	135	136	139
	BR	32.5	36.3	34.6	50.4
Com. m,	<i>n</i>	69	37	45	50
M, & S TBI	BR	27.5	43.2	35.6	38.0
χ^2		.60	.60	.02	2.26
<i>p</i>		.437	.440	.903	.133
Φ^2		.05	.06	.01	.11
RR (95% CI)		1.18 (.78, 1.82)	.84 (.55, 1.29)	.97 (.62, 1.53)	1.33 (.90, 1.96)

Note. Unc. mTBI = Uncomplicated mild traumatic brain injury; Com. M, M, & S TBI = Complicated mild, moderate, and severe traumatic brain injury; TOMM = Test of Memory Malinger (Tombaugh, 1996); EI-7_{VER} = Erdodi Index – Seven Verbal; EI-7_{VIS} = Erdodi Index – Seven Visuomotor; VI-10 = Validity Index – Ten; BR = Base rate of failure (≤ 44 TOMM Trial 2 or Retention, or ≥ 4 on EI-7_{VER}, EI-7_{VIS}, and VI-10); RR = Relative risk ratio; 95% CI = 95% Confidence interval.

Question 1

Would gender be associated with differences in BR_{FAIL} ?

Three independent samples t tests were conducted to compare of TOMM Trial 1, EI-7_{VER}, EI-7_{VIS}, and VI-10 scores between genders, and four 2x2 χ^2 analyses were conducted to compare TOMM, EI-7_{VER}, EI-7_{VIS}, and VI-10 failure rates across genders.

Levene's test was significant for t tests involving the EI-7_{VER} and VI-10. For these analyses, t tests that do not assume equality of variance are reported. Mann-Whitney U tests were not conducted, as sample sizes are roughly equal and t tests are robust to violations of assumptions of normality and equality of variance in this case (Skidmore & Thompson, 2013; Stevens, 2009). Results are presented in Tables 26 and 27 and indicate that there are no significant differences in TOMM, EI-7_{VER}, EI-7_{VIS}, or VI-10 performance across genders.

Table 26

Independent t Tests on TOMM 1, EI-7_{VER}, EI-7_{VIS}, and VI-10 Scores by Gender

	TOMM 1			EI-7 _{VER}			EI-7 _{VIS}			VI-10		
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>
Female	146	40.69	8.61	133	2.46	2.54	132	2.61	3.46	133	3.32	4.04
Male	154	39.77	9.24	127	2.95	3.52	135	2.83	3.53	134	4.05	4.79
<i>df</i>	298			228.37			265			258.27		
<i>t</i>	.90			1.29 ^a			.52			1.35 ^a		
<i>p</i>	.371			.198			.602			.180		
<i>d</i>	.10			.16			.06			.16		
95% CI	-1.11, 2.96			-.26, 1.25			-.62, 1.07			-.34, 1.80		

Note. ^a = Levene's test for equality of variance significant, *t* tests with equal variances not assumed are reported; TOMM 1 = Test of Memory Malinger Trial 1 Raw Score (Tombaugh, 1996); EI-7_{VER} = Erdodi Index – Seven Verbal; EI-7_{VIS} = Erdodi Index – Seven Visuomotor; VI-10 = Validity Index – Ten.

Table 27

TOMM, EI-7_{VER}, EI-7_{VIS}, and VI-10 BR_{FAIL} by Gender

		TOMM	EI-7 _{VER}	EI-7 _{VIS}	VI-10
	<i>N</i>	302	188	200	207
Female	<i>n</i>	148	95	98	97
	BR _{FAIL}	30.4	35.8	31.6	42.3
Male	<i>n</i>	154	93	102	110
	BR _{FAIL}	31.8	40.9	34.3	50.0
χ^2		.07	.51	.16	1.24
<i>p</i>		.791	.475	.687	.266
Φ^2		.02	.05	.03	.08
RR (95% CI)		.96 (.68, 1.34)	.88 (.61, 1.26)	.92 (.62, 1.37)	.85 (.63, 1.14)

Note. TOMM = Test of Memory Malingering (Tombaugh, 1996); EI-7_{VER} = Erdodi Index – Seven Verbal; EI-7_{VIS} = Erdodi Index – Seven Visuomotor; VI-10 = Validity Index – Ten; BR_{FAIL} = Base rate of failure (\leq 44 TOMM Trial 2 or Retention, or \geq 4 on EI-7_{VER}, EI-7_{VIS}, and VI-10); RR = Relative risk ratio; 95% CI = 95% Confidence interval.

Question 2

Is age related to differences in BR_{FAIL} ?

Examinees were divided into five groups by age. Cut-points for the groups were chosen to divide the examinees into groups based on decade of life. Three one-way ANOVAs were conducted to compare TOMM Trial 1, EI-7_{VER}, EI-7_{VIS}, and VI-10 scores across the age groups. Four 2x5 χ^2 analyses were conducted to compare TOMM, EI-7_{VER}, EI-7_{VIS}, and VI-10 BR_{FAIL} across age groups.

Levene's test was significant for the EI-7_{VIS} one-way ANOVA, indicating that the assumption of equality of variance was violated. In this case, group sizes were roughly comparable, and one-way ANOVAs are robust to violations of equality of variance when in this case (Skidmore & Thompson, 2013; Stevens, 2009). As a result, no Kruskal-Wallis tests were conducted. Results of one-way ANOVAs are presented in Table 28, and results of χ^2 analyses are presented in Table 29. Results indicate that examinees age 40 to 49 had significantly higher scores than examinees age 18 to 29 on the EI-7_{VIS}. No other comparisons were significant.

Table 28

One-way ANOVAs on TOMM Trial 1, EI-7_{VER}, EI-7_{VIS}, and VI-10 Scores by Age

Age (years)	TOMM 1			EI-7 _{VER}			EI-7 _{VIS}			VI-10		
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>
18-29	67	40.88	7.80	61	2.70	3.01	62	1.84 ^a	2.73	63	3.06	3.36
30-39	46	39.00	9.76	38	2.63	3.11	39	2.33	3.29	38	3.16	4.70
40-49	67	40.60	8.15	59	3.14	3.54	59	3.68 ^a	4.42	59	4.39	4.98
50-59	80	40.25	9.55	69	2.36	2.65	71	2.48	3.18	72	3.64	4.61
60-69	40	39.80	9.98	33	2.70	3.09	36	3.56	3.36	35	4.31	4.52
<i>df</i>			4			4			4			4
<i>F</i>			.36			.51			2.90 ^b			.99
<i>p</i>			.841			.729			.023			.412
Partial η^2			.01			.01			.04			.02

Note. ^b = Levene's test for equality of variance significant; Post hoc analyses used Tukey's HSD for control of Type 1 error. TOMM 1 = Test of Memory Malingering Trial 1 Raw Score (Tombaugh, 1996); EI-7_{VER} = Erdodi Index – Seven Verbal; EI-7_{VIS} = Erdodi Index – Seven Visuomotor; VI-10 = Validity Index – Ten; ^a = Significant post hoc comparison.

Table 29

TOMM, EI-7_{VER}, EI-7_{VIS}, and VI-10 BR_{FAIL} by Age

Age (years)		TOMM	EI-7 _{VER}	EI-7 _{VIS}	VI-10
	<i>N</i>	302	188	200	207
18-29	<i>n</i>	68	43	54	47
	BR _{FAIL}	29.4	39.5	22.2	40.4
30-39	<i>n</i>	46	23	28	29
	BR _{FAIL}	34.8	34.8	28.6	37.9
40-49	<i>n</i>	67	47	41	51
	BR _{FAIL}	32.8	42.6	41.5	52.9
50-59	<i>n</i>	81	51	50	49
	BR _{FAIL}	30.9	33.3	30.0	44.9
60-69	<i>n</i>	40	24	27	31
	BR _{FAIL}	27.5	41.7	51.9	54.8
	χ^2	.72	1.16	8.96	3.32
	<i>p</i>	.949	.885	.062	.506
	Φ^2	.05	.08	.21	.13

Note. TOMM = Test of Memory Malingering (Tombaugh, 1996); EI-7_{VER} = Erdodi Index – Seven Verbal; EI-7_{VIS} = Erdodi Index – Seven Visuomotor; VI-10 = Validity Index – Ten; BR_{FAIL} = Base rate of failure (\leq 44 TOMM Trial 2 or Retention, or \geq 4 on EI-7_{VER}, EI-7_{VIS}, and VI-10).

CHAPTER 8

Objective 1 Discussion

The current objective explored the effects of demographic, cultural, and linguistic variables on PVT performance. Specifically, the effects of limited English proficiency, time in Canada, education level, TBI severity, gender, and age were explored regarding the TOMM and three composite embedded validity indicator measures, the EI-7_{VER}, EI-7_{VIS}, and VI-10.

Limited English Proficiency and BR_{FAIL}

The findings regarding the effects of limited English proficiency on BR_{FAIL} were both surprising and counterintuitive. It was expected that examinees with limited English proficiency would have higher BR_{FAIL} on embedded validity indicators with high verbal mediation and perform as well as Anglophone Canadian examinees on the visuospatial and motor tasks involved in the EI-7_{VIS}. Instead, the opposite pattern was found. This is especially surprising considering previous findings that examinees with limited English proficiency struggle with verbally mediated neuropsychological tests in general (Poreh, Avital, Dines, & Levin, 2015), as well as embedded validity indicators with high verbal mediation specifically (Erdodi, Nussbaum, Sagar, Abeare, & Schwartz, 2017). EI-7_{VER} failure rates for the limited English proficiency group in the current study were comparable to those of the Canadians with English as a first language, and to verbally mediated BR_{FAIL} in a recent study (Erdodi, Nussbaum, et al., 2017) in which healthy Arabic-English bilinguals were tested in their dominant and nondominant language on a variety of PVTs and embedded validity indicators.

In contrast, in the current study, the EI-7_{VIS} and TOMM BR_{FAILS} for the limited English proficiency group were significantly higher than those of the Anglophonic Canadian group, with large effects for continuous EI-7_{VIS} scores and BR_{FAILS}. The limited English proficiency group in the current study scored in the expected range for people with limited English proficiency *with no motivation to appear impaired* on the EI-7_{VER}, i.e., similarly to the healthy Arabic-English bilinguals in the previous study (Erdodi, Nussbaum, et al., 2017). In other words, it appears that the limited English proficiency group scores on the EI-7_{VER} represent something close to credible best performance for people with limited English proficiency, whereas the comparable scores of the Canadian-born Anglophone group seem to indicate noncredible performance. The EI-7_{VIS} scores of both groups appear to indicate noncredible performance.

Several factors may partially account for this seemingly counterintuitive finding. The first is possible sampling bias. The data were collected from clinical assessments in which the battery was somewhat flexible, which resulted in systematically missing data in the current study. Examinees were able to refuse tests, and some may have systematically refused tests that they found particularly challenging. Although these difficulties are common to clinical settings, they may be even more prominent in this forensic setting, where some examinees may refuse tests to demonstrate impairment (e.g., telling the examiner that they cannot complete certain tests because the tests exacerbate their headaches). The neuropsychologist may also have had unstated biases that may have affected tests selection, or referral types.

The assessing neuropsychologist also decided to abbreviate test batteries for a variety of reasons. For example, if examinees had a very limited English proficiency, he

might remove many of the measures with high verbal mediation. Unfortunately, it is not possible to test this hypothesis, because the reasons for missing tests were not reliably documented for each case. The examinees with missing EI-7_{VER} data may have had higher EI-7_{VER} scores had they been administered all components. It is possible that some of these selection biases may have altered the distribution of the EI-7_{VER} for the limited English proficiency group such that these examinees had no EI-7_{VER} score, but had calculable and elevated EI-7_{VIS} scores.

A related, albeit potentially smaller factor, is instrumentation bias in Vocabulary minus Digit Span, an embedded validity indicator with particularly low BR_{FAIL} in the current study. This embedded validity indicator may favour individuals with limited English proficiency (as well as examinees of any linguistic background who suppress their performance throughout the examination). As a difference score, the logic of the Vocabulary minus Digit Span rests on the assumption that simulators, and by extension noncredible examinees, perform disproportionately worse on Digit Span compared to other intellectual functions (Mittenberg et al., 1995). Vocabulary was originally chosen as the comparison to Digit Span as Vocabulary scores closely relate to global intellectual function (Mittenberg et al., 1995). Immediate attention (i.e., Digit Span) performance remains intact relative to overall intelligence in concussion examinees (Mittenberg et al., 1995). Thus a large discrepancy between Digit Span and Vocabulary scores is indicative of noncredible performance. The higher the Vocabulary age-corrected scaled score is compared to the Digit Span age-corrected scaled score, the greater the confidence in noncredible performance. This paradigm has since been utilized successfully by several

researchers (Curtis, Greve, & Bianchini, 2009; Greve et al., 2003; Iverson & Tulskey, 2003; Schwarz, Gfeller, & Oliveri, 2006).

Despite the utility of Vocabulary minus Digit Span in research with examinees who are Anglophonic, individuals with limited English proficiency would be expected to perform more poorly on both Vocabulary (Poreh et al., 2015) and Digit Span (Erdodi, Nussbaum, et al., 2017). Vocabulary is more verbally complex, and therefore may be a conceptually *more* difficult task than Digit Span for individuals with limited English proficiency. This might lead to lower Vocabulary scores compared to Digit Span scores, which is the opposite discrepancy direction compared to the EVI. Both Vocabulary and Digit Span scores would also likely be lower in examinees with limited English proficiency, limiting the opportunity for discrepancy. One might therefore rationally expect lower rates of failure on this embedded validity indicator for examinees with limited English proficiency as the difference score would likely be low regardless of credible performance.

Cultural concepts of distress. An alternative and perhaps more appealing explanation for the EI-7_{VIS} difference is the possibility of an underlying factor increasing the limited English proficiency group EI-7_{VIS} BR_{FAIL}. It is possible that the motoric component of the tests had a differential effect for the limited English proficiency group compared to the Canadian-born Anglophonic group. Specifically, six of the seven EI-7_{VIS} embedded validity indicators require the examinee to interact with the test in some way that involves movement. In contrast, none of the EI-7_{VER} components involve motoric components beyond speaking and reading. It is possible that the differentially high BR_{FAIL} of these “motorically mediated” embedded validity indicators are expressions of

distress for many examinees with limited English proficiency. For example, examinees with limited English proficiency may not perceive repeating strings of numbers (Digit Span) as being connected to impairment and distress, but perceive tapping their finger quickly (Finger Tapping Test) as being closely connected to impairment and distress.

This explanation closely relates to cultural concepts of distress. Cultural concepts of distress are means of expressing distress that provide explanations of distress that are more socially acceptable and understandable within the examinee's culture (American Psychiatric Association, 2013). These symptoms are usually physical and serve as a means for the examinee to express complex cultural and social concerns (Kirmayer, Groleau, Looper, & Dao, 2004). For example, examinees who immigrated to Canada may have lost their previous source of income following the motor vehicle accident and may have had to become more dependent on their family financially and for activities of daily living. They may also not fully understand the nature of their injuries, treatments they receive, or instructions from their health care providers and lawyer. An examinee who immigrated to Canada may express these concerns through physical symptoms like motor slowing, headaches, back pain, and stomach upset. Conversely, a Canadian-born examinee may express similar concerns through emotional symptoms such as sadness and frustration or demonstrating disruptions in cognitive functioning (e.g., attention, memory).

Cultural concepts of distress are not synonymous with somatization (Kirmayer & Sartorius, 2008). Chiefly, individuals expressing concepts of distress may be aware of the social antecedents of the symptoms. They may be willing to explore these antecedents in a safe and supportive environment, but generally will not accept an intrapsychic

explanation for their symptoms (i.e., an explanation of somatization), as they may see these explanations as morally judgmental and stigmatizing (Kirmayer & Sartorius, 2008). Cultural concepts of distress also do not usually indicate psychopathology (Groleau & Kirmayer, 2004). These expressions of distress are not conscious efforts at deception (Young, 2008), but rather are culturally determined means of expressing existing suffering.

Applying this information to the current findings, examinees with limited English proficiency may have expressed cultural concepts of distress that involve physical symptoms (e.g., pain, motor retardation) that would interfere with motorically mediated neuropsychological tests, but would not affect verbally mediated tests. The tasks involved in verbally mediated measures, such as remembering and producing words, may not fall into the scope of socially relevant or acceptable expressions of distress for examinees with limited English proficiency, and as such, these examinees may not have expressed their distress in ways that were measured by the EI-7_{VER} embedded validity indicators.

It is also possible that examinees with limited English proficiency may have been less likely to grasp the nuances of test instructions compared to their Canadian-born Anglophonic peers. For example, on tests that require a focus on speed of performance (e.g., Coding, Symbol Search, Finger Tapping), examinees with limited English proficiency may have at times focused more on the accuracy of responses than the speed of output, leading to lower scores, and by extension higher PVT BR_{FAIL}.

In a related vein, it is also possible that the examinees with limited English proficiency performed comparatively well on the verbal subtests in a counterintuitive response to stereotype threat. Stereotype threat or diagnosis threat is the process in which

examinees perform worse or “choke” on tests when they are confronted with poor expectations of their subgroup (Silver, 2015). These threats affect examinees who are members of minority groups (Thames et al., 2013) and examinees who have had a TBI (Silver, 2015). It is possible that the examinees with limited English proficiency were very cognizant of TBI stereotypes and cultural concepts of distress. Although untested, this may have led to a reduced focus on or awareness of stereotypes about people with limited English proficiency during the examination. Examinees with limited English proficiency may, therefore, have performed comparatively well on tests with high verbal mediation because cultural stereotype threat had been overridden by TBI diagnosis threat and cultural concepts of distress.

Impact of interpreters. Another explanation that may occur to readers is that the use of interpreters may have improved the apparent performance of examinees with limited English proficiency on the EI-7_{VER} and not the EI-7_{VIS}. Previous research has indicated that the use of interpreters can improve verbally mediated test scores, and can increase variability in these test scores when compared to measures with lower verbal mediation (Casas et al., 2012). Although this is possible, only three examinees with interpreters had EI-7_{VER} data and were included in these analyses. Removal of these examinees from the EI-7_{VER} analyses did not alter the results appreciably.

Contrary to this explanation, examinees who had interpreters performed significantly more poorly on TOMM Trial 1, EI-7_{VIS}, and VI-10 when compared to examinees without interpreters, with medium to very large effect sizes. There are several possible explanations for these findings. Firstly, some of the interpreters might not have explained instructions effectively, which may have led to some confusion about

important technical details on the tests, resulting in poorer performance. For example, the interpreters may not have emphasized to the examinees that they should complete as many items as quickly as possible on Coding or Symbol Search. This explanation may be supported both by previous research that finds greater variability in test scores when utilizing interpreters (Casas et al., 2012), and the larger *SDs* for the group that had interpreters compared to those who did not for TOMM Trial 1, EI-7_{VIS}, and VI-10.

These exploratory findings may also support the explanation that cultural concepts of distress drive the differences in PVT performance between examinees with limited English proficiency and examinees who are Canadian-born. Acculturation is related to English language proficiency (Jia, Gottardo, Chen, Koh, & Pasquarella, 2016; Riccio, Yoon, & McCormick, 2014). The use of languages other than English for interview (i.e., low English language proficiency) has been found to be a better proxy of acculturation than other demographic factors (Lee, Nguyen, & Tsui, 2011).

Lower English language proficiency also predicts marginalisation—i.e., low identification with the person's own culture and the host culture (Shafaei, Abd Razak, & Nejati, 2016). Although acculturation was not measured directly in this study, it is likely that examinees with interpreters were the least acculturated to Canadian society in the sample. It follows that they may be the most likely to enact cultural concepts of distress that lead to motorically mediated PVT failure.

If this is the case, it is possible that examinees with the lowest acculturation (i.e., those requiring interpreters) did not complete enough of the composite measures for the EI-7_{VER} to be calculated. This may have led to selection bias for this analysis. Given that EI-7_{VER} values with more than two missing components were automatically excluded, the

inclusion criteria may have disproportionately affected examinees with the lowest levels of acculturation, ultimately contributing to the null findings for limited English proficiency for the EI-7_{VER}.

Time in Canada and BR_{FAIL}

In contrast to the findings comparing examinees with limited English proficiency and their Anglophone Canadian counterparts, the findings were somewhat different when examinees who immigrated to Canada were compared to examinees who were Canadian-born regardless of first language. Results of these analyses showed that the group who immigrated to Canada had significantly higher BR_{FAIL} on the TOMM, EI-7_{VER}, EI-7_{VIS}, and VI-10.

One consideration for the higher EI-7_{VER} scores in this analysis is the inclusion of examinees who immigrated to Canada ($n = 21$) from countries that are primarily Anglophonic, such as Jamaica ($n = 8$) and Guyana ($n = 4$). The dialects that these examinees speak and the education systems they encountered may have led to difficulties in understanding instructions and completing verbally mediated tests with examiners using Canadian English and Canadian test norms. Strategies such as supplementing test instructions to make testing more understandable may have been used less with these examinees when compared to examinees with limited English proficiency. This may have led to higher BR_{FAIL}. Regarding the differences on the EI-7_{VIS} and TOMM, the previous discussion of cultural concepts of distress would likely generalize similarly to these analyses that included examinees who are French Canadian and Anglophonic immigrants.

The results of these analyses highlight that living in Canada for at least 10 years does not close the gap in PVT performance between examinees who are Canadian-born and immigrants—even those who lived in Canada for a long time. Previous research has often demonstrated greater acculturation to the host culture as time in the host country increases (Chudek, Cheung, & Heine, 2015). However, some studies showed this pattern only for participants who immigrated before age 15 (e.g., Cheung et al., 2010). It is possible that the current research has exposed a particularly vulnerable subset of people who have immigrated to Canada—people who have not acculturated to the host culture and are having difficulty navigating the Canadian health care system after they sustain injuries in an accident.

Health literacy. Another consideration that may contribute to the differences found between Canadian-born examinees and examinees who immigrated to Canada may be differences in health literacy. Health literacy is the ability of a person to seek out, comprehend, and communicate about health services and specific information (Aldoori, 2017). It encompasses traditional concepts of literacy and numeracy for health information, such as being able to read and understand a pamphlet that explains risks for a particular illness. It also includes broader domains that can affect healthcare engagement, such as self-efficacy and knowledge and beliefs about health (Ishikawa & Kiuchi, 2010).

