Predictors of effort test failure during neuropsychological evaluation in cases of traumatic brain injury

James William Webb
MSocSc (Hons), PGDipPsych(Clin)

Macquarie University
Department of Psychology

This thesis is presented as a partial fulfilment to the requirements for the degree of Doctor of Philosophy in Psychology

November, 2013

Table of Contents

List of Figures and Tables

CHAPTER 1: THESIS OVERVIEW AND SETTING

CHAPTER 5: EFFORT TEST FAILURE: TOWARDS A PREDICTIVE MODEL

CHAPTER 6: PSYCHOLOGICAL PREDICTORS OF EFFORT TEST FAILURE

CHAPTER 7: ACCULTURATION AS A PREDICTOR OF EFFORT TEST FAILURE

CHAPTER 8: SELF-REPORTED DEPRESSIVE SYMPTOMATOLOGY AND EFFORT TEST FAILURE: A SYSTEMATIC REVIEW

APPENDIX A

Abstract

The application of suboptimal effort during neuropsychological assessment is frequently encountered. The variables that predict suboptimal effort are not well understood. To date individual variables that are associated with suboptimal effort have been examined in an ad hoc fashion and there has been no attempt to construct an empirical model of cognitive test taker effort. The aim of the current series of four studies was to examine the relationship between economic, demographic, psychological, and personal history variables and effort test failure (ETF), using a multivariable statistical technique (logistic regression) that has been only rarely employed in effort research. In study 1 the power of economic, demographic, psychological and behavioural variables to predict ETF was examined in an archival consecutive sample of mixed-severity adult traumatic brain injury patients (*N* = 555). In study 2 the predictive power of psychological and personal history variables to predict ETF was further examined while holding constant the statistical predictors identified in study 1. Study 3 comprised an exploration of the predictive relationship between a range of acculturation variables and ETF while holding constant the predictive variables identified in study 1. In study 4 the relationship between self-reported depressive symptomatology (SRDS) and ETF in compensation-seeking samples was examined by undertaking a systematic review of the literature between 1950 and 2012 (inclusive). A total of 9,501 articles were screened, of which 19 satisfied inclusion criteria.

The results of study 1 revealed ETF to be significantly associated with compensation-seeking, low education, self-reported mood disorder, exaggerated displays of behavior, psychotic illness, being foreign-born, having sustained a workplace accident, and mild as compared to severe traumatic brain injury. In study 2 it was demonstrated that, holding study 1 variables constant, of a range of psychological and personal history variables examined, only self-reported depressive disorder was predictive of ETF. Scores on the Beck Depression Inventory-II were predictive of ETF holding compensation-seeking constant. Of the range of

acculturation variables examined in study 3, only age at which English was learned was found to make a significant independent contribution to the predictive model established in study 1. The systematic review of the literature that examined SRDS and ETF revealed that studies were of high quality but typically afforded a low level of evidence. The results of those studies revealed a medium to large effect of SRDS on test taker effort. Psychological symptom reporting was found to be elevated in this population but frank malingering was not detected. Together, the studies indicate that ETF can be predicted by psychological symptom reporting, the display of abnormal behaviours, economic variables, demographic variables, injury-related variables, and workplace variables.

Statement of Candidate

I certify that the work in this thesis entitled "Predictors of effort test failure during neuropsychological evaluation in cases of traumatic brain injury" has not previously been submitted for a degree nor has it been submitted as part of the requirements for a degree to any other university or institution other than Macquarie University.

I also certify that the thesis is an original piece of research and it has been written by me. Any help and assistance that I have received in my research work and the preparation of the thesis itself has been appropriately acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

The research was approved by Macquarie University Ethics Review Committee, reference number, HE26JUN2009-D00005 on 8 July, 2009.

James William Webb (Student ID: 41879260)

1 November, 2013

Acknowledgements

I would like to thank my supervisors, Drs Jennifer Batchelor and Susanne Meares, for their support through this extended programme of study. They have each supported me with clarity of thought and patience, they have disciplined my mind to scientific methods, and together they have provided me with a pair of steady hands on the tiller. Studying by distance has its challenges for student and supervisor alike and I thank them for staying cool and settling my nerves via telephone and email through the tough times.

I thank Professor Nigel Marsh for providing the initial impetus to convert the clinical material contained in this thesis into an academic body of work and for his initial help laying out a framework for the study. I thank Professor Don Hines for his skill in teaching the complexities of multivariate data analysis in a way that penetrated and stuck. Particular thanks go to Associate Professor Arthur Shores for the opportunity to gain from his considerable knowledge in this field of effort research, for his breadth of clinical experience, and for providing me with the opportunity to study at Macquarie University.