Lower functional health literacy in the host country is related to poorer health (Mantwill & Schulz, 2016). Low health literacy also interferes with health care access and the ability to navigate the healthcare system (Yun et al., 2015). It is conceivable although untested that the PVT BR_{FAIL} in the group who immigrated to Canada represents

ineffective attempts to interface with the Canadian health care system. Examinees who are less familiar with the Canadian health care system may perceive it necessary to demonstrate their impairment emphatically to the assessor to secure benefits. They may not be aware that this behaviour is likely to result in their claims being dismissed as noncredible in the Canadian health care system.

To address noncredible performance that results from poor health literacy, it may help to use a more comprehensive informed consent process that includes explicit discussion of performance validity. Carone, Iverson, and Bush (2010) advocate for this approach with all examinees, both in clinical and forensic assessment contexts. They suggest that the informed consent process should include an explicit explanation that the purpose of the assessment is not to advocate for or against the examinee but to understand their neuropsychological functioning. They further suggest that the examiner explain that symptoms exaggeration and poor test engagement will be assessed as well and that noncredible performance can negatively impact financial or other claims.

It should be noted that this approach is controversial, and several other researchers strongly express that neuropsychologists should not warn examinees about PVTs and symptom validity tests (Boone, 2007; Youngjohn, Lees-Haley, & Binder, 1999). These authors are concerned that warning examinees about PVTs and symptom validity tests will lead to more sophisticated and effective dissimulation tactics rather than full and honest engagement in testing. There is, however, a consensus that examinees should never be informed about specific PVTs—e.g., that the test that they are about to complete is a PVT (Boone, 2007; Gervais, Green, Allen, & Iverson, 2001; Iverson, 2006).

Despite the above-noted controversy, warning examinees with low health literacy may be a solution to the difficulties that these individuals seem to face in post-injury assessment. Boone (2007) suggested a clause to include in consent forms that acknowledges that the examinee understands that exaggeration may “make my test profile more problematic to interpret” (p. 43). Examinees with high health literacy would then likely understand that it is in their best interest to engage fully in testing. For examinees with low functional Canadian health literacy, it may be necessary to have a more explicit conversation about the potential negative effects of exaggeration.

This is especially important considering that the Canadian Psychological Association Code of Ethics instructs psychologists to ensure that examinees understand their responsibilities, the risks and benefits of the assessment, the consequences of nonaction, and to “take whatever reasonable steps are needed to ensure that the information was, in fact, understood” (p. 11, Canadian Psychological Association, 2010). A more explicit explanation of the risks of exaggeration may provide examinees with low functional Canadian health literacy an equivalent amount of context for informed consent as a Canadian-born examinee would *already have* due to higher acculturation. Future research could examine this supposition through simulation designs in which examinees who are Canadian-born and those who immigrated to Canada would be randomly assigned to warning and nonwarning conditions before PVT testing.

Alternative explanations. It is possible that the effects of cultural concepts of distress and low health literacy cannot be controlled in PVT testing through more thorough consent processes or the development of normative corrections. Consider an analogy about the difficulty in interpreting CT-scans of the base of the brain. The base of

the brain rests on a thick and deeply ridged plate of bone. The bone is radiodense, leading to bright artefacts that obscure the brain tissue on CT images. No amount of signal manipulation allows the brain tissue to be well distinguished from the artefacts. As a result, CT-scans of the base of the brain are usually not helpful. Instead, the radiologist must use different imaging techniques such as MRI. MRI is not universally superior to CT, but it circumvents the problem of imaging the base of the brain because it does not detect bone well.

Similarly, it is possible that the artefacts of culture, which may include limited English proficiency, cultural concepts of distress and low functional Canadian health literacy, are so strong that they obscure the signal that PVTs attempt to detect—performance validity. If this is the case, researchers and clinicians may have to rethink how to measure performance credibility in these populations. New measures would have to be relatively insensitive to limited English proficiency, health literacy and cultural concepts of distress to detect performance credibility in examinees who have immigrated to Canada.

A final possible interpretation of the findings of higher BR_{FAIL} in examinees who immigrated to Canada is that there is a much higher proportion of malingering in this group compared to the Canadian-born group. There may be grave implications of this interpretation, which might include prejudice and discrimination against people who have immigrated to Canada and reduced access to health care and other benefits for examinees who are immigrants. There are also multiple factors, outlined above, that may account for the observed differences, and no independent evidence aside from the BR_{FAILS} indicates that malingering is the best explanation for the phenomenon. This interpretation,

therefore, should be considered only when other avenues of interpretation have been ruled out.

Education and BR_{FAIL}

There were no differences in TOMM BR_{FAIL} across education groups, consistent with previous research (Gervais, Rohling, Green, & Ford, 2004; Strauss et al., 2006). $EI-7_{VIS}$ scores and BR_{FAIL} were higher for examinees with less than high school education. Conversely, there were no overall differences in $EI-7_{VER}$. Follow-up comparison of those with ≤ 11 years of education with examinees who had had ≥ 16 years of education revealed higher $EI-7_{VER}$ BR_{FAIL} in the lower education group. One partial explanation for the differences is that only three of the seven $EI-7_{VER}$ embedded validity indicators were education corrected, whereas only one of seven $EI-7_{VIS}$ embedded validity indicators was education corrected. These corrections are designed to account for the effects of education on cognitive performance and may have had a differential effect on the $EI-7_{VIS}$ as compared to the $EI-7_{VER}$ (Lam et al., 2013) due to the number of education-corrected cutoffs in their components.

TBI severity and BR_{FAIL}

Contrary to the hypothesis, TBI severity had no significant relationships with EI scores or BR_{FAIL} of the TOMM or EIs although there was a nonsignificant trend toward examinees with moderate-to-severe TBI having better TOMM Trial 1 scores than examinees with mTBI. Previous research has been inconsistent about the relationship between TBI severity and PVT failure. Many studies found higher BR_{FAIL} for examinees with mTBI compared to those with moderate-to-severe TBI (Carone, 2008; Green et al., 1999,2001; Mittenberg et al., 2002; Sherer et al., 2015; Webb et al., 2012; West et al.,

2011). Some studies have shown similar BR_{FAIL} across TBI severity for some PVTs and/or cutoffs (Arnold et al., 2005; Curtis et al., 2008; Guise et al., 2014; Hampson, Kemp, Coughlan, Moulin, & Bhakta, 2014) and in some cases higher BR_{FAIL} are found for those with severe TBI compared to mTBI on embedded validity indicators (Erdodi, Abeare, et al., 2017).

The use of embedded validity indicators as opposed to stand-alone PVTs may have contributed to the null findings for this hypothesis. Embedded validity indicators are typically derived from normally distributed measures that are designed to be sensitive to impairment (Erdodi & Lichtenstein, 2017). This can lead to higher false-positive error with examinees who have severe neuropathology when using more liberal embedded validity indicator cutoffs designed for use with mTBI, i.e., misclassification of impairment as noncredible performance (Curtis et al., 2006).

That said, it is somewhat unusual that 20.7% of the moderate-to-severe TBI group failed the TOMM, as this instrument typically has low (sometimes 0%) BR_{FAIL} in participants with moderate-to-severe TBI, especially at the standard cutoffs used in the current study (Tombaugh, 1996). It is conceivable that demand characteristics in this study (i.e., being compensation-seeking claimants) contributed to the null findings regarding TBI severity. Relatedly, it is possible that examinees with moderate-to-severe TBI referred by insurance companies for independent medical examinations differ substantially from the majority of individuals with similar injuries in terms of base rate of feigned impairment or somatic symptom disorder. In other words, the same factors that contributed to noncredible performance in examinees with mTBI also contributed to noncredible performance in examinees with moderate-to-severe TBI in this sample.

Gender and BR_{FAIL}

With regard to gender, the null results were expected and consistent with previous research (Constantinou & McCaffrey, 2003; Donders, 2005; Rees et al., 1998; Webb et al., 2012). Gender adjusted norms for FAS, Animals, Trail Making Test-A, and gender-adjusted cutoffs for the Finger Tapping Test (Arnold et al., 2005; Axelrod et al., 2014), which account for the expected differences in raw scores across gender on these tasks, likely aided these findings.

Age and BR_{FAIL}

The only significant finding about age-related differences was in EI-7_{VIS} scores. This finding was somewhat unexpected and difficult to understand. If there were significant findings, one would expect that performance would decrease across the age groups, consistent with typical cognitive declines that happen over the lifespan (Strauss et al., 2006). These differences, however, are controlled for by age correction in three of the seven EI-7_{VIS} subtests, and four of the seven EI-7_{VER} subtests. The only significant pairwise comparison found in the age-related analyses was lower EI-7_{VIS} scores in the 18- to the 29-year-old group when compared to the 40- to 49-year-old group. In general, there was no obvious pattern of scores or BR_{FAIL} across the lifespan for any of the measures examined, including the EI-7_{VIS}. It is not clear what might have contributed to the one significant comparison, but the effect size was small ($\eta^2 = .04$). This finding may not be replicable.

CHAPTER 9

Objective 2: Psychiatric Symptoms and Performance Validity

Method

The relationship between BR_{FAIL} and depression, anxiety, PTSD, and dissociation were explored in this objective. Depression, anxiety, PTSD, and dissociative symptom severity were to be categorized according to the Trauma Symptom Inventory II Alternate manual cutoffs (TSI-II-A; Briere, 2011). It was expected that higher levels of PTSD and dissociation would be associated with higher BR_{FAIL} and that depression and anxiety symptoms would not be related to BR_{FAIL} .

Research Questions

1. Would self-reported depression symptoms be associated with higher BR_{FAIL} ?
2. Would self-reported anxiety symptoms be associated with higher BR_{FAIL} ?
3. Would self-reported PTSD symptoms be associated with higher BR_{FAIL} ?
4. Would self-reported dissociative symptoms be associated with higher BR_{FAIL} ?

Participants

Participants are described in the General Methods section of this document.

Measures

Performance validity tests. Specific information about the PVTs included in this study are described in the General Methods section of this document.

Symptom validity tests. Symptom validity tests measure noncredible symptom reporting (Morey, 1991). These measures are designed to detect patterns of symptom

endorsement that are rare in general and clinical populations (Strauss et al., 2006). The results of self-report questionnaires should not be interpreted when symptom validity tests exceed cutoffs provided in test manuals (Briere, 2011; Morey, 1991). This section introduces the symptom validity tests for this study and the inventories in which they are embedded. The clinical scales for the study will be explained in the next section.

Personality Assessment Inventory (PAI; Morey, 1991). The PAI is a 344-item self-report inventory with Likert-scale responses ranging from 1 (*false*) to 4 (*very true*). It measures multiple facets of personality and psychopathology, as well as symptom validity.

Negative Impression Management is a 9-item scale designed to detect exaggerated negative responding with low endorsement rates in clinical examinees. A *T*-score between 73 and 91 indicates some level of magnification of symptoms. A *T*-score \geq 92 indicates noncredible responding. Cronbach's alpha for a US census matched sample ($N = 1000$) was reported at .72, and for a clinical sample ($N = 1246$) was reported at .77 (Morey, 1991). Test-retest reliability (mean interval 24 days) in a community ($n = 75$) and college ($n = 80$) combined sample ($N = 155$) was $r = .75$.

Positive Impression Management is a 9-item scale designed to detect strongly favorable impression management or denial of common flaws. A *T*-score of 57 to 67 suggests responding with some denial of flaws, and a *T*-score \geq 68 represents noncredible responding. Cronbach's alpha was .71 for the census matched group and .77 for the clinical sample. Test-retest reliability in the combined sample was $r = .78$ (Morey, 1991).

Infrequency is an 8-item scale comprised of very unusual items that are rarely endorsed and unrelated to psychopathology. A *T*-score score between 60 and 67 indicates

idiosyncratic responding, and a T -score ≥ 74 indicates possible reading difficulties, random responding, confusion, or carelessness, and is indicative of noncredible responding. Cronbach's alpha was .45 for the census matched group and .23 for the clinical group. It should be noted that these unusual items are conceptually unrelated to one another, which may have contributed to the low inter-item consistency. Test-retest reliability in the combined sample was $r = .48$ (Morey, 1991).

Finally, Inconsistency is a scale based on 10 pairs of matched items used in the evaluation of the consistency of responses to items with very high correlations. A T -score from 64 to 72 suggests some level of inconsistent responding. A T -score ≥ 73 indicates inconsistent responses that suggest noncredible responding due to carelessness, reading difficulties, or confusion. Cronbach's alpha was .52 for the census matched group and .40 for the clinical group. Like the Infrequency scale, the pairs of items are similar within pairs but are unrelated between pairs, which may have contributed to low scale consistency. Test-retest reliability in the combined sample was $r = .31$ (Morey, 1991).

Trauma Symptom Inventory Second Edition Alternate (TSI-2-A; Briere, 2011).

The TSI-II-A is a 126-item self-report measure in which the examinee endorses the experience of trauma-related symptoms over the past six months on a scale from 0 (*never*) to 3 (*often*).

The Atypical Response subscale was designed to detect over-endorsed PTSD symptomatology (Gray, Elhai, & Briere, 2010). A raw score ≥ 15 on Atypical Response is representative of noncredible responding in clinical and forensic contexts (Briere, 2011). Cronbach's alpha in the standardization sample ($N = 678$) was .72, and test-retest coefficient ($N = 31$, mean interval one week) was $r = .66$ (Briere, 2011).

The Response Level scale is comprised of eight items that are unlikely to receive a score of zero in community or clinical contexts (Briere, 2011). A high score is representative of defensiveness or unwillingness to endorse items, and T -score > 75 is representative of invalid responding (Briere, 2011). Cronbach's alpha in the standardization sample was .81, and the test-retest coefficient was $r = .89$ (Briere, 2011).

Clinical scales.

Trauma Symptom Inventory Second Edition Alternate (TSI-2-A; Briere, 2011) Dissociation subscale. This scale measures self-reported dissociative symptomatology, including alterations in awareness, cognitive disengagement, depersonalization and derealization, and multiple personality experiences (as measured by one item in the scale; Briere, 2011). Examinees rate symptom frequency on a scale from 0 (*never*) to 3 (*often*) over the past month. T -scores from 60 to 64 indicate "problematic" levels of symptomatology, and T -scores ≥ 65 indicate "clinically elevated" symptomatology. Cronbach's alpha in the standardization sample was .86, and the test-retest coefficient was $r = .87$ (Briere, 2011).

Beck Depression Inventory Second Edition (BDI-II; Beck et al., 1996). The BDI-II is a 21-item self-report scale measuring cognitive, affective, and behavioural depression symptoms with item descriptions corresponding to endorsement levels ranging from zero to three, where three represents the most severe experience of the symptom. Raw scores ≤ 13 are considered to represent minimal probability of representing major depressive disorder, scores from 14 to 19 represent mild probability of representing major depressive disorder, scores from 20 to 28 represent moderate probability of major depressive disorder, and scores ≥ 29 represent severe probability of major depressive disorder.

Cronbach's alpha for 500 outpatients was .92, and test-retest reliability for 26 outpatients (mean interval one week) was $r = .93$ (Beck et al., 1996). Chronbach's alpha for the current data was .95.

Trauma Symptom Inventory Second Edition Alternate (TSI-2-A; Briere, 2011)
Depression subscale. This scale measures depressed mood and cognition (Briere, 2011). It does not include items that query about suicidality or self-harm behaviours, which are measured by the Suicidality and Tension Reduction Behaviour subscales, respectively (Briere, 2011). *T*-scores from 60 to 64 indicate "problematic" levels of symptomatology, and *T*-scores ≥ 65 indicate "clinically elevated" symptomatology. Cronbach's alpha in the standardization sample was .94, and the test-retest coefficient was $r = .94$ (Briere, 2011).

Beck Anxiety Inventory (BAI; Beck & Steer, 1993). This is a 21-item self-report measure of anxiety symptoms where each item is rated on a 4-point Likert-type scale from zero to four where four represents highest symptom severity over the past week. The items include somatic, affective, and cognitive symptoms. Raw scores ≤ 7 reflect minimal levels of anxiety, scores from 8-15 suggest mild anxiety, scores from 16-25 represent moderate anxiety, and scores ≥ 26 represent severe anxiety. Cronbach's alpha for 160 outpatients was .92, and test-retest reliability ($N = 83$, one week interval) was $r = .75$ (Beck & Steer, 1993).

Trauma Symptom Inventory Second Edition Alternate (TSI-2-A; Briere, 2011)
Anxious Arousal subscale. This scale measures symptoms of anxiety, including fear, panic, physiological symptoms, and phobias (Briere, 2011). These symptoms can be present in people who have been exposed to trauma but are not specific to trauma-related

disorders (Briere, 2011). *T*-scores from 60 to 64 indicate “problematic” levels of symptomatology, and *T*-scores ≥ 65 indicate “clinically elevated” symptomatology. Cronbach’s alpha in the standardization sample was .89, and the test-retest coefficient was $r = .87$ (Briere, 2011).

Trauma Symptom Inventory Second Edition Alternate (TSI-2-A; Briere, 2011)

Posttraumatic Stress factor. This factor consists of the following scales: Intrusive Experiences, Defensive Avoidance, Anxious Arousal, and Dissociation. The factor represents elevated symptoms of flashbacks, nightmares, intrusive memories, avoidance of traumatic events, hyperarousal, and dissociative symptoms. *T*-scores from 60-64 indicate “problematic” levels of symptomatology, and *T*-scores ≥ 65 indicated “clinically elevated” symptomatology. Cronbach’s alpha in the standardization sample was .93, and the test-retest coefficient was .93 (Briere, 2011).

Personality Assessment Inventory (PAI; Morey, 1991) Anxiety Related Disorders

Traumatic Stress. This subscale measures specific fears and distress that result from past traumatic events (Morey, 1991). Examinees endorse statements as *false, somewhat true, mostly true, or very true*. *T*-scores from 60-69 indicate that the examinee has some fears or worries. *T*-scores between 70 and 90 indicate that the examinee has impairment associated with these fears, and *T*-scores above 91 indicate wider ranging impairment and severe psychological suffering resulting from these fears. Cronbach's alpha in the standardization sample was .89, and test-retest coefficient was .82 (Morey, 1991).

CHAPTER 10

Results Objective 2

Initial Analyses

One-way ANOVAs were conducted to compare BDI-II and BAI scores across examinees who passed the TSI-II-A symptom validity tests, those who failed at least one of the TSI-II-A symptom validity tests, and those who did not have TSI-II-A data. These analyses were conducted to determine whether the TSI-II-A missing group could be considered a “healthy” group for subsequent analyses.

The assumption of normality of the BDI-II and BAI data was tested, and results are presented in Table 30. Shapiro-Wilk’s test was significant ($p < .001$ for both measures), indicating that neither variable was normally distributed. Visual inspection of the data indicated that BDI-II scores were bimodally distributed, and BAI scores were both bimodally distributed and positively skewed. Levene’s test was significant for both BDI-II ($p < .001$) and BAI ($p < .001$) for the one-way ANOVAs, indicating heteroscedasticity in the data. One-way ANOVAs are not robust to violations of assumptions of normality and equality of variance when sample sizes are not equal, as in this case (Skidmore & Thompson, 2013; Stevens, 2009). Kruskal-Wallis tests were therefore conducted to confirm findings through nonparametric testing. One-way ANOVA and Kruskal-Wallis results are presented in Table 31.

Post hoc analyses were conducted using Tukey’s HSD to control for Type-1 error, revealing that the TSI missing group had significantly lower scores than the TSI valid group for both BDI-II and BAI. Despite the differences in both BDI-II and BAI means scores across groups, it was decided that the TSI Missing group could not constitute a

“healthy” group, as their BDI-II and BAI mean scores were both in the moderate range (Beck et al., 1996; Beck & Steer, 1993).

Two 2x3 χ^2 analyses were conducted to compare TOMM completion and TOMM failure across examinees who passed, failed, or were missing TSI-II-A validity data. TOMM completion rates did not differ across TSI-II-A Pass, Fail, and Missing groups $\chi^2(2, N = 303) = 3.67, p = .160$. TOMM failure rates also did not differ across these groups $\chi^2(2, N = 302) = .86, p = .652$. Because cell sizes for TSI-II-A Fail group were expected to be smaller than five, Fisher’s Exact Test was calculated, yielding $p = .197$ for TOMM administration rate and $p = .676$ for TOMM failure rate.

Another challenge in using TSI-II-A data for this study is the large amount of missing TSI-II-A data (51.2% of examinees). As a result of the large amount of missing data, and the use of the missing data group as a “healthy” comparison group being untenable, it was decided that the BDI-II would be used for the analyses regarding depression ($N = 279$), the BAI would be used for the analyses regarding anxiety ($N = 285$), and PAI Anxiety Related Disorders Traumatic Stress Scale would be used for the analyses regarding PTSD ($N = 239$). It was decided that TSI-II-A data would be used for analyses regarding dissociation, as no other scale in the battery measures the construct. Examinees who failed PAI symptom validity tests (Infrequency, Inconsistency, Positive Impression Management, and Negative Impression Management) were excluded from analyses regarding depression, anxiety, and PTSD, and examinees who failed TSI-II-A symptom validity tests were excluded from analyses regarding dissociation due to small sample size (see Table 32).

It should be noted that symptom validity test failure was relatively rare in this dataset (ranging from 2.70% to 6.69% depending on the measure), which contrasts starkly with the very high rates of PVT failure previously discussed. One contributing factor may be selection bias, as only a subset of the sample completed the TSI-II-A and PAI. These measures were typically completed at the end of the battery. At times, the battery would be truncated or incomplete when examinees had poor attendance or engaged in extreme pain behaviour such as: curtailing assessment sessions after brief periods of testing (e.g., after only 30 minutes); taking long, frequent breaks; crying frequently and profusely; lying down on the floor with complaints of pain and fatigue; and completing items very slowly. When pain behaviour or poor attendance were present inventories including the TSI-II-A and PAI were the least likely to be completed. These examinees may have been more likely to fail PVTs (Webb et al., 2012). The absence of their data may have resulted in a sample that includes patients who are more likely to produce credible response sets on neuropsychological testing. Furthermore, several of the symptom validity tests do not measure symptom over-reporting, but rather are designed to detect failure to read or understand the items (e.g., Response Level and Inconsistency), or symptom under-reporting (i.e., Positive Impression Management). A combination of these factors likely contributed to the discrepancy between PVT and symptom validity test BR_{FAIL} in this data set.

Table 30

Sample Size, Mean, Standard Deviation, Range, Median, Skewness, Kurtosis, and

Shapiro-Wilk's Test for BDI-II and BAI

	<i>N</i>	<i>M</i>	Classification Range				SD	Range	<i>Mdn</i>	Skew	Kurtosis
			<i>Min</i>	<i>Mild</i>	<i>Mod</i>	<i>Sev</i>					
BDI-II	279	28.38	≤13	14-19	20-28	≥29	14.40	0,61	28	.183	-.866
BAI	285	22.28	≤7	8-15	16-25	≥26	14.82	0,62	21	.612	-.298

Note. BDI-II = Beck Depression Inventory – Second Edition (Beck et al., 1996); BAI = Beck Anxiety

Inventory (Beck & Steer, 1993); Min = Minimal; Mod = Moderate; Sev = Severe.

Table 31

One-way ANOVAs on BDI-II and BAI Scores for Examinees with Valid, Invalid, and Missing TSI-II-A Data

	BDI-II			BAI		
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>
TSI Pass < 15 ATR and < 76 RL	128	30.78 ^a	12.07	132	24.52 ^b	13.23
TSI Fail ≥ 15 ATR or ≥ 76 RL	11	34.64	22.98	11	28.45	24.04
TSI Missing	140	25.70 ^a	15.10	142	19.79 ^b	15.21
<i>df</i>			2			2
<i>F</i>			5.41			4.51
<i>p</i>			.005			.012
Partial η^2			.04			.03
Kruskal-Wallis <i>H</i>			10.89			10.99
<i>df</i>			2			2
<i>p</i>			.004			.004

Note. BDI-II = Beck Depression Inventory – Second Edition (Beck et al., 1996); BAI = Beck Anxiety Inventory (Beck & Steer, 1993); TSI = Trauma Symptom Inventory – Second Edition Alternate (Briere, 2011); TSI = Trauma Symptom Inventory – Second Edition (Briere, 2011); ATR = Atypical Responses Raw Score; RL = Response Level *T*-score; ^{a, b} = Significant post hoc comparisons.

Table 32

PAI and TSI-II-A Symptom Validity Test Failures

Test	SVT	<i>N</i>	<i>M</i>	<i>SD</i>	Cutoff	<i>BR_{FAIL}</i>
TSI-II-A	Response Level	148	52.22	8.92	≥ 76 T	2.70
	Atypical Responses		5.20	4.56	≥ 15 Raw	4.73
PAI	Inconsistency	239	54.17	9.33	≥ 73 T	5.86
	Infrequency		54.12	9.98	≥ 75 T	4.60
	Negative Impression Management		64.53	16.41	≥ 92 T	6.69
	Positive Impression Management		47.45	11.16	≥ 68 T	3.35

Note. TSI-II-A = Trauma Symptom Inventory – Second Edition Alternate (Briere, 2011); PAI = Personality Assessment Inventory (Morey, 1997); SVT = Symptom Validity Test; T = *T*-score; *BR_{FAIL}* = Base rate of failure.

Question 1

Would self-reported depression symptoms be associated with BR_{FAIL} ?

Three one-way ANOVAs were conducted to compare TOMM Trial 1, EI-7_{VER}, EI-7_{VIS}, and VI-10 scores in those with low (≤ 19 BDI-II raw score), moderate (20-28 BDI-II raw score), and severe (≥ 29 BDI-II raw score) self-reported depressive symptoms. The minimal and mild groups were collapsed due to small group sizes. Levene's Test was significant for the TOMM Trial 1 one-way ANOVA, indicating that the assumption of equality of variance was violated. Follow-up Kruskal-Wallis testing was conducted for this analysis. Four 2x3 χ^2 analyses were conducted to compare TOMM, EI-7_{VER}, EI-7_{VIS}, and VI-10 BR_{FAIL} for those with low, moderate, and severe self-reported depression symptoms. Results of the one-way ANOVAs are presented in Table 33, and results of the χ^2 analyses are presented in Table 34. Results of the one-way ANOVAs indicate that those with moderate and severe self-reported depressive symptoms had lower TOMM Trial 1 scores than those with low self-reported depressive symptoms and that examinees with severe self-reported depression scored higher than those with low self-reported depression on the EI-7_{VER}, with no other significant differences found. Results of χ^2 analyses indicate significant differences in TOMM, EI-7_{VER}, and VI-10 BR_{FAIL} across depression groups. Examination of standardized residuals indicated that those with low self-reported depression were significantly less likely to fail the TOMM than others and that those with severe self-reported depression were significantly more likely to fail the TOMM as compared to other groups.

Table 33

One-way ANOVA on TOMM Trial 1, EI-7_{VER}, EI-7_{VIS}, and VI-10 Scores in Examinees with Low, Moderate, and Severe Self-Reported Depression (BDI-II)

BDI-II Raw Scores	TOMM 1			EI-7 _{VER}			EI-7 _{VIS}			VI-10		
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>
Low ≤ 19	80	44.65 ^{ab}	6.03	75	2.05 ^c	2.95	78	1.85	2.72	77	2.69	3.73
Moderate 20-28	45	41.22 ^a	7.93	44	2.32	3.09	46	2.26	2.70	46	3.22	3.96
Severe ≥ 29	109	38.72 ^b	8.84	91	3.20 ^e	3.17	90	3.02	3.72	92	4.01	4.28
<i>df</i>			2			2			2			2
<i>F</i>			13.28 ^d			3.09			2.94			2.31
<i>p</i>			<.001			.048			.055			.102
Partial η^2			.10			.03			.03			.02
Kruskal-Wallis <i>H</i>			25.54									
<i>df</i>			2									
<i>p</i>			<.001									

Note. Post hoc analyses used Tukey's HSD for control of Type 1 error. TOMM 1 = Test of Memory Malinger Trial 1 Raw Score (Tombaugh, 1996); EI-7_{VER} = Erdodi Index – Seven Verbal; EI-7_{VIS} = Erdodi Index – Seven Visuomotor; VI-10 = Validity Index – Ten; BDI-II = Beck Depression Inventory – Second Edition (Beck et al., 1996); ^{a, b, c} = Significant post hoc analyses; ^d = Levene's test for equality of variance significant, Kruskal-Wallis test reported.

Table 34

TOMM, EI-7_{VER}, EI-7_{VIS}, and VI-10 BR_{FAIL} for Examinees with Low, Moderate, and Severe Self-Reported Depression (BDI-II)

BDI-II		TOMM	EI-7 _{VER}	EI-7 _{VIS}	VI-10
Raw Scores	<i>N</i>	236	154	162	168
Low	<i>n</i>	80	53	64	61
≤ 19	BR _{FAIL}	12.5	24.5	20.3	31.1
	<i>z</i>	-2.6**	-1.4	-1.4	-1.5
Moderate	<i>n</i>	47	35	35	34
20-28	BR _{FAIL}	19.1	28.6	28.6	47.1
	<i>z</i>	-1.1	-.8	-.1	.3
Severe	<i>n</i>	109	66	63	73
≥ 29	BR _{FAIL}	42.2	50.0	39.7	52.1
	<i>z</i>	2.9**	1.8	1.5	1.1
χ^2		22.47	9.43	5.74	6.14
<i>p</i>		<.001	.009	.057	.046
Φ^2		.31	.25	.19	.19

Note. TOMM = Test of Memory Malingering (Tombaugh, 1996); EI-7_{VER} = Erdodi Index – Seven Verbal; EI-7_{VIS} = Erdodi Index – Seven Visuomotor; VI-10 = Validity Index – Ten; BR_{FAIL} = Base rate of failure (≤ 44 TOMM Trial 2 or Retention, or ≥ 4 on EI-7_{VER}, EI-7_{VIS}, and VI-10); BDI-II = Beck Depression Inventory – Second Edition (Beck et al., 1996); *z* = standardized residual; ** = Significant at *p* < .01.

Question 2

Would self-reported anxiety symptoms be associated with BR_{FAIL} ?

Three one-way ANOVAs were conducted to compare TOMM Trial 1, EI-7_{VER}, EI-7_{VIS}, and VI-10 scores across examinees with low self-reported anxiety (BAI raw score ≤ 15), moderate self-reported anxiety (BAI raw score 16-25), and severe self-reported anxiety (BAI raw score ≥ 26). The minimal and mild self-reported anxiety groups were combined due to small sample sizes. Levene's Test was significant for TOMM Trial 1, EI-7_{VER} and EI-7_{VIS} one-way ANOVAs regarding anxiety, indicating a violation of the assumption of equality of variance. One-way ANOVAs are not robust to violations of assumptions of normality and equality of variance when sample sizes are not equal, which is the case here (Skidmore & Thompson, 2013; Stevens, 2009). As a result, follow-up Kruskal-Wallis tests were conducted, and are displayed in Table 35 alongside one-way ANOVA results. Results of the one-way ANOVAs indicate that those with low self-reported anxiety had higher TOMM Trial 1 scores than those with moderate or severe self-reported anxiety. Examinees with low self-reported anxiety also had lower EI-7_{VER} scores than those with severe self-reported anxiety, and those with low self-reported anxiety had lower EI-7_{VIS} scores than those with moderate self-reported anxiety.

Four $2 \times 3 \chi^2$ analyses were conducted to compare TOMM, EI-7_{VER}, EI-7_{VIS}, and VI-10 BR_{FAIL} across anxiety groups. These analyses, presented in Table 36, indicate significant differences in BR_{FAIL} for TOMM, EI-7_{VER}, EI-7_{VIS}, and VI-10 across groups. Examination of standardized residuals indicated that examinees with low self-reported anxiety were significantly less likely than other groups to fail the TOMM and that

examinees with severe self-reported anxiety were significantly more likely to fail the TOMM and the EI-7_{VER}.

Table 35

One-way ANOVAs on TOMM Trail 1, EI-7_{VER}, EI-7_{VIS}, and VI-10 Scores in Examinees with Low, Moderate, and Severe Self-Reported Anxiety (BAI)

BAIRaw Scores	TOMM1			EI-7 _{VER}			EI-7 _{VIS}			VI-10		
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>
Low ≤ 15	101	44.18 ^{ab}	6.27	95	2.01 ^c	2.79	99	1.80 ^d	2.54	97	2.61	3.67
Moderate 16-25	54	40.69 ^a	6.91	48	2.40	2.96	49	3.16 ^d	3.70	51	4.00	4.75
Severe ≥ 26	86	37.71 ^b	9.81	71	3.44 ^c	3.40	71	2.96	3.70	70	3.89	3.82
<i>df</i>			2			2			2			2
<i>F</i>			15.88 ^e			4.56 ^e			4.09 ^e			2.98
<i>p</i>			<.001			.011			.018			.053
Partial η^2			.12			.04			.04			.03
Kruskal-Wallis <i>H</i>			26.18			9.95			8.48			
<i>df</i>			2			2			2			
<i>p</i>			<.001			.007			.014			

Note. Post hoc analyses used Tukey's HSD for control of Type 1 error. TOMM 1 = Test of Memory Malingered Trial 1 Raw Score (Tombaugh, 1996); EI-7_{VER} = Erdodi Index – Seven Verbal; EI-7_{VIS} = Erdodi Index – Seven Visuomotor; VI-10 = Validity Index – Ten; BAI = Beck Anxiety Inventory (Beck & Steer, 1991); ^{a, b, c, d} = Significant post hoc comparisons; ^e = Levene's test of equality of variance was significant, Kruskal-Wallis test reported.

Table 36

TOMM, EI-7_{VER}, EI-7_{VIS}, and VI-10 BR_{FAIL} for Examinees with Low, Moderate, and Severe Self-Reported Anxiety (BAI)

BAI		TOMM	EI-7 _{VER}	EI-7 _{VIS}	VI-10
Raw Scores	<i>N</i>	243	156	167	168
Low	<i>n</i>	103	66	79	74
≤ 15	BR _{FAIL}	14.6	22.7	19.0	29.7
	<i>z</i>	-2.6**	-1.7	-1.8	-1.7
Moderate	<i>n</i>	54	39	39	38
16-25	BR _{FAIL}	29.6	33.3	41.0	50.0
	<i>z</i>	.2	-.2	1.3	.7
Severe	<i>n</i>	86	51	49	56
≥ 26	BR _{FAIL}	43.0	52.9	38.8	55.4
	<i>z</i>	2.6**	2.1*	1.1	1.4
χ^2		18.93	11.59	8.63	9.57
<i>p</i>		<.001	.003	.013	.008
Φ^2		.28	.27	.23	.24

Note. TOMM = Test of Memory Malinger (Tombaugh, 1996); EI-7_{VER} = Erdodi Index – Seven Verbal; EI-7_{VIS} = Erdodi Index – Seven Visuomotor; VI-10 = Validity Index – Ten; BR_{FAIL} = Base rate of failure (≤ 44 TOMM Trial 2 or Retention, or ≥ 4 on EI-7_{VER}, EI-7_{VIS}, and VI-10); BAI = Beck Anxiety Inventory (Beck & Steer, 1991); *z* = standardized residual; * = Significant at *p* < .05; ** = Significant at *p* < .01.

Question 3

Would self-reported PTSD symptoms be associated with BR_{FAIL}?

Three one-way ANOVAs were conducted comparing TOMM Trial 1, EI-7_{VER}, EI-7_{VIS}, and VI-10 scores across examinees with no self-reported PTSD symptoms (PAI Anxiety Related Disorders Traumatic Stress Scale T -score ≤ 59), mild self-reported PTSD symptoms (T -score 60-69), and moderate-to-severe self-reported PTSD symptoms (T -score ≥ 70). The moderate and severe categories were collapsed due to the inadequate sample size of the severe group ($n = 8$). Levene's Test was significant for the EI-7_{VIS} one-way ANOVA in this section, indicating that the assumption of equality of variance was violated. A follow-up Kruskal-Wallis test was conducted, as ANOVAs are not robust to violations of assumptions of equality of variance when group sizes are unequal (Skidmore & Thompson, 2013; Stevens, 2009). The results of the ANOVAs and Kruskal-Wallis test are displayed in Table 37. Results indicate that examinees with no self-reported PTSD symptoms had significantly lower scores on the EI-7_{VIS} and higher scores on TOMM Trial 1 as compared to examinees with moderate to severe self-reported PTSD symptoms.

Four $2 \times 4 \chi^2$ analyses were conducted to compare TOMM, EI-7_{VER}, EI-7_{VIS}, and VI-10 BR_{FAIL} across the PTSD groups. Results, presented in Table 38, indicate that TOMM BR_{FAIL} differs across PTSD severity. Examination of standardized residuals indicates that examinees with moderate to severe self-reported PTSD symptoms are significantly more likely to fail the TOMM than other groups.

Table 37

One-way ANOVAs on TOMM Trial 1, EI-7_{VER}, EI-7_{VIS}, and VI-10 Scores in Examinees with Low, Mild, and Moderate-to-Severe Self-Reported Posttraumatic Symptoms (PAI Anxiety Related Disorders Traumatic Stress Subscale)

PAITS	TOMM 1			EI-7 _{VER}			EI-7 _{VIS}			VI-10		
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>
None ≤59	92	42.90 ^a	6.63	89	2.16	2.91	91	1.77 ^b	2.16	90	2.58	2.99
Mild 60-69	35	42.11	7.79	31	1.77	1.86	34	1.88	3.15	34	3.09	3.97
Mod-Sev ≥70	66	39.55 ^a	8.07	56	2.61	2.28	58	3.02 ^b	3.61	56	3.14	2.92
<i>df</i>			2			2			2			2
<i>F</i>			4.09			1.13			3.57 ^c			.66
<i>p</i>			.018			.324			.030			.516
<i>η</i> ²			.04			.01			.04			.01
Kruskal-Wallis <i>H</i>									8.15			
<i>df</i>									2			
<i>p</i>									.017			

Note. Post hoc analyses conducted with Tukey's HSD to control for Type 1 error. TOMM 1 = Test of Memory Malinger Trial 1 Raw Score (Tombaugh, 1996); EI-7_{VER} = Erdodi Index – Seven Verbal; EI-7_{VIS} = Erdodi Index – Seven Visuomotor; VI-10 = Validity Index – Ten; PAI TS *T*-score = PAI Anxiety Related Disorders Traumatic Stress Subscale *T*-score (Morey, 1997); ^{a, b} = Significant post hoc comparisons; ^c = Levene's test of equality of variance was significant, Kruskal-Wallis test reported.

Table 38

TOMM, EI-7_{VER}, EI-7_{VIS}, and VI-10 BR_{FAIL} for Examinees with Low, Mild, and Moderate-to-Severe Self-Reported Posttraumatic Symptoms (PAI Anxiety Related Disorders Traumatic Stress Subscale)

PAI TS		TOMM	EI-7 _{VER}	EI-7 _{VIS}	VI-10
<i>T</i> -score	<i>N</i>	195	128	141	140
None ≤ 59	<i>n</i>	94	68	71	69
	BR _{FAIL}	18.1	29.4	22.5	34.8
	<i>z</i>	-1.6	-.5	-.7	-.8
Mild 60-69	<i>n</i>	35	22	29	26
	BR _{FAIL}	17.1	22.7	20.7	46.2
	<i>z</i>	-1.1	-.8	-.6	.4
Mod-Sev ≥ 70	<i>n</i>	66	38	41	45
	BR _{FAIL}	43.9	44.7	39.0	46.7
	<i>z</i>	2.7**	1.3	1.5	.6
χ^2		15.23	3.82	4.32	1.99
<i>p</i>		<.001	.148	.116	.371
Φ^2		.28	.17	.18	.12

Note. TOMM = Test of Memory Malingering (Tombaugh, 1996); EI-7_{VER} = Erdodi Index – Seven Verbal; EI-7_{VIS} = Erdodi Index – Seven Visuomotor; VI-10 = Validity Index – Ten; BR_{FAIL} = Base rate of failure (≤ 44 TOMM Trial 2 or Retention, or ≥ 4 on EI-7_{VER}, EI-7_{VIS}, and VI-10); PAI TS *T*-score = PAI Anxiety Related Disorders Traumatic Stress Subscale *T*-score (Morey, 1997); *z* = standardized residual; ** = Significant at *p* < .01.

Question 4

Would self-reported dissociative symptoms be associated with BR_{FAIL}?

As a result of the small sample size of clinically elevated dissociation symptoms ($n = 19$, TSI-II-A Dissociation T -score 60-64), it was decided that the elevated and clinical level groups would be collapsed for these analyses. Three independent samples t tests were conducted to compare TOMM Trial 1, EI-7_{VER}, EI-7_{VIS}, and VI-10 scores across examinees with normal self-reported dissociative symptoms (TSI-II-A Dissociation T -score ≤ 59) and moderate-to-severe self-reported dissociative symptoms (TSI-II-A Dissociation T -score ≥ 60). Levene's Test was conducted to examine the assumption of equality of variance and was significant for EI-7_{VER} t test. Mann-Whitney U tests were not conducted, as t tests are robust to violations of assumptions of normality and equality of variance when sample sizes are roughly equal, as in this case (Skidmore & Thompson, 2013; Stevens, 2009). Table 39 displays results of t tests. Results indicate that examinees with normal levels of dissociation had higher scores on TOMM Trial 1 and lower scores on the EI-7_{VER} than those with moderate-to-severe levels of self-reported dissociation, and this finding was upheld in nonparametric testing.

Four 2x2 χ^2 analyses were conducted to compare TOMM, EI-7_{VER}, EI-7_{VIS}, and VI-10 BR_{FAIL} across dissociative symptom groups. Results are displayed in Table 40 and indicate that those with moderate-to-severe self-reported dissociation symptoms were significantly more likely to fail the TOMM and EI-7_{VER} than those with normal levels of dissociative symptoms.

Table 39

Independent t Tests on TOMM Trial 1, EI-7_{VER}, EI-7_{VIS}, and VI-10 Scores in Examinees with Normal and Elevated Self-reported Dissociative Symptoms (TSI-II-A Dissociation)

TSI-II-A Dissociation T-score	TOMM 1			EI-7 _{VER}			EI-7 _{VIS}			VI-10		
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>
Normal ≤ 59	66	41.86	7.44	60	2.05	2.51	64	2.67	3.60	60	2.98	3.56
Elevated ≥ 60	68	38.35	8.36	58	3.34	3.31	62	2.94	3.21	63	4.10	3.80
<i>df</i>	132			106.32			124			121		
<i>t</i>	2.57			2.39 ^a			.43			1.67		
<i>p</i>	.011			.019			.665			.097		
<i>d</i>	.44			.44			.08			.30		
95% CI	.803, 6.22			.22, 2.37			-.94, 1.47			-.20, 2.43		

Note. ^a = Levene's test of equality of variance was significant, *t* test with equality of variance not assumed is reported; TOMM 1 = Test of Memory Malingering Trial 1 Raw Score (Tombaugh, 1996); EI-7_{VER} = Erdodi Index – Seven Verbal; EI-7_{VIS} = Erdodi Index – Seven Visuomotor; VI-10 = Validity Index – Ten; 95% CI = 95% Confidence interval; TSI-II-A Dissociation *T*-score = Self-reported Trauma Symptom Inventory – Second Edition Alternate Dissociation Scale *T*-score (Briere, 2011).