I would also like to thank Dr Alan Taylor whose knowledge of statistics and methods of data analysis was invaluable during the statistical analyses undertaken in the studies included in this thesis.

Special thanks go to the Macquarie University Library for their exceptional online facilities and database access, without which this project could not have succeeded. Thanks to the HDR office and Erin Thatcher specifically for flexibility and understanding the peculiar needs of an external student.

Finally, greatest thanks to my wife, Katia, and children, Amelie and Jules, for their steadfast support. Katia, Amelie and Jules were willing to forego vast quantities of my time to support this project, and despite the cost of time, income lost, and my endless fretting, Katia never once said, "Will you please just get back to work!"

CHAPTER 1: THESIS OVERVIEW AND SETTING

Structure of the thesis

This thesis examines neuropsychological test-taker effort in cases of traumatic brain injury. The broadest aim of the thesis is to identify variables that might be predictive of a patient affording a suboptimal effort during cognitive testing. To that end the archival records of a sample of 555 patients will be examined, and the data from those records will be analysed using modern and powerful statistical methods.

This thesis is presented as a thesis by publication, and includes three empirical studies examining predictors of low test taker effort, and one systematic review of the literature describing depressive symptomatology and cognitive test-taker effort. Each publication will be accompanied by its own relevant reference list.

Table 1 describes the sample of participants employed for each of the three empirical studies, the inclusion and exclusion criteria employed and any subgroups of participants included. The entire sample of 555 participants was included for studies 1 and 3, while a subgroup of those with psychotic illness was excluded from the pool for study 2 on the basis that psychosis was specifically examined as a risk factor for effort test failure in study 1. Study 2 included two subgroups of the sample pool – those participants that had completed self-report affect measures (Beck Depression Inventory-2 and State Trait Anxiety Inventory).

Table 1

Participant characteristics for each study

Study	Inclusion criteria	Exclusion criteria	Total *N*	Subgroup characteristics
1 (Chapter 5)	Consecutive adult TBI referrals 2001-2007 inclusive.	Pre-existing history of mental retardation or dementia.	555	None
2 (Chapter 6)	Consecutive adult TBI referrals 2001-2007 inclusive.	Pre-existing history of mental retardation, pre-existing or post-injury psychotic illness, or dementia.	540	*n* = 207 (participants completed BDI-2). *n* = 90 (participants completed the STAI).
3 (Chapter 7)	Consecutive adult TBI referrals 2001-2007 inclusive.	Pre-existing history of mental retardation or dementia.	555	None

Note. TBI = traumatic brain injury. BDI-2 = Beck Depression Inventory – second edition. STAI = State Trait Anxiety Inventory.

One study (Chapter 5) has been published in the international peer-reviewed journal *The Clinical Neuropsychologist* (Webb et al., 2012), Chapter 8 has been submitted for publication to the journal *Brain Injury* and is currently under review, and the remainder of the studies have been prepared for submission for publication to international peer-reviewed journals, in due course.

The setting

The sample of participants represents consecutive referrals of patients to a private neuropsychology practice for cognitive assessment following traumatic brain injury. All patients have been individually assessed by the author who is a licensed clinical psychologist. The author's practice is based in Auckland, New Zealand, and the practice receives referrals from a variety of sources, however for the purposes of the studies comprising this thesis, all participants were referred by one agency – the Accident Compensation Corporation.

New Zealand has, since 1974, adopted a pure no-fault no litigation state-administered worker's compensation scheme. After a number of minor revisions the system is, in essence, unchanged since 1974 and the most recent Accident Compensation Act 2001 establishes that the Accident Compensation Corporation manages the rehabilitation and compensation of all individuals who sustain accidental injuries. Litigation for compensation for injury is completely precluded. In New Zealand worker's compensation is paid to injured people by the ACC irrespective of the cause of injury (but it is not paid for non-injury medical illness) and irrespective of fault. Worker's compensation is paid at a rate of 80% of gross yearly income and is capped at a total income of $NZ 88,000 per annum.

The majority of the sample of participants employed in the studies of this thesis (85%) were compensation-seeking; that is, seeking to gain access to or maintain access to worker's compensation payments. A substantial minority of participants (15%) were inelligible for worker's compensation because they were not in paid employment at the time of their accidental injury (e.g., university students, full-time home-makers or parents).