Table 40

TOMM, EI-7_{VER}, EI-7_{VIS}, and VI-10 BR_{FAIL} for Examinees with Normal and Elevated Self-Reported Dissociative Symptoms (TSI-II-A Dissociation)

TSI-II-A		TOMM	EI-7 _{VER}	EI-7 _{VIS}	VI-10
Dissociation	<i>N</i>	136	91	92	100
<i>T</i> -score					
Normal ≤ 59	<i>n</i>	68	45	48	46
	BR _{FAIL}	25.0	28.9	31.3	41.3
Elevated ≥ 60	<i>n</i>	68	46	44	54
	BR _{FAIL}	42.6	50.0	43.2	57.4
χ^2		4.73	4.24	1.40	2.58
<i>p</i>		.030	.039	.236	.108
Φ^2		.19	.22	.12	.16
RR (95% CI)		1.31 (1.02, 1.67)	1.42 (1.01, 2.01)	1.21 (.88, 1.67)	1.38 (.93, 2.04)

Note. TOMM = Test of Memory Malinger (Tombaugh, 1996); EI-7_{VER} = Erdodi Index – Seven Verbal; EI-7_{VIS} = Erdodi Index – Seven Visuomotor; VI-10 = Validity Index – Ten; BR = Base rate of failure (≤ 44 TOMM Trial 2 or Retention, or ≥ 4 on EI-7_{VER}, EI-7_{VIS}, and VI-10); TSI-II-A Dissociation *T*-score = Self-reported Trauma Symptom Inventory – Second Edition Alternate Dissociation Scale *T*-score (Briere, 2011). RR = Relative risk ratio; 95% CI = 95% confidence interval.

CHAPTER 11

Discussion Objective 2

The results of questions regarding self-reported depression, anxiety, and PTSD symptoms will be addressed first as a group, followed by a discussion of self-reported dissociation symptoms.

Self-reported Depression, Anxiety, and PTSD Symptoms and BR_{FAIL}

As was previously discussed, the intersection of mood disorders, TBI, and compensation-seeking is complex. Research suggests that it is likely that psychopathology often predates mTBI and contributes to poor prognosis following the injury (Moore et al., 2006), but that TBI itself likely contributes to the development of mood symptoms as well (Hesdorffer et al., 2009). Confounded with this, researchers disagree about whether PTSD can develop in the presence of loss of consciousness (Hesdorffer et al., 2009). PTSD is often claimed as a psychological injury following TBI. As such, most PTSD PVT research is aimed at identifying noncredible PTSD presentation (e.g., Young, 2015a), rather than exploring credible PTSD might affect PVT performance, as is the case with other diagnoses such as major depressive disorder and schizophrenia (e.g., Schroeder & Marshall, 2011).

The results of the current study indicate that there is a significant effect of self-reported depressive symptoms on EI-7_{VER} scores, with severe self-reported depressive symptoms being associated with higher EI-7_{VER} scores. Higher self-reported depressive symptoms were also associated with higher BR_{FAIL} on the TOMM, EI-7_{VER}, and VI-10. Similarly, low self-reported anxiety symptoms were associated with lower EI-7_{VER} and

EI-7_{VIS} scores, and higher anxiety was associated with higher BR_{FAIL} on the TOMM, EI-7_{VER}, EI-7_{VIS}, and VI-10 when the data were analyzed dichotomously.

Regarding self-reported PTSD symptoms, there was an effect on EI-7_{VIS} scores with lower self-reported PTSD symptoms being associated with lower EI-7_{VIS} scores and an effect of self-reported PTSD symptoms on TOMM BR_{FAIL} with high self-reported PTSD symptoms being associated with higher BR_{FAIL}. It appears that a relationship between self-reported PTSD symptoms and BR_{FAIL} only exists on the TOMM, which may indicate that the composite EIs are less susceptible to PTSD symptoms than is the TOMM, although replication is necessary to confirm this result. Composite EIs may be less susceptible to PTSD because they measure a continuum of credibility over the duration of the assessment, rather than a window into credibility at a specific time point. This may make these measures more resistant to transient experiences that interfere with test engagement, such as the intrusive symptoms of PTSD (e.g., flashbacks, intense distress) that may affect the examinee in brief, discrete periods during the assessment. If replicated, this would raise confidence in the emerging methods based on the underlying proposition that systematically interpreting findings *across* multiple PVTs is more sensitive and specific than interpreting single, more well-established PVTs (Berthelson et al., 2013; Boone, 2009; Erdodi & Lichtenstein, 2017; Odland et al., 2015).

The greater number of significant effects and larger effects sizes when outcomes were measured dichotomously (i.e., Pass/Fail) as compared to continuously highlights issues that Bigler (2012, 2015) raises about “near-pass” PVT performance. Although there is an underlying continuum of performance validity (Erdodi & Lichtenstein, 2017), differences are easier to determine and interpret when those with borderline PVT

performance are removed from analysis, and only those with clearly credible performance and those with unambiguously noncredible performance are compared.

When this method is employed, the effect sizes are larger and more interpretable.

As with any classification system, data points that are very close to the cutoff are more ambiguous than more extreme values. In the case of the current research, for example, a score of zero on the VI-10 would be a strong indication that the results of the examination are credible and a score of one would be slightly ambiguous but still indicate credible results. A score of two or three would be ambiguous enough that it is challenging to classify the performance as either credible or noncredible. Likewise, as the VI-10 score continues to increase, so does the confidence that the examinee's performance is not credible, such that a score of 30 (i.e., a maximum score) would be much more compelling evidence of noncredible performance than a score of five.

Collapsing the distribution allows the researcher to evaluate the data with less impact of influential outliers, and removing ambiguous performers reduces ambiguity stemming from intermediate scores and incorrectly classified data. However, these statistical methods do not solve the problem raised by Bigler (2015) of clinical interpretation with the ambiguous/indeterminate group he dubs "near pass," more recently labelled "soft fail" (Erdodi & Lichtenstein, 2017). In this regard, the work of the neuropsychologist extends beyond the arithmetic to clinical interpretation. Clinicians must take into account the correspondence between the objective severity of the injury and the neuropsychological data and consistency between the reported symptom severity, the internal consistency of neuropsychological data, and objective level of functional impairment (Slick et al., 1999).

In addition, as explored in the discussion of the first objective, it is important for the clinician to take into account contextual and cultural factors when interpreting these indeterminate cases. The clinician should also consider the possible consequences of a determination of noncredible performance. For example, if the clinician finds indeterminate PVT performance in healthy athletes during baseline testing for sports, the consequences of noncredible determination would perhaps include some mild embarrassment, re-testing, and ultimately greater safety for athletes should they be injured. Conversely, if the clinician assesses a refugee with a complex trauma history following a motor vehicle accident, the consequences of a noncredible determination may be more damaging. Effects could include shame, alienation from the dominant culture, and denial of benefits with resultant significant harm to examinees and their family due to lost income and poorer mental and physical health.

Ambiguous results are likely to remain problematic in any classification system that seeks to classify inherently continuous data into dichotomous categories. Nevertheless, multivariate PVT classification methods such as the EI model have merit in harnessing the sensitivity of multiple measures while controlling for false-positive errors leading to better overall accuracy (Erdodi & Lichtenstein, 2017; Odland et al., 2015).

The results that self-reported depression, anxiety, and PTSD symptoms have a significant, albeit small, effect on BR_{FAIL} preliminarily indicates that mood symptoms should be taken into account in the determination of credibility. More specifically, in some cases, if an examinee has very high levels of self-reported symptomatology and ambiguous “near pass” PVT failure, the mood symptoms may have contributed to the PVT failure. As the number and severity of PVT failures increases, so does the

confidence in noncredible determination, even in the presence of high self-reported symptomatology.

These findings may not be generalizable to cases in which there is strong objective evidence of severe psychiatric pathology. For example, an examinee who fails PVTs but is in an inpatient facility for major depressive disorder with severe vegetative symptoms following a suicide attempt and may still have credible performance. Examinees with such severe symptoms may not be capable of engaging adequately in testing at that point. The results of neuropsychological testing may be representative of their best performance at that time, but this performance may have been temporarily compromised by factors like inattention or behavioural problems (Marcopulos et al., 2014). Nonetheless, even examinees with severe self-reported depressive symptoms usually pass PVTs (O'Bryant et al., 2007; Schroeder & Marshall, 2011; Yanez et al., 2006). Thus, in the absence of clear, objective evidence of severe psychiatric pathology, unambiguous PVT failure can be confidently used for determining noncredible performance.

Dissociation and BR_{FAIL}

Self-reported dissociation was found to have a small-to-medium effect on $EI-7_{VER}$ scores, and a small effect on TOMM BR_{FAIL} and $EI-7_{VER} BR_{FAIL}$, with no significant findings regarding $EI-7_{VIS}$ or VI-10. Previous research into the association between dissociative symptoms and neuropsychological functioning indicates that dissociative symptoms negatively affect a variety of cognitive domains including attention; executive function; and working, verbal and visual memory (Haaland & Landro, 2009; McKinnon et al., 2016; Parlar et al., 2016). Results of the current study may suggest that these

cognitive deficits can interfere with embedded validity indicators that tap into higher order or more complex cognitive functioning that require greater working memory load and which are clustered in the EI-7_{VER} (e.g., Controlled Oral Word Association, Digit Span) while sparing relatively simplistic and motorically mediated tasks that are clustered in the EI-7_{VIS} (e.g., Finger Tapping Test, Trail Making Test-A).

Alternatively, there may be some aspect of verbally engaging with the examiner or even simple attention to and from the examiner that may contribute to lower scores in those with elevated dissociative symptoms. The higher TOMM BR_{FAIL} in those with elevated self-reported dissociative symptoms supports this explanation, as the TOMM also requires the examinee to indicate their choices verbally to the examiner. Some preliminary evidence that suggests dissociative symptoms interfere with social cognition (Nazarov et al., 2015) also supports this explanation.

Previously noted difficulties in disentangling premorbid psychiatric symptoms, effects of the TBI and motivation to appear impaired all apply in the case of self-reported dissociative symptoms in the same way that they do for the other self-reported psychiatric symptoms. The use of a the TSI-II-A Dissociation subscale, which has relatively little research, and a sample with complex interacting factors contributing to PVT failure also limits the generalizability of the findings.

CHAPTER 12

General Discussion

The current study explored the relationship between cultural, demographic, linguistic, and psychiatric factors and PVT performance in a sample of compensation-seeking examinees who had been involved in motor vehicle accidents in Southern Ontario. The study used composite embedded validity indicator measures that allowed a multivariate approach to assessing performance validity. One significant challenge in the current research was the unexpectedly high BR_{FAIL} across embedded validity indicators and the TOMM when using a priori cutoff scores based on the previous literature.

It was necessary to adjust the cutoff scores to perform meaningful analyses. As previously discussed, the development of cutoff scores is influenced by BR_{FAIL} of criterion measures and experimental measures, as well as the study design. The evaluation context matters as well, where clinical evaluations tend to yield higher BR_{FAIL} than research assessments in otherwise similar samples (McCormick, Yoash-Gantz, McDonald, Campbell, & Tupler, 2013).

There is no obvious remedy for this challenge. One strategy may be to use the same cutoff scores across various populations without regard to the demand characteristics of the situation. This strategy assumes that scores that exceed the cutoff are universally indicative of noncredible performance and that scores below the cutoff universally indicate credible performance.

If psychologists were to use this indiscriminate strategy, an artificially low proportion of athletes would be identified as “sandbagging” baseline (i.e., preseason) neurocognitive testing to return to play faster following a concussion. If this were the

case, postconcussion testing of true impairment would be compared to noncredible baseline performance, which may lead to a return to play before recovery from the concussion. If the athlete were to sustain another head injury during play before recovery from the first, the damage from the second injury may be far worse than expected and can be fatal in rare cases (Cantu & Gean, 2010). Meanwhile, a high proportion of personal injury and disability claim examinees might be denied benefits and support that they need and to which they are entitled for legitimate impairment (Bigler, 2012). This problem is significantly compounded by the growing evidence of the impact of cultural and linguistic factors on PVT BR_{FAIL} (Erdodi, Nussbaum, et al., 2017; Nijdam-Jones & Rosenfeld, 2017).

An alternative, which the EI model provides, is to look at the continuum of performance validity. As previously outlined, the consideration of a gradation of confidence in noncredible scores across multiple PVTs allows the researcher and clinician to adopt a more nuanced perspective on noncredible performance in the neuropsychological assessment. The current research, along with the previously described range of BR_{FAIL} across samples, also highlights the need to develop cutoffs and/or algorithms that are appropriate both to the personal and cultural characteristics of the examinee and to the evaluation context.

Of particular note, in this case, are the similarities and differences between the current findings and a recent study with an Arabic-English bilingual community sample (Erdodi, Nussbaum, et al., 2017). In the current study, there was no significant difference between the limited English proficiency and the Anglophone Canadian groups on $EI-7_{VER}$ scores or BR_{FAIL} . Additionally, the limited English proficiency group's verbal embedded

validity indicator BR_{FAIL} were similar to those in Erdodi, Nussbaum, et al.'s study (2017). Conversely, in the current study, the limited English proficiency group had much higher BR_{FAIL} on the EI-7_{VIS} than the Anglophone Canadian group, whereas in the previous study language dominance did not affect nonverbal PVTs (Erdodi, Nussbaum, et al., 2017).

This discrepancy suggests that *context matters*. Being evaluated following a motor vehicle accident changes the relationship between limited English proficiency and PVT performance. As previously discussed, it is possible that the unusually high BR_{FAIL} on EI-7_{VIS} and TOMM among examinees with limited English proficiency represents a cultural concept of distress. All of the examinees in the current study were in a motor vehicle accident, and likely experienced some form of injury in the accident. The expression of impairment by the examinees is likely influenced by the demand characteristics of the evaluation. In other words, to be approved for benefits, they must communicate to the examiner that they have significant deficits. This interacts with cultural perceptions of impairment, which is understood in most of the world as involving primarily physical limitations (Rohlof, Knipscheer, & Kleber, 2014). It also interacts with the examinee's understanding of, and ability to engage with, the Canadian health care system.

Nevertheless, in the context of compensation-seeking, multiple factors can also contribute to the maintenance and exacerbation of somatic complaints. These factors include intrapsychic factors such as attribution of sensations to pathology resulting from the accident, attention to symptoms and emotional arousal (Kirmayer & Sartorius, 2008). They can also include help-seeking behaviour and iatrogenic effects of the insurance

process, family system reinforcement of distress, and the sick role (Kirmayer & Sartorius, 2008; Young, 2008).

All of these factors create feedback loops in which the individual learns to attribute sensations to pathology, catastrophizes these sensations, avoids activity and ultimately becomes deconditioned, leading to greater unpleasant sensations and greater disability (Young, 2008). In many countries, mental health professionals are seen as being exclusively present for the treatment of psychosis, which further dissuades people from presenting or interpreting their difficulties as having a psychological component, which in their perception would mean they are “crazy” (Rohlof et al., 2014).

Limited English proficiency examinees, because of these interconnected factors, may express cultural concepts of distress through pain behaviours and motor slowing that result in EI-7_{VIS} failure. This interpretation is bolstered by the finding of better than expected performance of examinees with limited English proficiency on the EI-7_{VER} when compared to Canadian-born examinees and previous research, which would imply that these examinees were engaging to the best of their ability in some portions of the testing. If these findings are replicated in future studies, the counterintuitive conclusion may be that in examinees with limited English proficiency involved in motor vehicle accidents PVTs with high verbal mediation are better representations of performance validity than PVTs with low verbal mediation. It also suggests the need to validate and/or develop PVTs that are relatively impervious to the effects of limited English proficiency, cultural concepts of distress, and health literacy. Further, it reinforces the importance of testing even seemingly obvious assumptions about the psychometric properties of a given instrument.

Another important finding in the current research is the relationship between self-reported dissociative symptoms and poor EI-7_{VER} and TOMM performance. To this author's knowledge, this is the first study to explore the effects of dissociative symptoms on PVT performance. Considering the fairly well-established link between dissociative symptoms and neuropsychological impairment, both in the general population (Ozdemir, Ozdemir, Boysan, & Yilmaz, 2015) and with clinical samples (Haaland & Landro, 2009; Parlar et al., 2016), it is imperative for researchers to explore this area further.

There are important implications of dissociative symptoms affecting PVT performance. Dissociative symptoms and disorders have been strongly associated with the experience of trauma (American Psychiatric Association, 2013; Giesbrecht et al., 2008), especially complex childhood abuse (Nazarov et al., 2015), and other forms of extreme stress (Sandole & Auerbach, 2013). Dissociative symptoms are also associated with many forms of psychopathology, from anxiety disorders (Belli, 2014) to psychotic disorders (Sar et al., 2010), and predict greater morbidity and suicidality (Stein et al., 2013) and poorer treatment response (Bae, Kim, & Park, 2015).

Furthermore, practitioners are unlikely to screen for dissociative symptoms compared to other forms of psychopathology (Steinberg & Schnall, 2003). It is especially important for clinicians and researchers to screen for dissociative symptoms, pre-accident histories of childhood trauma, and exposure to extreme stress considering the prevalence of dissociative symptoms (Soffer-Dudek, 2014), their link to poor neuropsychological performance and past trauma. When motor vehicle accident examinees pass symptom validity tests and endorse dissociative symptoms, clinicians

should consider the possible contributions of these alterations in consciousness to PVT failure.

It is important that further studies be conducted on PVT performance with the use of more well-researched measures of dissociation such as the Dissociative Experiences Scale (Bernstein & Putnam, 1986). In concert with the larger field of PVT research, future studies should include simulation studies and clinical studies with a variety of populations. These populations should include individuals with dissociative disorders compared to other mental illnesses, dissociative vs. nondissociative subtypes of PTSD, and compensation-seeking clients with and without dissociative symptoms. A series of these studies would help elucidate the relationship between dissociative symptoms and PVT failure. This information would then help clinicians to make accurate distinctions between PVT failure related to dissociative disorders' hallmark alterations in consciousness and PVT failure indicative of noncredible performance.

Strengths of the Present Research

This research is the first, to the author's knowledge, to explore the effects of limited English proficiency on PVT performance in a forensic sample. The findings, as previously discussed, have implications for the interpretation of PVT results in clinical assessments. This research also reinforces the need to address demographic factors in research studies across contexts to facilitate valid and equitable PVT interpretation. The results of the first objective of this study suggest that clinicians and researchers should not generalize results of studies of the effect of limited English proficiency on PVT performance using healthy participants to forensic samples without further validation.

The first objective of this study suggests that when immigrant examinees are assessed following injuries in the context of forensic assessment in Ontario, their patterns of performance do not follow those of healthy examinees assessed in a pure research context in the same province. Further, the pattern of performance preliminarily suggests that cultural concepts of distress and health literacy might play a role, such that forensic examinees who immigrated to Canada and whose first language is not English are more likely to present with higher BR_{FAIL} in motorically mediated embedded validity indicators, but perform as well as Anglophone Canadians on verbally mediated embedded validity indicators.

This is also the first research to explore the effects of dissociative symptoms on PVT performance in any setting. This research is long overdue, as dissociative symptoms include alterations in consciousness, attention, and consolidation of information that interfere with functioning and lead to lower neuropsychological test scores. The findings from the current research suggest that individuals with elevated self-reported dissociative symptoms perform more poorly on verbally mediated embedded validity indicators as compared to visuomotor embedded validity indicators.

Underlying factors that may contribute to this are the heavier reliance of these tasks on verbal memory and working memory and/or direct verbal engagement with the examiner that is necessary for these tasks compared to visuomotor measures. Further research is needed to replicate these findings with forensic, clinical, and simulation samples. If the results of the second objective are replicated, they would indicate that dissociative symptoms interfere with performance in neuropsychological assessment and that PVTs should be interpreted with caution in the presence of a premorbid history of

complex child maltreatment and/or extreme stress resulting from events such as war, refugee experiences, and natural disasters. Further, this study highlights the importance of neuropsychologists screening for both premorbid trauma histories and the presence of significant dissociative experiences, as these factors can significantly affect client well-being and neuropsychological test results.

Limitations of the Present Research

Several factors limit the inferences drawn from the present study. The main one is the unusually high rates of PVT failure at traditional cutoff scores. The reasons for the abnormally poor performance of the sample as a whole are not clear. The province in which the data were collected, characterized by limited referral sources leading to selection bias, and/or contextual factors in the motor vehicle insurance system of Ontario may have affected these results. The use of data collected in a single neuropsychologist's practice may have also affected results. The restricted data source may have contributed to sampling bias, as well as introducing the possibility of site-specific biases that may have affected results.

Selection biases may also limit the findings. Examinees were all assessed as part of evaluations for compensation-seeking following motor vehicle accidents. The study excluded data from individuals who declined the use of their data in research, those who did not attend their scheduled assessments, and those who did not complete enough measures for their data to be included in analyses. These selection biases limit the generalizability of the obtained results.

The inherent constraints of using secondary data also limited the amount and types of information obtained. The inclusion of more stand-alone PVTs, better-

established measures of dissociation such as the DES, more detailed demographic information and the inclusion of measures of health literacy, acculturation, and acculturative stress would have widened the scope of analyses.

Another limitation to the current research regarding psychiatric symptoms is the contextual similarity between incentives to appear neuropsychologically impaired and to appear psychiatrically impaired to secure benefits following motor vehicle accident. This constrains the confidence in the effect of self-reported psychiatric symptoms on PVTs. It also restricts the range of scores obtained on these measures, with few examinees obtaining scores in the normal range. Extending the current work to research with clinical samples without incentive to appear impaired would strengthen conclusions about the relationship between PVT performance and psychiatric variables.

Future Directions

The current study provides important preliminary findings about the effects of demographic and psychiatric variables on PVT performance. Future studies that evaluate the impact of health literacy, acculturation, enculturation, and stereotype threat on PVT performance might give insight into more specific acculturative factors that contribute to PVT performance. Future simulation studies might include measures of language proficiency to assess the effect of language proficiency on PVT performance more directly. Measurement of language proficiency in forensic studies would likely be confounded by performance credibility, limiting the utility of that line of research.

Future research studies with large samples from specific cultural groups could help develop specific normative data and cutoff scores for PVTs. Different cultural groups may be affected to differing extents by the impacts of limited English proficiency,

cultural concepts of distress, and health literacy. For example, examinees who emigrated from Germany may have comparable levels of English proficiency as examinees who emigrated from India but may have very different cultural concepts of distress and/or functional Canadian health literacy. Development of appropriate normative data and cutoff scores for particular groups may help to mitigate some of the disparity between the guidelines for culturally competent neuropsychological assessment (Board of Directors, 2007; Canadian Psychological Association, 2000) and the dearth of culturally appropriate tests and norms for the completion of said assessments.

Future research should examine the utility of warning examinees of the role of credibility assessment. This may help to reduce the disparity between the informational context that examinees born in Canada and examinees born in other countries have when providing informed consent and engaging in neuropsychological assessment.

Regarding future psychiatric symptom research, it would be important to explore PVT performance with clinical samples that have varying incentives to appear impaired. Specifically, more research is necessary in PVT performance of clinical samples with primary dissociative disorders without external incentives and comparisons of PVT performance in PTSD samples with and without the dissociative subtype.

Conclusion

Results of the current study suggest that the relationship between limited English proficiency and PVT BR_{FAIL} functions differently in examinees assessed for forensic purposes following a motor vehicle accident as opposed to healthy examinees assessed in a research laboratory. These findings suggest that demand characteristics and cultural concepts of distress may contribute to a PVT pattern in which people with limited

English proficiency are more likely to fail motorically mediated PVTs. Simultaneously, examinees with limited English proficiency performed better than expected BR_{FAIL} on verbally mediated PVTs. Should these findings be replicable, they indicate the need for careful consideration of the indirect impact of culture on the nuanced expression of impairment following injury.

The first objective of this study highlights the need to develop algorithms or cutoff adjustments that take into account not only the direct impact of language proficiency on PVTs but also the intersection of cultural concepts of distress, health literacy and the context of the assessment.

Results of the second objective of this study give preliminary support to the notion that the negative effects of dissociative symptoms on neuropsychological test results generalize to PVT performance, especially in the case of verbally mediated PVTs and those that require the examinee to provide verbal answers to the assessor. Further research is necessary to confirm that these patterns replicate in other forensic samples, and generalize with clinical nonforensic samples.

References

- Albrecht, J.S., Liu, X., Smith, G. S., Baumgarten, M., Rattinger, G.B., Gambert, S. R., Langenberg, P., & Zuckerman, I. H. (2015). Stroke incidence following traumatic brain injury in older adults. *Journal of Trauma and Rehabilitation, 30*(2), E62-E67. doi: 10.1097/htr.0000000000000035
- Aldoory, L. (2017). The status of health literacy research in health communication and opportunities for future scholarship. *Health Communication, 32*(2), 211-218. doi: 10.1080/10410236.2015.1114065
- Allen, L. M., Conders, R. L., Green, P., & Cox, D. R. (1997). *CARB 97 manual for the computerized assessment of response bias*. Durham, NC: CogniSyst.
- American Congress of Rehabilitation Medicine; Mild Traumatic Brain Injury Committee (1993). Definition of mild traumatic brain injury. *Journal of Head Trauma Rehabilitation, 8*(3), 86-87. doi: 10.1097/00001199-199309000-00010
- American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders, 5th edition*. Washington, DC: Author.
- Archila-Suerte, P., Zevin, J., & Hernandez, A. E. (2015). The effect of age of acquisition, socioeducational status, and proficiency on the neural processing of second language speech sounds. *Brain and Language, 141*, 35-49. doi: 10.1016/j.bandl.2014.11.005
- Armistead-Jehle, P. (2010). Symptom validity test performance in U.S. veterans referred for evaluation of mild TBI. *Applied Neuropsychology, 17*(1), 52-59. doi: 10.1080/09084280903526182

- Arnold, G., Boone, K. B., Lu, P., Dean, A., Wen, J., Nitch, S., & McPherson, S. (2005). Sensitivity and specificity of finger tapping test scores for the detection of suspect effort. *The Clinical Neuropsychologist*, *19*(1), 105-120. doi: 10.1080/13854040490888567
- Ashendorf, L., Clark, E. L., & Sugarman, M. A. (2017). Performance validity and processing speed in a VA polytrauma sample. *The Clinical Neuropsychologist*, *31*(5), 857-866. doi: 10.1080/13854046.2017.1285961
- Ashendorf, L., Constantinou, M., & McCaffrey, R. J. (2004). The effect of depression and anxiety on the TOMM in community-dwelling older adults. *Archives of Clinical Neuropsychology*, *19*(1), 125-130. doi: 10.1093/arclin/19.1.125
- Ashendorf, L., O'Bryant, S. E., McCaffrey, R. J. (2003). Specificity of malingering detection strategies in older adults using the CVLT and WCST. *The Clinical Neuropsychologist*, *17*(2), 255-262. doi: 10.1076/clin.17.2.255.16502
- Axelrod, B. N., Fichtenberg, N. L., Millis, S. R., & Wertheimer, J. C. (2006). Detecting incomplete effort with digit span from the Wechsler Adult Intelligence Scale – Third edition. *The Clinical Neuropsychologist*, *20*(3), 513-523. doi: 10.1080/13854040590967117
- Axelrod, B. N., Meyers, J. E., & Davis, J. J. (2014). Finger Tapping Test performance as a measure of performance validity. *The Clinical Neuropsychologist*, *28*(5), 876-888. doi: 10.1080/13854046.2014.907583
- Babikian, T., Boone, K. B., & Arnold, G. (2006). Sensitivity and specificity of various digit span scores in the detection of suspect effort. *The Clinical Neuropsychologist*, *20*(1), 145-159. doi: 10.1080/13854040590947362

- Bae, H., Kim, D., & Park, Y. C. (2015). Dissociation predicts treatment response in eye movement desensitization and reprocessing for posttraumatic stress disorder. *Journal of Trauma and Dissociation, 17*(1), doi: 10.1080/15299732.2015.1037039
- Bashem, J. R., Rapport, L. J., Miller, J. B., Hanks, R. A., Axelrod, B. N., & Millis, S. R. (2014). Comparisons of five performance validity indices in bona fide and simulated traumatic brain injury. *The Clinical Neuropsychologist, 28*(5), 851-875. doi: 10.1080/13854046.2014.927927
- Bauer, L., Yantz, C. L., Ryan, L. M., Warden, D. L., & McCaffrey, R. L. (2005). An examination of the California Verbal Learning Test II to detect incomplete effort in a traumatic brain-injury sample. *Applied Neuropsychology, 12*(4), 202-207. doi: 10.1207/s15324826an1204_3
- Bay, E. & Donders, J. (2008). Risk factors for depressive symptoms after mild-to-moderate traumatic brain injury. *Brain Injury, 22*(3), 233-241. doi: 10.1080/02699050801953073
- Beck, A. T. (1987). *Beck Depression Inventory*. San Antonio, TX: Psychological Corporation.
- Beck, A. T., & Steer, R. A. (1993). *Beck Anxiety Inventory Manual*. San Antonio, TX: Psychological Corporation.
- Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Beck Depression Inventory II Manual*. San Antonio, TX: Psychological Corporation.
- Belanger, H. G., Curtiss, G., Demery, J. A., Lebowitz, B. K., & Vanderploeg, R. D. (2005). Factors moderating neuropsychological outcomes following mild

- traumatic brain injury: A meta-analysis. *Journal of the International Neuropsychological Society*, 11(3), 215-227. doi: 10.1017/S1355617705050277
- Belli, H. (2014). Dissociative symptoms and dissociative disorders comorbidity in obsessive compulsive disorder: Symptom screening, diagnostic tools and reflections on treatment. *World Journal of Clinical Cases*, 2(8), 327–331. doi: 10.12998/wjcc.v2.i8.327
- Benton, A. L., & Hamsher, K. D. (1978). *Multilingual aphasia examination: Manual*. Iowa City, IA: University of Iowa.
- Benton, A. L., Hamsher, K. D., Varney, N. R., & Spreen, O. (1983). *Judgment of line orientation*. New York, NY: Oxford University Press.
- Berna, G., Vaiva, G., Ducrocq, F., Duhem, S., & Nandrino, J. L. (2012). Categorical and dimensional study of the predictive factors of the development of a psychotrauma in victims of car accidents. *Journal of Anxiety Disorders*, 26(1), 239-245. doi: 10.1016/j.janxdis.2011.11.011
- Bernstein, E. M., & Putnam, F. W. (1986). Development, reliability, and validity of a dissociation scale. *Journal of Nervous and Mental Disease*, 174(12), 727-735. doi: 10.1097/00005053-198612000-00004
- Berthelson, L., Mulchan, S. S., Odland, A. P., Miller, L. J., & Mittenberg, W. (2013). False positive diagnosis of malingering due to the use of multiple effort tests. *Brain Injury*, 27(7-8), 909-916. doi: 10.3109/02699052.2013.793400
- Bigler, E. D. (2015). Neuroimaging as a biomarker in symptom validity and performance validity testing. *Brain Imaging and Behavior*, 9(3), 421-444. doi: 10.1007/s11682-015-9409-1

- Bigler, E. D. (2008). Neuropsychology and clinical neuroscience of persistent post-concussive syndrome. *Journal of the International Neuropsychological Society*, 14(1), 1-22. doi: 10.1017/S135561770808017X
- Bigler, E. D. (2012). Symptom validity testing, effort, and neuropsychological assessment. *Journal of the International Neuropsychological Society*, 18(4), 632-642. doi: 10.1017/S1355617712000252
- Bigler, E. D. (2013). Neuroimaging biomarkers in mild traumatic brain injury (mTBI). *Neuropsychological Review*, 23(3), 169-209. doi: 10.1007/s11065-013-9237-2
- Bigler, E. D. (2014). Effort, symptom validity testing, performance validity testing and traumatic brain injury. *Brain Injury*, 28(13-14), 1623-1638. doi: 10.3109/02699052.2014.947627
- Binder, L. M., Iverson, G. L., & Brooks, B. L. (2009). To err is human: "Abnormal" neuropsychological scores and variability are common in healthy adults. *Archives of Clinical Neuropsychology*, 24(1), 31-46. doi: 10.1093/arclin/acn001
- Blake, D. D., Weathers, F. W., Nagy, L. M., Kaloupek, D. G., Gusman, F. D., Charney, D. S., & Keane, T. M. (1995). The development of a clinician-administered PTSD scale. *Journal of Traumatic Stress*, 8(1), 75-90. doi: 10.1002/jts.2490080106
- Board of Directors (2007). American Academy of Clinical Neuropsychology (AACN) practice guidelines for neuropsychological assessment and consultation. *The Clinical Neuropsychologist*, 21(2), 209-231. doi: 10.1080/13825580601025932
- Boone, K. B. (2007). A reconsideration of the Slick et al. (1999) criteria for malingered neurocognitive dysfunction. In K. B. Boone (Eds.), *Assessment of Feigned*

Cognitive Impairment: A Neuropsychological Perspective. 29-49. New York, NY: Guildford Press.

- Boone, K. B. (2009). The need for continuous and comprehensive sampling of effort/response bias during neuropsychological examinations. *The Clinical Neuropsychologist*, 23(4), 729-741. doi: 10.1080/13854040802427803
- Boone, K. B. (2002). *The Dot Counting Test: Manual*. Los Angeles: Western Psychological Services.
- Boone, K. B., Lu, P., & Herzberg, D. S. (2002). *The b Test Manual*. Los Angeles, CA: Western Psychological Services.
- Boone, K. B., Victor, T. L., Wen, J., Razani, J., & Ponton, M. (2007). The association between neuropsychological scores and ethnicity, language, and acculturation variables in a large patient population. *Archives of Clinical Neuropsychology*, 22(3), 355-365. doi: 10.1016/j.acn.2007.01.010
- Bortnik, K. E., Boone, K. B., Marion, S. D., Amano, S., Ziegler, E., Victor, T. L., & Zeller, M. A. (2010). Examination of various WMS-III logical memory scores in the assessment of response bias. *The Clinical Neuropsychologist*, 24(2), 344-357. doi: 10.1080/13854040903307268
- Brennan, A. M., Meyer, S., David, E., Pella, R., Hill, B. D., & Gouvier, W. D. (2009). The vulnerability to coaching across measures of effort. *The Clinical Neuropsychologist*, 23(2), 314-328. doi: 10.1080/13854040802054151
- Briere, J. (2011). *Trauma Symptom Inventory, 2nd ed. (TSI-2) professional manual*. Lutz, FL: Psychological Assessment Resources.

- Brooks, B. L., & Ploetz, D. M. (2015). Embedded Performance Validity on the CVLT-C for Youth with Neurological Disorders. *Archives of Clinical Neuropsychology*, 30(3), 200-206. doi: 10.1093/arclin/acv017
- Brown, L., Sherbenou, R. J., & Johnson, S. K. (1997). *Test of Non-verbal Intelligence*. Austin, TX: Pro-Ed.
- Burton, V., Vilar-Lopez, R., & Puente, A.E. (2012). Measuring effort in neuropsychological evaluations of forensic cases of Spanish speakers. *Archives of Clinical Neuropsychology*, 27(3), 262-267. doi: 10.1093/arclin/acs026
- Calamia, M., Markon, K., & Tranel, D. (2013). The robust reliability of neuropsychological measures: Meta-analyses of test-retest correlations. *The Clinical Neuropsychologist*, 27(7), 1077-1105. doi: 10.1080/13854046.2013.809795
- Canadian Psychological Association (2000). *Canadian Code of Ethics for Psychologists, third edition*. Ottawa, ON: Canadian Psychological Association.
- Cantu, R. C., & Gean, A. D. (2010). Second-impact syndrome and a small subdural hematoma: An uncommon catastrophic result of repetitive head injury with a characteristic imaging appearance. *Journal of Neurotrauma*, 27(9), 1557-1564. doi: 10.1089/neu.2010.1334
- Carter-Allison, S. N., Potter, S., & Rimes, K. (2016). Diagnosis threat and injury beliefs after mild traumatic brain injury. *Archives of Clinical Neuropsychology*, 31(7), 727-737. doi: 10.1093/arclin/acw062
- Casas, R., Guzmán-Vélez, E., Cardona-Rodríguez, J., Rodríguez, N., Quiñones, G., Izaguirre, B., & Tranel, D. (2012). Interpreter-mediated neuropsychological

testing of monolingual Spanish speakers. *The Clinical Neuropsychologist*, 26(1), 88-101. doi: 10.1080/13854046.2011.640641

Cassidy, J. D., Cancelliere, C., Carroll, L. J., Cote, P., Hincapie, C. A., Holm, L. W., Hartvigsen, J., Donovan, J., Nygren-de Boussard, C., Kristman, V. L., & Borg, J. (2014). Systematic review of self-reported prognosis in adults after mild traumatic brain injury: Results of the International Collaboration on Mild Traumatic Brain Injury Prognosis. *Archives of Physical Medicine and Rehabilitation*, 95(3 Suppl 2), 132-151. doi:10.1016/j.apmr.2013.08.299

Castaneda, A. E., Tuulio-Henriksson, A., Marttunen, M., Suvisaari, J., & Lonnqvist, J. (2008). A review on cognitive impairments in depressive and anxiety disorders with a focus on young adults. *Journal of Affective Disorders*, 106(1-2), 1-27. doi: 10.1016/j.jad.2007.06.006

Carone, D. A. (2008). Children with moderate/severe brain damage/dysfunction outperform adults with mild-to-no brain damage on the Medical Symptom Validity Test. *Brain Injury*, 22(12), 960-971. doi: 10.1080/02699050802491297

Carone, D. A., Iverson, G. L., & Bush, S. S. (2010). A model to approaching and providing feedback to patients regarding invalid test performance in clinical neuropsychological evaluations. *The Clinical Neuropsychologist*, 24(5), 759-778. doi: 10.1080/13854041003712951

Carroll, L. J., Cassidy, J. D., Cancelliere, C., Cote, P., Hincapie, C. A., Kristman, V. L., Holm, L. W., Borg, J., Nygren-de Boussard, C., & Hartvigsen, J. (2014). Systematic review of the prognosis after mild traumatic brain injury in adults: Cognitive, psychiatric, and mortality outcomes: Results of the International

Collaboration on Mild Traumatic Brain Injury Prognosis. *Archives of Physical Medicine and Rehabilitation*, 95(3 Suppl 2), 152-173.

doi:10.1016/j.apmr.2013.08.300

Chan, A. S., & Kwok, I. (1999). *Hong Kong List Learning Test: Manual and Preliminary Norm*. Hong Kong: Chinese University of Hong Kong.

Chang, S. Y. S. (2006). *Development of a test battery for assessing memory malingering in Hong Kong and its application on depressed patients*. (Doctoral dissertation. Retrieved from Proquest Dissertation Publishing. (3254615)

Chen, A., Bushmeneva, K., Zagorski, B., Colantonio, A., Parsons, D., & Wodchis, W. P. (2012). Direct cost associated with acquired brain injury in Ontario. *BMC Neurology*, 12(1), 1-12. doi:10.1186/1471-2377-12-76

Cheung, B. Y., Chudek, M., & Heine, S. J. (2010). Evidence for a sensitive period for acculturation: Younger immigrants report acculturating at a faster rate. *Psychological Science*, 22(2), 147-152. doi: 10.1177/0956797610394661

Chiu, H. F. K., Lee, H. B., Chung, W. S., & Kwong, P. K. (1994). Reliability and validity of the Cantonese version of Mini-Mental State Examination: A preliminary report. *Journal of Hong Kong College of Psychiatrists*. 4(2), 25-28. Retrieved from <http://dspace.lib.cuhk.edu.hk/handle/2006/25085>

Chiu, V. Y., & Lee, T. C. (2002). Detection of malingering behavior at different levels of task difficulty in Hong Kong Chinese. *Rehabilitation Psychology*, 47(2), 194-203. doi: 10.1037/0090-5550.47.2.194

Chudek, M., Cheung, B. Y., & Heine, S. J. (2015). US immigrants' patterns of acculturation are sensitive to their age, language, and cultural contact but show no

evidence of a sensitive window for acculturation. *Journal of Cognition and Culture*, 15(1-2), 174-190. doi: 10.1163/15685373-12342145

Cirlugea, O. (2014). Assessing effort and malingering with the Asian client. In: Benuto, L. T., Thaler, N., & Leany, B. D. (Eds.), *Guide to psychological assessment with Asians*. 165-180. New York, NY: Springer.

Citizenship Act, Revised Statutes of Canada (1985 c. C-29). Retrieved from <http://laws.justice.gc.ca/eng/acts/C-29/page-2.html#docCont>

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*, 2nd Edition. Hillsdale, NJ: Erlbaum.

Colantonio, A., Saverine, C., Zagorski, B., Swaine, B., Lewko, J., Jaglal, S., & Vernich, L. (2010). Hospitalizations and emergency department visits for TBI in Ontario. *Canadian Journal of Neurological Science*, 37(6), 783-790. doi: 10.1017/s0317167100051441

College of Psychologists of Ontario (2017). *Standards of Professional Conduct*. Toronto, ON: Author.

Considine, C. M., Weisenbach, S. L., Walker, S. J., McFadden, E. M., Franti, L. M., Bieliauskas, L. A., Maixner, D. F., Giordani, B., Berent, S., & Langenecker, S. A. (2011). Auditory memory decrements, without dissimulation, among patients with major depressive disorder. *Archives of Clinical Neuropsychology*, 26(5), 445-453. doi: 10.1093/arclin/acr041

Constantinou, M. & McCaffrey, R. J. (2003). Using the TOMM for evaluating children's effort to perform optimally on neuropsychological measures. *Child Neuropsychology*, 9(2), 81-90. doi: 10.1076/chin.9.2.81.14505

- Cottingham, M. E., Victor, T. L., Boone, K. B., Ziegler, E. A., & Zeller, M. (2014). Apparent effect of type of compensation seeking (disability versus litigation) on performance validity test scores may be due to other factors. *The Clinical Neuropsychologist*, 28(6), 1030-1047. doi: 10.1080/13854046.2014.951397
- Curtis, K. L., Greve, K. W., & Bianchini, K. J. (2009). The Wechsler Adult Intelligence Scale—III and malingering in traumatic brain injury. *Assessment*, 16(4), 401-414. doi: 10.1177/1073191109338161
- Curtis, K. L., Thompson, L. K., Greve, K. W., & Bianchini, K. J. (2008). Verbal fluency indicators of malingering in traumatic brain injury: Classification accuracy in known groups. *The Clinical Neuropsychologist*, 22(5), 930-945. doi: 10.1080/13854040701563591
- Dekeyser, R., Alfi-Shabtay, I., & Ravid, D. (2010). Cross-linguistic evidence for the nature of age effects in second language acquisition. *Applied Psycholinguistics*, 31(3), 413-438. doi: 10.1017/S0142716410000056
- De Leon, M. B., Kirsch, N. L., Maio, R. F., Tan-Schriner, C. U., Millis, S. R., Frederiksen, S., Tanner, C. L., & Breer, M. L. (2009). Baseline predictors of fatigue 1 year after mild head injury. *Archives of Physical Medicine and Rehabilitation*, 90(6), 956-965. doi:10.1016/j.apmr.2008.12.016
- Delis, D. C., Kramer, J. J., Kaplan, E., & Ober, B. (2000). *California verbal learning test – Second Edition*. San Antonio, TX: Psychological Corporation.
- Delis, D. C., Kramer, J. J., Kaplan, E., & Ober, B. (1987). *California verbal learning test*. San Antonio, TX: Psychological Corporation.

- DiCarlo, M. A., Gfeller, J. D., Oliveri, M.V. (2000). Effects of coaching on detecting feigned cognitive impairment with the Category Test. *Archives of Clinical Neuropsychology*, 15(5), 399-413. doi: 10.1016/S0887-6177(99)00031-1
- Donders, J. (2005). Performance on the Test of Memory Malinger in a mixed pediatric sample. *Child Neuropsychology*, 11(2), 221-227. doi: 10.1080/09297040490917298
- Donders, J., & Boonstra, T. (2007). Correlates of invalid neuropsychological test performance after traumatic brain injury. *Brain Injury*, 21(3), 319-316. doi: 10.1080/02699050701253129
- Donders, J., & Strong, C. A. H. (2011). Embedded effort indicators on the California Verbal Learning Test – Second Edition (CVLT-II): An attempted cross-validation. *The Clinical Neuropsychologist*, 25(1), 173-184. doi: 10.1080/13854046.2010.536781
- Drob, S. L., Meehan, K. B., & Waxman, S. E. (2009). Clinical and conceptual problems in the attribution of malingering in forensic evaluations. *Journal of the American Academy of Psychiatry and the Law*, 37, 98-106. Retrieved from <http://www.jaapl.org/content/37/1/98.abstract>
- Education Testing Service (2016). About the TOEFL test. Retrieved from <http://www.ets.org/toefl/institutions/about>.
- Eglit, G. M. L., Lynch, J. K., & McCaffrey, R. J. (2016). Not all performance validity tests are created equal: The role of recollection and familiarity in the Test of Memory Malinger and Word Memory Test. *Journal of Clinical and*

Experimental Neuropsychology, 39(2), 173-189. doi:

10.1080/13803395.2016.1210573

Elbers, N. A., Hulst, L., Cuijpers, P., Akkermans, A. J., & Bruinvels, D. J. (2013). Do compensation processes impair mental health? A meta-analysis. *Injury*, 44(5), 674-683. doi: 10.1016/j.injury.2011.11.025

Elbulok-Charcape, M. M., Rabin, L. A., Spadaccini, A. T., & Barr, W. B. (2014). Trends in the neuropsychological assessment of ethnic/racial minorities: A survey of clinical neuropsychologists in the United States and Canada. *Cultural Diversity and Ethnic Minority Psychology*, 20(3), 353-361. doi: 10.1037/a0035023

Erdodi, L. A., Abeare, C. A., Lichtenstein, J. D., Tyson, B. T., Kucharski, B., Zuccato, B. G., & Roth, R. M. (2017). Wechsler Adult Intelligence Scale – Fourth Edition (WAIS-IV) processing speed scores as measures of noncredible responding: The third generation of embedded performance validity indicators. *Psychological Assessment*, 29(2), 148-157. doi: 10.1037/pas0000319

Erdodi, L. A., Jongsma, K. A., & Issa, M. (2017). The 15-Item Version of the Boston Naming Test as an Index of English Proficiency. *The Clinical Neuropsychologist*, 31(1), 168-178. doi: 10.1080/13854046.2016.1224392

Erdodi, L. A., Kirsch, N. L., Lajiness-O'Neill, R., Vingilis, E., & Medoff, B. (2014). Comparing the recognition memory test and the word choice test in a mixed clinical sample: Are they equivalent? *Psychological Injury and Law*, 7(3), 255-263. doi: 10.1007/s12207-014-9197-8

- Erdodi, L. A., & Lichtenstein, J. D. (2017). Invalid before impaired: An emerging paradox of embedded validity indicators. *The Clinical Neuropsychologist*, Advance online publication. doi: 10.1080/13854046.2017.1323119
- Erdodi, L. A., Nussbaum, S., Sagar, S., Abeare, C. A., & Schwartz, E. A. (2017). Limited English proficiency increases failure rates on performance validity tests with high verbal mediation. *Psychological Injury & Law*, 10(1), 96-103. doi: 10.1007/s12207-017-9282-x
- Erdodi, L. A., & Rai, J. K. (2017). A single error is one too many: Examining alternative cutoffs on Trial 2 of the TOMM, *Brain Injury*, Advance online publication. doi: 10.1080/02699052.2017.1332386
- Erdodi, L. A., & Roth, R. (2017). Low scores on BDAE Complex Ideational Materials associated with invalid performance in adults without aphasia. *Applied Neuropsychology: Adult*, 24(3), 264-274. doi: 10.1080/23279095.2016.1154856
- Erdodi, L. A., Roth, R. M., Kirsch, N. L., Lajiness-O'Neill, R., & Medoff, B. (2014). Aggregating validity indicators embedded in Conners' CPT-II outperforms individual cutoffs at separating valid from invalid performance in adults with traumatic brain injury. *Archives of Clinical Neuropsychology*, 29(5), 456-466. doi:10.1093/arclin/acu026
- Erdodi, L.A., Seke, K. R., Shahein, A., Tyson, B. T., Sagar, S., & Roth, R. M. (2017). Low scores on the Grooved Pegboard Test are associated with invalid responding and psychiatric symptoms. *Psychology & Neuroscience*, 10(3), 325-344. doi: 10.1037/pne0000103

- Erdodi, L. A., Tyson, B. T., Abeare, C. A., Lichtenstein, J. D., Pelletier, C. L., Rai, J. K., & Roth, R. M. (2016). The BDAE complex ideational material – a measure of receptive language or performance validity? *Psychological Injury and Law, 9*(2), 112-120. doi: 10.1007/s12207-016-9254-6
- Erdodi, L. A., Tyson, B. T., Shahein, A. G., Lichtenstein, J. D., Abeare, C. A., Pelletier, C. L., Zuccato, B. G., Kucharski, B., & Roth, R. M. (2017). The power of timing: Adding a time-to-completion cutoff to the Word Choice Test and Recognition Memory Test improves classification accuracy. *Journal of Clinical and Experimental Neuropsychology, 39*(4), 369-383. doi: 10.1080/13803395.2016.1230181
- Etherton, J. L., Bianchini, K. J., Ciota, M. A., & Greve, K. W. (2005). Reliable Digit Span is unaffected by laboratory-induced pain: Implications for clinical use. *Assessment, 12*(1), 101–106. doi: 10.1177/1073191104270789
- Etherton, J. L., Bianchini, K. J., Greve, K. W., & Ciota, M. A. (2005). Test of memory malingering performance is unaffected by laboratory-induced pain: Implications for clinical use. *Archives of Clinical Neuropsychology, 20*(3), 375–384. doi: 10.1016/j.acn.2004.09.007
- Etherton, J. L., Bianchini, K. J., Heinly, M. T., & Greve, K. W. (2006). Pain, malingering, and performance on the WAIS-III Processing Speed Index. *Journal of Clinical and Experimental Neuropsychology, 28*(7), 1218-1237. doi: 10.1080/13803390500346595
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical

sciences. *Behavior Research Methods*, 39(2), 175-191. doi:
10.3758/BF03193146

Field, A. (2009). *Discovering statistics using SPSS – Third edition*. Thousand Oaks, CA:
Sage.

Financial Services Commission of Ontario (2011). Auto insurance regulatory oversight.
Retrieved from

<http://www.auditor.on.ca/en/content/annualreports/arreports/en13/401en13.pdf>

Fitzpatrick, A. L., Rapp, S. R., Luchsinger, J., Hill-Briggs, F., Alonso, A., Gottesman, R.,
Lee, H., Camethon, M., Liu, K., Williams, K., Sharrett, A. R., Frazier-Wood, A.,
Lyketsos, C., & Seeman, T. (2015). Sociodemographic correlates of cognition in
the Multi-Ethnic Study of Atherosclerosis (MESA). *The American Journal of
Geriatric Psychiatry*, 23(7), 684-697. doi: 10.1016/j.jagp.2015.01.003

Gaetz, M. (2004). The neurophysiology of brain injury. *Clinical Neurophysiology*,
115(1), 4-18. doi: 10.1016/s1388-2457(03)00258-x

Gasquoine, P. G., Croyle, K. L., Cavazos-Gonzalez, C., & Sandoval, O. (2007).
Language of administration and neuropsychological test performance in
neurologically intact Hispanic American bilingual adults. *Archives of Clinical
Neuropsychology*, 22(8), 991-1001. doi: 10.1016/j.acn.2007.08.003

Gervais, R. O., Green, P., Allen, L. M., & Iverson, G. L. (2001). Effects of coaching on
symptom validity testing in chronic pain patients presenting for disability
assessments. *Journal of Forensic Neuropsychology*, 2(2), 1-19. doi:
10.1300/j151v02n02_01

- Gervais, R. O., Rohling, M. L., Green, P., & Ford, W. (2004). A comparison of the WMT, CRAB, and TOMM failure rates in non-head injury disability claims. *Archives of Clinical Neuropsychology, 19*(4), 475-487. doi: 10.1016/j.acn.2003.05.001
- Giesbrecht, T., Lynn, S. J., Lilienfeld, S. O., & Merckelbach, H. (2008). Cognitive processes in dissociation: An analysis of core theoretical assumptions. *Psychological Bulletin, 134*(5), 617-647. doi: 10.1037/0033-2909.134.5.617
- Goldberg, H. E., Back-Madruga, C., Boone, K. B. (2007). The impact of psychiatric disorders on cognitive symptom validity test scores. In K. B. Boone (Eds.), *Assessment of Feigned Cognitive Impairment: A Neuropsychological Perspective*. 281-309. New York, NY: Guildford Press.
- Golden, C. J., & Freshwater, S. M. (2002). *Stroop Color and Word Test: Revised examiner's manual*. Wood Dale, IL: Stoelting Co.
- Gray, M. J., Elhai, J. D., & Briere, J. (2010). Evaluation of the Atypical Response scale of the Trauma Symptom Inventory-2 in detecting simulated posttraumatic stress disorder. *Journal of Anxiety Disorders, 24*(5), 447-451. doi: 10.1016/j.janxdis.2010.02.011
- Green, P., Allen, L. M., & Astner, K. (2003). *The word memory test*. Edmonton, AB: Author.
- Green, P. (2004). *Green's Medical Symptom Validity Test (MSVT): User's manual*. Edmonton, AB: Author.
- Green, P. (2007a). *Manual for the Nonverbal Medical Symptom Validity Test*. Edmonton, AB: Author.

- Green, P. (2007b). The pervasive influence of effort on neuropsychological tests. *Physical Medicine and Rehabilitation Clinics of North America*, 18(1), 43–68. doi: 10.1016/j.pmr.2006.11.002
- Green, P. (2013). Spoiled for choice: Making comparisons between forced choice effort tests. In K. B. Boone (Ed.), *Clinical practice of forensic neuropsychology*. New York, NY: Guilford Press.
- Green, P., Iverson, G. L., & Allen, L. M. (1999). Detecting malingering in head injury litigation with the Word Memory Test. *Brain Injury*, 13(10), 813-819. doi: 10.1080/026990599121205
- Green, P., Rohling, M. L., Lees-Haley, P. R., & Allen, L. (2001). Effort has a greater effect on test scores than severe brain injury in compensation claims. *Brain Injury*, 15(12), 1045-1060. doi: 10.1080/02699050110088254
- Green, R. E. A., Colella, B., Maller, J. J., Bayley, M., Glazer, J., & Mikulis, D. J. (2014). Scale and pattern of atrophy in the chronic stages of moderate-severe TBI. *Frontiers in Human Neuroscience*, 8, Article 67. doi: 10.3389/fnhum.2014.00067
- Greiffenstein, M. F., & Baker, W. J. (2008). Validity testing in dually diagnosed posttraumatic stress disorder and mild closed head injury. *The Clinical Neuropsychologist*, 22(3), 565-582. doi: 10.1080/13854040701377810
- Greiffenstein, M. F., Greve, K. W., Bianchini, K. J., & Baker, W. J. (2008). Test of Memory Malingering and Word Memory Test: A new comparison of failure concordance. *Archives of Clinical Neuropsychology*, 23(7-8), 801-807. doi: 10.1016/j.acn.2008.07.005

- Greve, K. W., & Bianchini, K. J. (2004). Setting empirical cutoffs on psychometric indicators of negative response bias: A methodological commentary with recommendations. *Archives of Clinical Neuropsychology, 19*(4), 533-541. doi: 10.1016/j.acn.2003.08.002
- Greve, K. W., Bianchini, K. J., & Brewer, S. T. (2013). The assessment of performance and self-report validity in persons claiming pain-related disability. *The Clinical Neuropsychologist, 27*(1), 108-137. doi: 10.1080/13854046.2012.739646
- Greve, K. W., Bianchini, K. J., & Doane, B. M. (2006). Classification accuracy of the Test of Memory Malingering in traumatic brain injury: Results of a known-groups analysis. *Journal of Clinical and Experimental Neuropsychology, 28*(7), 1176-1190. doi: 10.1080/13803390500263550
- Greve, K. W., Bianchini, K. J., Mathias, C. W., Houston, R. J., & Crouch, J. A. (2003). Detecting malingered performance on the Wechsler Adult Intelligence Scale validation of Mittenberg's approach in traumatic brain injury. *Archives of Clinical Neuropsychology, 18*(3), 245-260. doi: 10.1016/s0887-6177(02)00137-3
- Greve, K. W., Curtis, K. L., Bianchini, K. J., & Ord, J. S. (2009). Are the original and second edition of the California Verbal Learning Test equally accurate in detecting malingering? *Assessment, 16*(3), 237-248. doi: 10.1177/1073191108326227
- Greve, K. W., Ord, J., Curtis, K. L., Bianchini, K. J., & Brennan, A. (2008). Detecting malingering in traumatic brain injury and chronic pain: A comparison of three forced-choice symptom validity tests. *The Clinical Neuropsychologist, 22*(5), 896-918. doi: 10.1080/13854040701565208

- Groleau, D., & Kirmayer, L. (2004). Sociosomatic theory in Vietnamese immigrants' narratives of distress. *Anthropology and Medicine, 11*(2), 117-134. doi: 10.1080/13648470410001678631
- Grote, C. L., Kooker, E. K., Garron, D. C., Nyenhuis, D. L., Smith, C. A., & Mattingly, M. L. (2000). Performance of compensation seeking and non-compensation seeking samples on the Victoria Symptom Validity Test: Cross-validation and extension of a standardization study. *Journal of Clinical and Experimental Neuropsychology, 22*(6), 709-719. doi: 10.1076/jcen.22.6.709.958
- Grysiewicz, R. A., Thomas, K., & Pandey, D. K. (2008). Epidemiology of ischemic and hemorrhagic stroke: Incidence, prevalence, mortality, and risk factors. *Neurologic Clinics, 26*(4), 871-895. doi: 10.1016/j.ncl.2008.07.003
- Guise, B. J., Thompson, M. D., Greve, K. W., Bianchini, K. J., & West, L. (2014). Assessment of performance validity in the Stroop Color and Word Test in mild traumatic brain injury patients: A criterion-groups validation design. *Journal of Neuropsychology, 8*(1), 20-33. doi: 10.1111/jnp.12002
- Gunstad, J., & Suhr, J. A. (2001). "Expectation as etiology" versus "the good old days": Postconcussion syndrome symptom reporting in athletes, headache sufferers, and depressed individuals. *Journal of the International Neuropsychological Society, 7*(3), 323-333. doi: 10.1017/s1355617701733061
- Haaland, V. O., Landro, N. I. (2009). Pathological dissociation and neuropsychological functioning in borderline personality disorder. *Acta Psychiatrica Scandinavica, 119*(5), 383-392. doi: 10.1111/j.1600-0447.2008.01323.x

- Haber, A. H., & Fichtenberg, N. L. (2006). Replication of the Test of Memory Malingering (TOMM) in a traumatic brain injury and head injury sample. *The Clinical Neuropsychologist*, 20(3), 524-532. doi: 10.1080/13854040590967595
- Hampson, N. E., Kemp, S., Coughlan, A. K., Moulin, C. J. A., & Bhakta, B. B. (2014). Effort test performance in clinical acute brain injury, community brain injury, and epilepsy populations. *Applied Neuropsychology: Adult*, 21(3), 183-194. doi: 10.1080/09084282.2013.787425
- Hanlon, R. E., Demery, J. A., Martinovich, Z., & Kelly, J. P. (1999). Effects of acute injury characteristics on neuropsychological status and vocational outcome following mild traumatic brain injury. *Brain Injury*, 13(11), 873-887. doi: 10.1080/026990599121070
- Heaton, R. K., Miller, S. W., Taylor, M. J., & Grant, I. (2004). *Revised comprehensive norms for an expanded Halstead-Reitan battery: Demographically adjusted neuropsychological norms for African American and Caucasian Adults*. Odessa, FL: Psychological Assessment Resources, Inc.
- Heinly, M. T., Greve, K. W., Bianchini, K. J., Love, J. M., & Brennan, A. (2005). WAIS Digit Span-based indicators of malingered neurocognitive dysfunction: Classification accuracy in traumatic brain injury. *Assessment*, 12(4), 429-444. doi: 10.1177/1073191105281099
- Hernandez, A. E., & Li, P. (2007). Age of acquisition: Its neural and computational mechanisms. *Psychological Bulletin*, 133(4), 638-650. doi: 10.1037/0033-2909.133.4.638

- Hesdorffer, D. C., Rauch, S. L., & Tamminga, C. A. (2009). Long-term psychiatric outcomes following traumatic brain injury. *Journal of Head Trauma Rehabilitation, 24*(6), 452-459. doi: 10.1097/htr.0b013e3181c133fd
- Hiscock, M., & Hiscock, C. K. (1989). Refining the forced-choice method for the detection of malingering. *Journal of Clinical and Experimental Neuropsychology, 11*(6), 967-974. doi: 10.1080/01688638908400949
- Holm, L., Cassidy, J. D., Carroll, L. J., & Borg, J. (2005). Summary of the WHO collaborating centre for neurotrauma task force on mild traumatic brain injury. *Journal of Rehabilitation Medicine, 37*(3), 137-141. doi: 10.1080/16501970510027321
- Howe, L. L., Anderson, A. M., Kaufman, D. A., Sachs, B. C., & Loring, D. W. (2007). Characterization of the Medical Symptom Validity Test in evaluation of clinically referred memory disorders clinic patients. *Archives of Clinical Neuropsychology, 22*(6), 753-761. doi: 10.1016/j.acn.2007.06.003
- Hua, M. S., Chang, B. S., Lin, K. N., Yang, C. M., Lu, L. H. J., & Chen, H. Y. (2005). *Wechsler Memory Scale- III (Chinese version): Administration and scoring manual*. Taipei: Chinese Behavioral Science Corporation.
- Hughes, D. G., Jackson, A., Mason, D. L., Berry, E., Hollis, S., & Yates, D. W. (2004). Abnormalities on magnetic resonance imaging seen acutely following mild traumatic brain injury: Correlation with neuropsychological tests and delayed memory. *Neuroradiology, 46*(7), 550-558. doi: 10.1007/s00234-004-1227-x
- Hunt, S., Root, J. C., & Bascetta, B. L. (2014). Effort testing in schizophrenia and schizoaffective disorder: Validity Indicator Profile and Test of Memory

- Malingering performance characteristics. *Archives of Clinical Neuropsychology*, 29(2), 164-172. doi: 10.1093/arclin/act069
- Ishikawa, H., & Kiuchi, T. (2010). Health literacy and health communication. *BioPsychoSocial Medicine*, 4(1), 18-22. doi:10.1186/1751-0759-4-18
- Isokuortti, H., Iverson, G. L., Kataja, A., Brander, A., Ohman, J., & Luoto, T. M. (2016). Who gets head trauma or recruited in mild traumatic brain injury research? *Journal of Neurotrauma*, 33(2), 232-241. doi: 10.1089/neu.2015.3888
- Iverson, G. L., (2006). Ethical issues associated with the assessment of exaggeration, poor effort, and malingering. *Applied Neuropsychology*, 13(2), 77-90. doi: 10.1207/s15324826an1302_3
- Iverson, G. L. (2005). Outcome from mild traumatic brain injury. *Current Opinion in Psychiatry*, 18(3), 301-317. doi: 10.1097/01.yco.0000165601.29047.ae
- Iverson, G. L., Lange, R. T., Green, P., & Franzen, M. D. (2002). Detecting exaggeration and malingering with the Trail Making Test. *The Clinical Neuropsychologist*, 16(3), 398-406. doi: 10.1076/clin.16.3.398.13861
- Iverson, G. L., & Tulskey, D. S. (2003). Detecting malingering on the WAIS-III unusual Digit Span performance patterns in the normal population and in clinical groups. *Archives of Clinical Neuropsychology*, 18(1), 1-9. doi: 10.1093/arclin/18.1.1
- James, B. D., & Schneider, J. A. (2010). Increasing incidence of dementia in the oldest old: Evidence and implications. *Alzheimer's Research & Therapy*, 2(3), 9. doi: 10.1186/alzrt32

- Jamora, C. W., Schroeder, S. C., & Ruff, R. M. (2013). Pain and mild traumatic brain injury: The implications of pain severity on emotional and cognitive functioning. *Brain Injury, 27*(10), 1134-1140. doi: 10.3109/02699052.2013.804196
- Jamora, C. W., Young, A., & Ruff, R. M. (2012). Comparison of subjective cognitive complaints with neuropsychological tests in individuals with mild vs more severe traumatic brain injuries. *Brain Injury, 26*(1), 36-47. doi: 10.3109/02699052.2011.635352
- Jasinski, L. J., Berry, D. T. R., Shandera, A. L., & Clark, J. A. (2011). Use of the Wechsler Adult Intelligence Scale Digit Span subtest for malingering detection: A meta-analytic review. *Journal of Clinical and Experimental Neuropsychology, 33*(3), 300-314. doi: 10.1080/13803395.2010.516743
- Jia, F., Gottardo, A., Chen, X., Koh, P., & Pasquarella, A. (2016). English proficiency and acculturation among Chinese immigrant youth in Canada: a reciprocal relationship. *Journal of Multilingual and Multicultural Development, 37*(8), 774-782. doi: 10.1080/01434632.2015.1133630
- Jones, A. (2013). Test of Memory Malingering: Cutoff scores for psychometrically defined malingering groups in a military sample. *The Clinical Neuropsychologist, 27*(6), 1043-1059. doi: 10.1080/13854046.2013.804949
- Karr, J. E., Areshenkoff, C. N., & Garcia-Barrera, M. A. (2014). The neuropsychological outcomes of concussion: A systematic review of meta-analyses on the cognitive sequelae of mild traumatic brain injury. *Neuropsychology, 28*(3), 321-336. doi: 10.1037/neu0000037

- Kashluba, S., Hanks, R. A., Casey, J. E., & Millis, S. R. (2008). Neuropsychological and functional outcome after complicated mild traumatic brain injury. *Archives of Physical Medicine and Rehabilitation*, 89(5), 904-1011. doi: 10.1016/j.apmr.2007.12.029
- Kaufman, D. M., & Milstein, M. J. (2007). *Kaufman's clinical neurology for psychiatrists* (7th edition). New York, NY: Elsevier.
- King, N. S., Crawford, S., Wender, F. J., Moss, N. E. G., & Wade, D. T. (1995). The Rivermead Post Concussion Symptoms Questionnaire: a measure of symptoms commonly experienced after head injury and its reliability. *Journal of Neurology*, 242(9), 587-592. doi: 10.1007/bf00868811
- Kirmayer, L. J., Groleau, D., Looper, K. J., & Dao, M. D. (2004). Explaining medically unexplained symptoms. *Canadian Journal of Psychiatry*, 49(10), 663-671. doi: 10.1177/070674370404901003
- Kirmayer, L. J., & Sartorius, N. (2008). Cultural models and somatic syndromes. *Psychosomatic Medicine*, 69(9), 832-840. doi: 10.1097/PSY.0b013e31815b002c
- Labarge, A. S., McCaffrey, R. J., & Brown, T. A. (2003). Neuropsychologists' abilities to determine the predictive value of diagnostic tests. *Archives of Clinical Neuropsychology*, 18(2), 165-175. doi: 10.1093/arclin/18.2.165
- Lam, M., Eng, G. K., Rapisarda, A., Subramaniam, M., Kraus, M., Keefe, R., & Collinson, S. L. (2013). Formulation of the age-education index: Measuring age and education effects in neuropsychological performance. *Psychological Assessment*, 25(1), 61-70. doi: 10.1037/a0030548

- Lange, R. T., Iverson, G. L., Brooks, B. L., & Rennison, L. A. (2010). Influence of poor effort on self-reported symptoms and neurocognitive test performance following mild traumatic brain injury. *Journal of Clinical and Experimental Neuropsychology*, 32(9), 961-972. doi: 10.1080/13803391003645657
- Lange, R. T., Iverson, G. L., & Rose, A. (2010). Postconcussion symptom reporting and the “good-old-days bias” following mild traumatic brain injury. *Archives of Clinical Neuropsychology*, 25(5), 442-450. doi: 10.1093/arclin/acq031
- Langlois Orman, J. A., Kraus, J. F., Zaloshnja, E., & Miller, T. (2011). Epidemiology. In J. M. Silver, T. W. McAllister, & S. C. Yudofsky (Eds.), *Textbook of Traumatic Brain Injury Second Edition*. 3-22. Washington, DC: American Psychiatric Publishing, Inc.
- Lanius, R. A., Vermetten, R., Loewenstein, R. J., Brand, B., Schmahl, C., Bremner, J. D., & Spiegel, D. (2010). Emotion modulation in PTSD: Clinical and neurobiological evidence for a dissociative subtype. *American Journal of Psychiatry*, 167(6), 640-647. doi: 10.1176/appi.ajp.2009.09081168
- La Parra-Cassado, D., Stornes, P., & Solheim, E. F. (2017). Self-rated health and wellbeing among the working-age immigrant population in Western Europe: Findings from the European social survey (2014) special module on the social determinants of health. *European Journal of Public Health*, 27(suppl 1), 40-46. doi: 10.1093/eurpub/ckw221
- Larrabee, G. J. (2003). Detection of malingering using atypical performance patterns on standard neuropsychological tests. *The Clinical Neuropsychologist*, 17(3), 410-425. doi: 10.1076/clin.17.3.410.18089

- Larrabee, G. J. (2008). Aggregation across multiple indicators improves the detection of malingering: Relationship to likelihood ratios. *The Clinical Neuropsychologist*, 22(4), 666-679. doi: 10.1080/13854040701494987
- Larrabee, G. J. (2012). Performance validity and symptom validity in neuropsychological assessment. *Journal of the International Neuropsychological Society*, 18(4), 625-631. doi: 10.1017/s1355617712000240
- Larrabee, G. J. (2014). False-positive rates associated with the use of multiple performance and symptom validity tests. *Archives of Clinical Neuropsychology*, 29(4), 364–373. doi: 10.1093/arclin/acu019
- Larrabee, G. J., Binder, L. M., Rohling, M. L., & Ploetz, D. M. (2013). Meta-analytic methods and the importance of non-TBI factors related to outcome in mild traumatic brain injury: Response to Bigler et al., 2013. *The Clinical Neuropsychologist*, 27(2), 215-237. doi: 10.1080/13854046.2013.769634
- Larrabee, G. J., Millis, S. R., & Meyers, J. E. (2009). 40 plus or minus 10, a new magical number: Reply to Russel. *The Clinical Neuropsychologist*, 23(5), 841-849. doi: 10.1080/13854040902796735
- Larrabee, G. J., & Rohling, M. L. (2013). Neuropsychological differential diagnosis of mild traumatic brain injury. *Behavioural Sciences and the Law*, 31(6), 686-701. doi: 10.1002/bsl.2087
- Le Bihan, D., Mangin, J. F., Poupon, C., Clark, C. A., Pappata, S., Molko, N., & Chabriat, H. (2001). Diffusion tensor imaging: Concepts and applications. *Journal of Magnetic Resonance Imaging*, 13(4), 534-546. doi: 10.1002/jmri.1076

- Lee, S., Nguyen, H. A., & Tsui, J. (2011). Interview language: A proxy measure for acculturation among Asian Americans in a population-based survey. *Journal of Immigrant and Minority Health, 13*(2), 244-252. doi: 10.1007/s10903-009-9278-z
- Lensvelt-Mulders, G., van der Hart, O., van Ochten, J. M., van Son, M. J. M., Steele, K., & Breeman, L. (2008). Relations among peritraumatic dissociation and posttraumatic stress: A meta-analysis. *Clinical Psychology Review, 28*(7), 1138-1151. doi: 10.1016/j.cpr.2008.03.006
- Lezak, M. D., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012), *Neuropsychological Assessment, 5th edition*. Oxford, UK: Oxford University Press.
- Lichtenstein, J. D., Erdodi, L. A., Rai, J. K., Mazur-Mosiewicz, A., & Flaro, L. (2016). Wisconsin Card Sorting Test embedded validity indicators developed for adults can be extended to children. *Child Neuropsychology*, Online Article. doi: 10.1080/09297049.2016.1259402
- Losoi, H., Silverberg, N. D., Waljas, M., Turunen, S., Rosti-Otajarvi, E., Helminen, M., Luoto, T. M., Jukunen, J., Ohman, J., & Iverson, G. L. (2016). Recovery from mild traumatic brain injury in previously healthy adults. *Journal of Neurotrauma, 33*(1), 1-11. doi: 10.1089/neu.2015.4070
- Lu, P. H., Boone, K. B., Cozolino, L., & Mitchell, C. (2003). Effectiveness of the Rey-Osterrieth Complex Figure Test and the Meyers and Meyers Recognition Trial in the detection of suspect effort. *The Clinical Neuropsychologist, 17*(3), 426-440. doi: 10.1076/clin.17.3.426.18083

- MacFarlane, M. P., & Glenn, T. C. (2015). Neurochemical cascade of concussion. *Brain Injury*, 29(2), 139-153. doi: 10.3109/02699052.2014.965208
- Manly, J. J., Byrd, D. A., Touradji, P., & Stern, Y. (2004). Acculturation, reading level, and neuropsychological test performance among African American elders. *Applied Neuropsychology*, 11(1), 37-46. doi: 10.1207/s15324826an1101_5
- Mantwill, S., & Schulz, P. J. (2016). Does acculturation narrow the health literacy gap between immigrants and non-immigrants—an explorative study. *Patient Education and Counseling*, 100(4), 760-767. doi: 10.1016/j.pec.2016.10.021
- Marcopulos, B. A., Caillouet, B. A., Bailey, C. M., Tussey, C., Kent, J. A., & Frederick, R. (2014). Clinical decision making in response to performance validity test failure in a psychiatric setting. *The Clinical Neuropsychologist*, 28(4), 633-652. doi: 10.1080/13854046.2014.896416
- Matsumoto, D., & Juang, L. (2016). *Culture and psychology*. Toronto, ON: Nelson Education.
- McAllister, T. W. (2011). Mild Brain Injury. In J. M. Silver, T. W. McAllister, & S. C. Yudofsky (Eds.), *Textbook of Traumatic Brain Injury 2nd Edition*. 239-264. Washington, DC: American Psychiatric Publishing, Inc.
- McClintock, S. M., Husain, M. M., Greer, T. L., & Cullum, C M. (2010). Association between depression severity and neurocognitive function in major depressive disorder: A review and synthesis. *Neuropsychology*, 24(1), 9-34. doi: 10.1037/a0017336
- McCormick, C. L., Yoash-Gantz, R. E., McDonald, S. D., Campbell, T. C., & Tupler, L. A. (2013). Performance on the Green Word Memory Test following Operation

Enduring Freedom/Operation Iraqi Freedom-era military service: Test failure is related to evaluation context. *Archives of Clinical Neuropsychology*, 28(8), 808-823. doi: 10.1093/arclin/act050

McKinnon, M. C., Boyd, J. E., Frewen, P. A., Lanius, U. F., Jetly, R., Richardson, J. D., & Lanius, R. A. (2016). A review of the relation between dissociation, memory, executive functioning and social cognition in military members and civilians with neuropsychiatric conditions. *Neuropsychologia*, 90, 210-234. doi: 10.1016/j.neuropsychologia.2016.07.017

Menon, D. K., Schwab, K., Wright, D. W., Maas, A. I., & The Demographics and Clinical Assessment Working Group of the International and Interagency Initiative toward Common Data Elements for Research on Traumatic Brain Injury and Psychological Health. (2010). Position statement: Definition of traumatic brain injury. *Archives of Physical Medicine and Rehabilitation*, 91(11), 1637-1640. doi: 10.1016/j.apmr.2010.05.017

Merckelbach, H., Giesbrecht, T., van Heugten-van der Kloet, D., de Jong, J., Meyer, T., & Rietman, K. (2015). The overlap between dissociative symptoms and symptom over-reporting. *The European Journal of Psychiatry*, 29(3), 165-172. doi: 10.4321/s0213-61632015000300001

Merckelbach, H., Horselenberg, R., & Schmidt, H. (2002). Modeling the connection between self-reported trauma and dissociation in a student sample. *Personality and Individual Differences*, 32(4), 695-705. doi: 10.1016/s0191-8869(01)00070-8

Merckelbach, H., Langeland, W., de Vries, G., & Draijer, N. (2014). Symptom overreporting obscures the dose-response relationship between trauma severity

and symptoms. *Psychiatry Research*, 217(3), 215-219. doi:
10.1016/j.psychres.2014.03.018

Merten, T., & Merckelbach, H. (2013). Symptom validity testing in somatoform and dissociative disorders: A critical review. *Psychological Injury and Law*, 6(2), 122-137. doi: 10.1007/s12207-013-9155-x

Merten, T., Thies, E., Schneider, K., & Stevens, A. (2009). Symptom validity testing in claimants with alleged posttraumatic stress disorder: Comparing the Morel Emotional Numbing Test, the Structured Inventory of Malingered Symptomatology, and the Word Memory Test. *Psychological Injury and the Law*, 2(3-4), 284-293. doi: 10.1007/s12207-009-9057-0

Meyers, J. E., & Meyers, K. R. (1995). *Rey Complex Figure Test and Recognition Trial: Professional Manual*. Lutz, FL: Psychological Assessment Resources.

Mindt, M. R., Byrd, D., Saez, P., & Manly, J. (2010). Increasing culturally competent neuropsychological services for ethnic minority populations: A call to action. *The Clinical Neuropsychologist*, 24(3), 429-453. doi: 10.1080/13854040903058960

Ministry of Finance (2014). National household survey highlights. *Office of Economic Policy Labour and Demographic Analysis Branch*. Retrieved from <http://www.fin.gov.on.ca/en/economy/demographics/census/>

Mitrushina, M, Boone, K. B., Razani, J., & D'Elia, L. F. (2005). *Handbook of normative data for neuropsychological assessment, 2nd edition*. Oxford, UK: Oxford University Press.

- Mittenberg, W., DiGuilio, D., Perrin, S., & Bass, A. (1992). Symptoms following mild head injury: Expectation as aetiology. *Journal of Neurology, Neurosurgery, and Psychiatry*, 55(3), 200–204. doi: 10.1136/jnnp.55.3.200
- Mittenberg, W., Patton, C., Canyock, E. M., & Condit, D. C. (2002). Base rates of malingering and symptom exaggeration. *Journal of Clinical and Experimental Neuropsychology*, 24(8), 1094-1102. doi: 10.1076/jcen.24.8.1094.8379
- Mittenberg, W., Theroux-Fichera, S., Zielinski, R., & Heilbronner, R. L. (1995). Identification of malingered head injury on the Wechsler Adult Intelligence Scale—Revised. *Professional Psychology: Research and Practice*, 26(5), 491-498. doi: 10.1037/0735-7028.26.5.491
- Montese, A., Pere, V., Carme, J., Francese, V., & Eduardo, T. (2001). Visuospatial deficits in Parkinson's Disease assessed by Judgment of Line Orientation Test: Error analysis and practice effects. *Journal of Clinical and Experimental Neuropsychology*, 23(5), 592-598. doi: 10.1076/jcen.23.5.592.1248
- Moore, E. L., Terryberry-Spohr, L., & Hope, D. A. (2006). Mild traumatic brain injury and anxiety sequelae: A review of the literature. *Brain Injury*, 20(2), 117-132. doi: 10.1080/02699050500443558
- Morey, L. C. (1991). *Personality Assessment Inventory (PAI) Professional Manual*. Lutz, FL: Psychological Assessment Resources.
- Murgatroyd, D. F., Casey, P. P., Cameron, I. D., & Harris, I. A. (2015). The effect of financial compensation on health outcomes following musculoskeletal injury: Systematic review. *PLoS ONE*, 10(2), 1-33. doi: 10.1371/journal.pone.0117597

- Murray, J., Ehlers, A., & Mayou, R. A. (2002). Dissociation and posttraumatic stress disorder: Two prospective studies of road traffic accident survivors. *British Journal of Psychiatry, 180*(4), 363-368. doi: 10.1192/bjp.180.4.363
- Naim, R., Wald, I., Lior, A., Pine, D. S., Fox, N. A., Sheppes, G., Halpern, P., & Bar-Haim, Y. (2014). Perturbed threat monitoring following a traumatic event predicts risk for posttraumatic stress disorder. *Psychological Medicine, 44*(1), 2077-2084. doi: 10.1017/s0033291713002456
- Nampiaparampil, D. E. (2008). Prevalence of chronic pain after traumatic brain injury: A systematic review. *Journal of the American Medical Association, 300*(6), 711-719. doi: 10.1001/jama.300.6.711
- Nazarov, A., Frewen, P., Oremus, C., Schellenberg, E. G., Mckinnon, M. C., & Lanius, R. (2015). Comprehension of affective prosody in women with posttraumatic stress disorder related to childhood abuse. *Acta Psychiatrica Scandinavia 131*(5), 342-349. doi: 10.1111/acps.12364.
- Nicholson, A. A., Densmore, M., Frewen, P. A., Theberge, J., Neufeld, R. W. J., McKinnon, M. C., & Lanius, R. A. (2015). The dissociative subtype of posttraumatic stress disorder: Unique resting-state functional connectivity of basolateral and centromedial amygdala complexes. *Neuropsychopharmacology, 40*(10), 2317-2326. doi: 10.1038/npp.2015.79
- Nijdam-Jones, A., & Rosenfeld, B. (2017). Cross-cultural feigning assessment: A systematic review of feigning instruments used with linguistically, ethnically, and culturally diverse samples. *Psychological Assessment*. Advance online publication. doi: 10.1037/pas0000438

- Nogueras, D. J., Postma, J., & Van Son, C. (2016). Why didn't I know? Perspectives from adult children of elderly parents with dementia. *Journal of the American Association of Nurse Practitioners*, 28(12), 668-674. doi: 10.1002/2327-6924.12382
- Nuechterlein, K. H., Green, M. F., Kern, R. S., Baade, L. E., Barch, D. M., Cohen, J. D., ... Marder, S. R. (2008). The MATRICS consensus cognitive battery, part 1: Test selection, reliability, and validity. *American Journal of Psychiatry*, 165(2), 203-213. doi: 10.1176/appi.ajp.2007.07010042
- O'Bryant, S. E., Finlay, C. G., & O'Jile, J. R. (2007). TOMM performances and self-reported symptoms of depression and anxiety. *Journal of Psychopathological Behavioural Assessment*, 29(2), 111-114. doi: 10.1007/s10862-006-9034-9
- Odland, A. P., Lammy, A. B., Martin, P. K., Grote, C. L., & Mittenberg, W. (2015). Advanced administration and interpretation of multiple validity tests. *Psychological Injury and Law*, 8(1), 46-63. doi: 10.1007/s12207-015-9216-4
- Ozdemir, O., Ozdemir, P. G., Boysan, M., & Yilmaz, E. (2015). The relationships between dissociation, attention, and memory dysfunction. *Archives of Neuropsychiatry*, 51(1), 36-41. doi: 10.5152/npa.2015.7390
- Ozen, L. J., & Fernandes, M. A. (2011). Effects of "diagnosis threat" on cognitive and affective functioning long after mild head injury. *International Neuropsychological Society*, 17(2), 219-229. doi: 10.1017/S135561771000144X
- Paniagua, F. A. (2005). *Assessing and treating culturally diverse clients: A practical guide, third edition*. Thousand Oaks, CA: Sage Publications Inc.

- Parlar, M., Frewen, P. A., Oremus, C., Lanius, R. A., & McKinnon, M. C. (2016). Dissociative symptoms are associated with reduced neuropsychological performance in patients with recurrent depression and a history of trauma exposure. *European Journal of Psychotraumatology*, 7(1), Article 29061. doi: 10.3402/ejpt.v7.29061
- Pearson, N. C. S. (2009). *Advanced clinical solutions for WAIS-IV and WMS-IV: Administration and scoring manual*. San Antonio, TX: The Psychological Corporation.
- Pertab, J. L., James, K. M., & Bigler, E. D. (2009). Limitations of mild traumatic brain injury meta-analyses. *Brain Injury*, 23(6), 498-508. doi: 10.1080/02699050902927984
- Ponsford, J., Willmott, C., Rothwell, A., Cameron, P., Kelly, A. M., Nelms, R., Curran, C., & Ng, K. (2000). Factors influencing outcome following mild traumatic brain injury in adults. *Journal of the International Neuropsychological Society*, 6(5), 568-579. doi: 10.1017/s1355617700655066
- Poreh, A. M., Avital, R., Dines, P. L., & Levin, J. B. (2015). The effects of age of language acquisition on verbal memory tests in a sample of older adults immigrants. *Psychology and Neuroscience*, 8(1), 66-74. doi: 10.1037/h0100350
- Prieto de Estebecorena, M. J. (2007). *Evaluating the role of education and acculturation in the performance of Hispanics on a non-verbal test*. (Doctoral dissertation). Retrieved from ProQuest Dissertations Publishing. (3255534)
- Qualls, C. E., Bliwise, N. G., & Stringer, A. Y. (2000). Short forms of the Benton judgment of line orientation test: Development and psychometric properties.

Archives of Clinical Neuropsychology, 15(2), 159–163. doi:
10.1093/arclin/15.2.159

Reedy, S. D., Boone, K. B., Cottingham, M. E., Glaser, D. F., Lu, P. H., Victor, T. L.,
Ziegler, E. A., Zeller, M. A., & Wright, M. J. (2013). Cross validation of the Lu et
al. (2003) Rey-Osterrieth Complex Figure Test effort equation in a large known-
group sample. *Archives of Clinical Neuropsychology*, 28(1), 30-37. doi:
10.1093/arclin/acs106

Rees, L. M., Tombaugh, T. N., Gansler, D. A., & Moczynski, N. P. (1998). Five
validation experiments of the Test of Memory Malingering (TOMM).
Psychological Assessment, 10(1), 10-20. doi: 10.1037/1040-3590.10.1.10

Reitan, R. M. (1969). *Manual for administration of neuropsychological test batteries for
adults and children*. Indianapolis, IN.

Reitan, R. M. (1992). *Trail Making Test: Manual for administration and scoring*. Tucson,
AZ: Reitan Neuropsychology Laboratory.

Rey, A. (1941). L'examen psychologique dans les cas d'encephalopathie traumatique.
Archives de Psychologie, 28, 286-340.

Rey, A. (1964). *L'examen clinique en psychologie*. Paris, France: Presses Universitaires
de France.

Riccio, C. A., Yoon, H., & McCormick, A. S. (2014). Neuropsychological test selection
with clients who are Asian. In J. M. Davis & R. C. D'Amato (Eds.),
*Neuropsychology of Asians and Asian-Americans: Practical and theoretical
considerations*. 151-174. New York, NY: Springer.

- Rock, P. L., Roiser, J. P., Riedel, W. J., & Blackwell, A. D. (2014). Cognitive impairment in depression: A systematic review and meta-analysis. *Psychological Medicine*, 44(10), 2029-2040. doi: 10.1017/S0033291713002535
- Rohling, M. L., Larrabee, G. J., & Millis, S. R. (2012). The “miserable minority” following mild traumatic brain injury: Who are they and do meta-analyses hide them? *The Clinical Neuropsychologist*, 26(2), 197-213. doi: 10.1080/13854046.2011.647085
- Rohlf, H. G., Knipscheer, J. W., & Kleber, R. J. (2014). Somatization in refugees: A review. *Social Psychiatry and Psychiatric Epidemiology*, 49(11), 1793-1804. doi: 10.1007/s00127-014-0877-1
- Root, J. C., Robbins, R. N., Chang, L., & VanGorp, W. G. (2006). Detection of inadequate effort on the California Verbal Learning Test – Second edition: Forced choice recognition and critical item analysis. *Journal of the International Neuropsychological Society*, 12(5), 688-696. doi: 10.1017/s1355617706060838
- Rubenzler, S. (2009). Posttraumatic stress disorder: Assessing response style and malingering. *Psychological Injury and Law*, 2(2), 114-142. doi: 10.1007/s12207-009-9045-4
- Ruff, R. M., Camenzuli, L., & Mueller, J. (1996). Miserable minority: Emotional risk factors that influence the outcome of a mild traumatic brain injury. *Brain Injury*, 10(8), 551-565. doi: 10.1080/026990596124124
- Ruff, R. M., Iverson, G. L., Barth, J. T., Bush, S. S., & Broshek, D. K. (2009). Recommendations for diagnosing a mild traumatic brain injury: a National

- Academy of Neuropsychology education paper. *Archives of Clinical Neuropsychology*, 24(1), 3-10. doi: 10.1093/arclin/acp006
- Ruff, R. M., Light, R. H., Parker, S. B., & Levin, H. S. (1996). Benton controlled oral word association test: Reliability and updated norms. *Archives of Clinical Neuropsychology*, 11(4), 329-338. doi: 10.1016/0887-6177(95)00033-x
- Salazar, X. F., Lu, P. H., Wen, J., & Boone, K. B. (2007). The use of effort tests in ethnic minorities and in non-English-speaking and English as a second language populations. In K. B. Boone (Eds.), *Assessment of Feigned Cognitive Impairment: A Neuropsychological Perspective*. 405-427. New York, NY: Guildford Press.
- Sandole, D. H., & Auerbach, C. F. (2013). Dissociation and identity transformation in female survivors of the genocide against the Tutsi in Rwanda: A qualitative research study. *Journal of Trauma & Dissociation*, 14(2), 127-137. doi: 10.1080/15299732.2013.724345
- Sar, V., Taycan, O., Bolat, N., Ozmen, M., Duran, A., Oztürk, E., & Ertem-Vehid, H. (2010). Childhood trauma and dissociation in schizophrenia. *Psychopathology* 43(1), 33-40. doi: 10.1159/000255961
- Schroeder, R. W., & Marshall, P. S. (2011). Evaluation of the appropriateness of multiple symptom validity indices in psychotic and non-psychotic psychiatric populations. *The Clinical Neuropsychologist*, 25(3), 437-453.
<http://dx.doi.org.ezproxy.uwindsor.ca/10.1080/13854046.2011.556668>
- Schroeder, R. W., Twumasi-Ankrah, P., Baade, L. E. & Marshall, P. S. (2012). Reliable Digit Span: A systematic review and cross-validation study. *Assessment*, 19(1), 21-30. doi: 10.1177/1073191111428764

- Schutte, C., & Axelrod, B. N. (2012). Use of embedded cognitive symptom validity measures in mild traumatic brain injury cases. In: Carone, D., & Bush, S. S. *Mild traumatic brain injury: Symptom validity assessment and malingering*. 159-181. New York, NY: Springer Publishing Company.
- Schwarz, L. R., Gfeller, J. D., & Oliveri, M. V. (2006). Detecting feigned impairment with the Digit Span and Vocabulary subtests of the Wechsler Adult Intelligence Scale—Third Edition. *The Clinical Neuropsychologist*, 20(4), 741-753. doi: 10.1080/13854040590967054
- Scott, J. C., Matt, G. E., Wrocklage, K. M., Crnich, C., Jordan, J., Southwick, S. M., Krystal, J. H., & Schweinsburg, B. C. (2015). A quantitative meta-analysis of neurocognitive functioning in posttraumatic stress disorder. *Psychological Bulletin*, 14(1), 105-140. doi: 10.1037/a0038039
- Schagen, S., Schmand, B., de Sterke, S., & Lindeboom, J. (1997). Amsterdam Short-Term Memory test: A new procedure for the detection of feigned memory deficits. *Journal of Clinical and Experimental Neuropsychology*, 19(1), 43-51. doi: 10.1080/01688639708403835
- Shafaei, A., Abd Razak, N., & Nejati, M. (2016). Integrating two cultures successfully: Factors influencing acculturation attitude of international postgraduate students in Malaysia. *Journal of Research in International Education*, 15(2), 137-154. doi: 10.1177/1475240916653566
- Sherer, M., Davis, L. C., Sander, A. M., Nick, T. G., Luo, C., Pastorek, N., & Hanks, R. (2015). Factors associated with Word Memory Test performance in persons with

- medically documented traumatic brain injury. *The Clinical Neuropsychologist*, 29(4), 522-541. doi: 10.1080/13854046.2015.1052763
- Sherman, D. S., Boone, K. B., Lu, P., & Razani, J. (2002). Re-examination of a Rey Auditory Verbal Learning Test/Rey Complex Figure discriminant function to detect suspect effort. *The Clinical Neuropsychologist*, 16(3), 242-250. doi: 10.1076/clin.16.3.242.13860
- Sierra, M., & Berrios, G. E. (2000). The Cambridge Depersonalisation Scale: A new instrument for the measurement of depersonalisation. *Psychiatry Research*, 93(2), 153-164. doi: 10.1016/s0165-1781(00)00100-1
- Silver, J. M. (2015). Invalid symptom reporting and performance: What are we missing? *NeuroRehabilitation*, 36(4), 463-469. doi: 10.3233/nre-151234
- Silverberg, N. D., Iverson, G. L., Brubacher, R., Holland, E., Hoshino, L. C., Aquino, A., & Lange, R. T. (2016). The nature and clinical significance of preinjury recall bias following mild traumatic brain injury. *Journal of Head Trauma Rehabilitation*, 21(6), 388-396. doi: 10.1097/HTR.0000000000000198
- Skidmore, S. T., & Thompson, B. (2013). Bias and precision of some classical ANOVA effect sizes when assumptions are violated. *Behaviour Research Methods*, 45(2), 536-546. doi: 10.3758/s13428-012-0257-2
- Slick, D., Hopp, G., Strauss, E., & Thompson, G. (1997). *The Victoria Symptom Validity Test*. Odessa, FL: Psychological Assessment Resources, Inc.
- Slick, D. J., Sherman, E. M. S., & Iverson, G. L. (1999). Diagnostic criteria for malingered neurocognitive dysfunction: Proposed standards for clinical practice

and research. *The Clinical Neuropsychologist*, 13(4), 545-561. doi: 10.1076/1385-4046(199911)13:04;1-y;ft545

Smith, C. (2011). Neuropathology. In J. M. Silver, T. W. McAllister, & S. C. Yudofsky (Eds.), *Textbook of Traumatic Brain Injury 2nd Edition*. 23-36. Washington, DC: American Psychiatric Publishing, Inc.

Smith, G. P., & Burger, G. K. (1997). Detection of malingering: Validation of the Structured Inventory of Malingered Symptomatology (SIMS). *Journal of the American Academy of Psychiatry and the Law*, 25(2), 183-189. doi: 10.1037/t04573-000

Smith, K., Boone, K. B., Victor, T., Miora, D., Cottingham, M., Ziegler, E., Zeller, M., & Wright, M. (2014). Comparison of credible patients of very low intelligence and noncredible patients on neurocognitive performance validity indicators. *The Clinical Neuropsychologist*, 28(6), 1048-1070. doi: 10.1080/13854046.2014.931465

Soffer-Dudek, N. (2014). Dissociation and dissociative mechanisms in panic disorder, obsessive-compulsive disorder, and depression: A review and heuristic framework. *Psychology of Consciousness: Theory, Research, and Practice*. 1(3), 243-270. doi: 10.1037/cns0000023

Sollman, M. J., & Berry, D. T. R. (2011). Detection of inadequate effort on neuropsychological testing: A meta-analytic update and extension. *Archives of Clinical Neuropsychology*, 26(8), 774-789. doi: 10.1093/arclin/acr066

- Spearing, N. M., Connelly, L. B., Gargett, S., & Sterling, M. (2012). Does injury compensation lead to worse health after whiplash? A systematic review. *Pain, 153*(6), 1274-1282. doi: 10.1016/j.pain.2012.03.007
- Spenceley, L. M. (2013). *The impact of incentives on neuropsychological test performance: An analog study*. (Doctoral dissertation). Retrieved from ProQuest Dissertations Publishing. (3510000)
- Spencer, S. J., Logel, C., & Davies, P. G. (2016). Stereotype threat. *Annual Review of Psychology, 67*, 415-437. doi: 10.1146/annurev-psych-073115-103235
- Statistics Canada (2011). *Linguistic Characteristics of Canadians* (Catalogue no. 98-314-X2011001). Retrieved from <http://www12.statcan.gc.ca/census-recensement/2011/as-sa/98-314-x/98-314-x2011001-eng.pdf>
- Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist, 52*(6), 613-629. doi:10.1037/0003-066X.52.6.613
- Stein, D. J., Koenen, K. C., Friedman, M. J., Hill, E., McLaughlin, K. A., Petukhova, M., ... Kessler, R. C. (2013). Dissociation in Posttraumatic Stress Disorder: Evidence from the World Mental Health Surveys. *Biological Psychiatry, 73*(4), 302-312. doi: 10.1016/j.biopsych.2012.08.022
- Steinberg, M., & Schnall, M. (2003). *The stranger in the mirror: Dissociation – The hidden epidemic*. New York, NY: HarperCollins.
- Stevens, J. P. (2009). *Applied Multivariate Statistics for the Social Sciences, 5th Edition*. New York, NY: Routledge.

- Strauss, E., Sherman, E. M., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary*. Oxford University Press, USA.
- Studerus-Germann, A. M., Thiran, J. P., Daducci, A., & Gautschi, O. P. (2016). Diagnostic approaches to predict persistent posttraumatic symptoms after mild traumatic brain injury – a literature review. *International Journal of Neuroscience*, 126(4), 289-298. doi: 10.3109/00207454.2015.1033620
- Stulemeijer, M., Andriessen, T. M. J. C., Brauer, J. M. P., Vos, P. E., & van der Werf, S. P. (2007). Cognitive performance after mild traumatic brain injury: The impact of poor effort on test results and its relation to distress, personality and litigation. *Brain Injury*, 21(3), 309-318. doi: 10.1080/02699050701209980
- Sugarman, M. A., & Axelrod, B. N. (2015). Embedded measures of performance validity using verbal fluency tests in a clinical sample. *The Clinical Neuropsychologist*, 22(2), 141-146. doi: 10.1080/23279095.2013.873439
- Sugarman, M. A., Holcomb, E. M., Axelrod, B. N., Meyers, J. E., & Liethen, P. C. (2015). Embedded measures of performance validity in the Rey Complex Figure Test in a clinical sample of veterans. *Applied Neuropsychology: Adult*, 23(2), 105-114. doi: 10.1080/23279095.2015.1014557
- Suhr, J. A., & Gunstad, J. (2000). The effects of coaching on the sensitivity and specificity of malingering measures. *Archives of Clinical Neuropsychology*, 15(5), 415-424. doi: 10.1016/s0887-6177(99)00033-5
- Suhr, J. A., & Gunstad, J. (2002). “Diagnosis threat”: The effect of negative expectations on cognitive performance in head injury. *Journal of Clinical and Experimental Neuropsychology*, 24(4), 448-457. doi: 10.1076/jcen.24.4.448.1039

- Teasdale, G., Maas, A., Lecky, F., Manley, G., Stocchetti, N., & Murray, G. (2014). The Glasgow Coma Scale at 40 years: Standing the test of time. *The Lancet Neurology*, 13(8), 844-854. doi: 10.1016/s1474-4422(14)70120-6
- Teng, E. L., Hasegawa, K., Homma, A., Imai, Y., Larson, E., Graves, A., Sugimoto, K., Yamaguchi, T., Sasaki, H., Chiu, D., Chiu, D., & White, L. R. (1994). The Cognitive Abilities Screening Instrument (CASI): A practical test for cross-cultural epidemiological studies of dementia. *International Psychogeriatrics*, 6(1), 45-58. doi: 10.1017/s1041610294001602
- Thames, A. D., Hinkin, C. H., Byrd, D. A., Bilder, R. M., Duff, K. ., Mindt, M. R., Arentoft, A., & Streiff, V. (2013). Effects of stereotype threat, perceived discrimination, and examiner race on neuropsychological performance: Simple as black and white? *Journal of the International Neuropsychological Society*, 19(5), 583-593. doi: 10.1017/S1355617713000076
- Thompson, J., O'Donnell, M., Stafford, L., Nordfjaern, T., & Berk, M. (2014). Association between attributions of responsibility for motor vehicle crashes, depressive symptoms, and return to work. *Rehabilitation Psychology*, 59(4) 376-385. doi: 10.1037/rep0000012
- Tombaugh, T. N. (1996). *Test of Memory Malingering (TOMM)*. Toronto, ON: Multi-Health Systems.
- Turcotte, M., & Schellenberg, G. (2007). *A portrait of seniors in Canada, 2006*. Ottawa, ON: Statistics Canada, Social and Aboriginal Statistics Division.
- United States Census Bureau (2010). State & County QuickFacts. Retrieved from <http://quickfacts.census.gov/qfd/states/00000.html>

- Uskul, A. K., & Greenglass, E. (2005). Psychological wellbeing in a Turkish-Canadian sample. *Anxiety, Stress, & Coping, 18*(3), 269-287. doi: 10.1080/10615800500205983
- Vang, Z., Sigouin, J., Flenon, A., & Gagnon, A. (2015). The healthy immigrant effect in Canada: A systematic review. *Population Change and Lifecourse Strategic Knowledge Cluster Discussion Paper Series, 3*(1), Article 4. Retrieved from <http://ir.lib.uwo.ca/pclc/vol3/iss1/4/>
- Vermeer, S.E., Longstreth, W.T., & Koudstaal, P.J. (2007). Silent brain infarcts: A systematic review. *Lancet Neurology, 6*(7), 611-619. doi: 10.1016/s1474-4422(07)70170-9
- Vickery, C.D., Berry, D.T.R., Inman, T.H., Harris, M.J., & Orey, S.A. (2001). Detection of inadequate effort on neuropsychological testing: A meta-analytic review of selected procedures. *Archives of Clinical Neuropsychology, 16*(1), 45-73. doi: 10.1093/arclin/16.1.45
- Victor, T. L., Boone, K. B., Serpa, J. G., Buehler, J., & Ziegler, E. A. (2009). Interpreting the meaning of multiple symptom validity test failure. *The Clinical Neuropsychologist, 23*(2), 297-313. doi: 10.1080/13854040802232682
- Vilar-Lopez, R., Gomez-Rio, M., Santiago-Romajo, S., Rodriguez-Fernandez, A., Puente, A. E., & Perez-Garcia, M. (2008). Malingering detection in a Spanish population with a known-groups design. *Archives of Clinical Neuropsychology, 23*(4), 365-377. doi: 10.1016/j.acn.2008.01.007
- Vilar-Lopez, R., Gomez-Rio, M., Caracuel-Romero, A., Llamas-Elvira, J., & Perez-Garcia, M. (2008). Use of specific malingering measures in a Spanish sample.

Journal of Clinical and Experimental Neuropsychology, 30(6), 710-722. doi:
10.1080/13803390701684562

Vilar-Lopez, R., Santiago-Ramajo, S., Gomez-Rio, M., Verdejo-Garcia, A., Llamas, J. M., & Perez-Garcia, M. (2007). Detection of malingering in a Spanish population using three specific malingering tests. *Archives of Clinical Neuropsychology*, 22(3), 379-388. doi: 10.1016/j.acn.2007.01.012

Votruba, K. L., Persad, C., & Giordani, B. (2016). Cognitive deficits in healthy elderly population with “normal” scores on the Mini-Mental State Examination. *Journal of Geriatric Psychiatry and Neurology*, 29(3), 126-132. doi:
10.1177/0891988716629858

Waljas, M., Iverson, G. L., Lange, R. T., Hakulinen, U., Dastidar, P., Huhtala, H., Limatainen, S., Hartikainen, K., & Ohman, J. (2015). A prospective biopsychosocial study of the persistent post-concussive symptoms following mild traumatic brain injury. *Journal of Neurotrauma*, 32(5), 534-547. doi:
10.1089/neu.2014.3339

Warrington, E. K. (1984). *Recognition Memory Test Manual*. Windsor, England: NFER-Nelson.

Webb, J. W., Batchelor, J., Meares, S., Taylor, A., & Marsh, N. V. (2012). Effort test failure: Toward a predictive model. *The Clinical Neuropsychologist*, 26(8), 1377-1396. doi: 10.1080/13854046.2012.728248

Wechsler, D. (2011). *Wechsler Abbreviated Scale of Intelligence – Second Edition*. San Antonio, TX: Pearson.

- Wechsler, D. (2008). *WAIS-IV: Wechsler Adult Intelligence Scale*. San Antonio, TX: Pearson.
- Wechsler, D. (1981). *Wechsler Adult Intelligence Scale – Revised*. New York, NY: The Psychological Corporation.
- Wechsler, D. (1997a). *Wechsler Adult Intelligence Scale, Third Edition*. New York, NY: The Psychological Corporation.
- Wechsler, D. (1997b). *Wechsler Memory Scale – Third Edition. Administration and Scoring Manual*. San Antonio, TX: The Psychological Corporation.
- Weinborn, M., Woods, S., Nulsen, C., & Leighton, A. (2012). The effects of coaching on the Verbal and Nonverbal Medical Symptom Validity Tests. *The Clinical Neuropsychologist*, 26(5), 832–849. doi: 10.1080/13854046.2012.686630
- Weiss, R., & Rosenfeld, B. (2010). Cross-cultural validity in malingering assessment: The Dot Counting Test in a rural Indian sample. *Journal of Forensic Mental Health*, 9(4), 300-307. doi: 10.1080/14999013.2010.526680
- West, L. K., Curtis, K. L., Greve, K. W., & Bianchini, K. J. (2011). Memory in traumatic brain injury: The effects of injury severity and effort on the Wechsler Memory Scale-III. *Journal of Neuropsychology*, 5(1), 114-125. doi: 10.1348/174866410x521434
- Whiteside, D. M., Kogan, J., Wardin, L., Phillips, D., Franzwa, M. G., Rice, L., Basso, M., & Roper, B. (2015). Language-based embedded performance validity measures in traumatic brain injury. *Journal of Clinical and Experimental Neuropsychology*, 37(2), 220-227. doi: 10.1080/13803395.2014.1002758

- Whiteside, D., Wald, D., & Busse, M. (2011). Classification accuracy of multiple visual spatial measures in the detection of suspect effort. *The Clinical Neuropsychologist*, 25(2), 287-301. doi: 10.1080/13854046.2010.538436
- Wilkinson, G. S., & Robertson, G. J. (2006). *Wide Range Achievement Test (WRAT4)*. Odessa, FL: Psychological Assessment Resources.
- Williams, D. H., Levin, H. S., & Eisenberg, H. M. (1990). Mild head injury classification. *Neurosurgery*, 27(3), 422-428. doi: 10.1097/00006123-199009000-00014
- Wisdom, N. M., Pastorek, N. J., Miller, B. I., Booth, J. E., Romesser, J. M., Linck, J. F., & Sim, A. H. (2014). PTSD and cognitive functioning: Importance of including performance validity testing. *The Clinical Neuropsychologist*, 28(1), 128-145. doi: 10.1080/13854046.2013.863977
- Wolfe, P. L., Millis, S. R., Hanks, R. A., Fichtenberg, N., Larrabee, G. J., & Sweet, J. J. (2010). Effort indicators within the California Verbal Learning Test – II (CVLT-II). *The Clinical Neuropsychologist*, 24(1), 153-158. doi: 10.1080/13854040903107791
- Wood, R. L. (2004). Understanding the “miserable minority”: A diathesis-stress paradigm for post-concussional syndrome. *Brain Injury*, 18(11), 1135-1153. doi: 10.1080/02699050410001675906
- World Health Organization (2004). *World report on road traffic injury prevention*. Geneva, Switzerland: WHO Library.

- Yamaguchi, T. (2005). *Detecting malingered memory impairment using the Rey 15-Item Memory Test and the Wechsler Digit Span subtest in a Japanese population*. UMI Dissertation Services, ProQuest Information and Learning, Ann Arbor, MI.
- Yanez, Y., Tami, F., Tennant, J., Strunk, J., & Coker, K. (2006). Effects of severe depression on TOMM performance among disability-seeking outpatients. *Archives of Clinical Neuropsychology*, 21(2), 161-165. doi: 10.1016/j.acn.2005.07.009
- Yang, C. C., Kao, C. J., Cheng, T. W., Yang, C. C., Wang, W. H., Yu, R. L., Hsu, Y. H., & Hua, M. S. (2012). Cross-cultural effect of suboptimal effort detection: An example of the Digit Span subtest of the WAIS-III in Taiwan. *Archives of Clinical Neuropsychology*, 27(11), 869-878. doi: 10.1093/arclin/acs081
- Ylioja, S. G., Baird, A. D., & Podell, K. (2009). Developing as spatial analogue of the Reliable Digit Span. *Archives of Clinical Neuropsychology*, 24(8), 729-739. doi: 10.1093/arclin/acp078
- Young, G. (2008). Somatization and medically unexplained symptoms in psychological injury: Diagnoses and dynamics. *Psychological Injury and Law*, 1(4), 224-242. doi: 10.1007/s12207-008-9021-4
- Young, G. (2015a). Detection system for malingered PTSD and related response biases. *Psychological Injury and Law*, 8(2), 169-183. doi: 10.1007/s12207-015-9226-2
- Young, G. (2015b). Malingering in forensic disability-related assessments: Prevalence 15 +/- 15%. *Psychological Injury and Law*, 8(3), 188-199. doi: 10.1007/s12207-015-9232-4

- Young, G. (2015c). Towards balanced VA and SSA policies in psychological injury disability assessment. *Psychological Injury and Law*, 8(3), 200-218. doi: 10.1007/s12207-015-9230-6
- Youngjohn, J. R., Lees-Haley, P. R., & Binder, L. M. (1999). Comment: Warning malingerers produces more sophisticated malingering. *Archives of Clinical Neuropsychology*, 14(6), 511-515. doi: 10.1093/arclin/14.6.511
- Yun, K., Paul, P., Subedi, P., Kuikel, L., Nguyen, G. T., & Barg, F. K. (2016). Help-seeking behavior and health care navigation by Bhutanese refugees. *Health Promotion and Disease Prevention*, 41(3), 526-534.
- Zagorchev, L., Meyer, C., Stehle, T., Wenzel, F., Young, S., Peters, J. ... McAllister, T. (2016). Differences in regional brain volumes two months and one year after mild traumatic brain injury. *Journal of Neurotrauma*, 33(1), 29-34. doi: 10.1089/neu.2014.3831
- Zhang, S., Carroll, L. J., Cassidy, J. D., & Paniak, C. (2009). Factors influencing self-rated health in traffic-related mild traumatic brain injury. *Journal of Rehabilitation Medicine*, 41(13), 1062--1067. doi: 10.2340/16501977-0462

Vita Auctoris

NAME: Shayna Hannah Nussbaum

PLACE OF BIRTH: Brantford, ON

YEAR OF BIRTH: 1988

EDUCATION: Tiferes Bais Yaakov High School, Toronto, ON,
2007

York University, Hon. B.Sc., Toronto, ON,
2010

University of Windsor, M.A., Windsor, ON,
2014