New Zealand is a culturally diverse Pacific Island nation of approximately four million inhabitants (Statistics New Zealand, 2013). Indigenous New Zealand Maori comprise 15% of the population while other people of Pacific Island ethnicity comprise 7%. People of Asian ethnicity comprise about 12% of the population while Caucasian people of European/White ethnicity comprise about 74% of the population (Statistics New Zealand, 2013; census allows people to identify more than one ethnic group and as such percentages do not add to 100).

The sample of participants employed in this thesis is similarly ethnically diverse. European/White participants comprise 75% of the sample, Maori/Pacific Island comprise 29% of the sample, while those of Asian descent comprise 6%. When compared with census data it is apparent that the sample included here contains a slightly higher proportion of people of Maori/Pacific Island ethnicity relative to those of Indo/Asian ethnicity. Although the ethnic proportions contained in this participant sample do not exactly reflect the ethnic proportions seen in New Zealand they are consistent with the proportions found in New Zealand injury epidemiology studies and reflect the higher rates of accidental injury seen in Maori/Pacific Island people relative to those of Asian descent (Feigin et al., 2013; Hosking, Ameratunga, Exeter & Stewart, 2013).

References

Accident Compensation Act 2001 No 49, Public Act.

Feigin, V. L., Theadom, A., Barker-Collo, S., Starkey, N. J., McPherson, K., Kahan, M., Dowell, A., Brown, P., Parag, V., Kydd, R., Jones, K., Jones, A., & Ameratunga, S. (2013). Incidence of traumatic brain injury in New Zealand: a population-based study. *The Lancet Neurology*, *12,* 53-64.

Hosking, J., Ameratunga, S., Exeter, D., Stewart, J., & Bell, A. (2013). Ethnic, socioeconomic and geographical inequalities in road traffic injury rates in the Auckland region. *Australian and New Zealand Journal of Public Health*, *37*, 162-167.

Statistics New Zealand. (2013). *2013 census QuickStats about national highlights.* Retrieved from http://owll.massey.ac.nz/referencing/referencing-online-material-in-apa.php

CHAPTER 2: EFFORT TESTING AND SYMPTOM EXAGGERATION

Clinical neuropsychology is a science that has as its primary focus, the behavioural expression of brain dysfunction (Lezak, Howieson & Loring, 2004). Eliciting a patient's optimal cognitive test performance is necessary for accurate behavioural assessment (Lezak, Howieson & Loring, 2004); however, maximal test performance may not be forthcoming for a variety of reasons. Faust, Ahern, Bridges & Yonce (2012) have outlined reasons for suboptimal test performance including methodological weaknesses of the neuropsychological testing process (e.g., measurement error, non-standard administration of the measure), and extraneous factors (e.g., medication side effects). Other reasons for suboptimal effort are associated with characteristics of the patient and those include the effort that the examinee applies to the neuropsychological task.

The four studies that comprise the current thesis examine the phenomenon of cognitive test taker effort with the aim of identifying individual characteristics that are predictive of low effort during neuropsychological evaluation. The introductory chapter comprises a preliminary review of the effort literature, including consideration of the terminology that is associated with the measurement of effort during neuropsychological testing, the professional position held in neuropsychology on effort testing, base rates of effort test failure (ETF), the impact of effort on neuropsychological test outcomes and methods of effort testing. The phenomenon of symptom exaggeration is reviewed and differentiated from ETF and the abnormal nonverbal behaviours that can be associated with symptom exaggeration are considered. Finally, the distinction between ETF and malingering is discussed.

Test Taker Effort

Test taker effort has seldom been formally defined in the neuropsychological literature, with most uses of the term relying on common understandings. Slick and Sherman (2013) describe effort in terms of the amount of exertion or the amount of mental and/or physical energy expended toward a task. Effort has also been described as “investment in performing at capacity levels,” and an attempt to perform well (Bush et al., 2005, p.420). Iverson (2010) defines effort in terms of outcome, stating that a person has afforded suboptimal effort if they have “underperformed during testing” (p. 99).

The neuropsychological literature employs various terms to describe effort-related constructs. Terms such as suboptimal effort, insufficient effort, inadequate effort and poor effort are employed and there is no consensus on the preferred descriptor at this time (Heilbronner, Sweet, Morgan, Larrabee, & Millis, 2009). For the purposes of the current thesis the term *suboptimal effort* will be employed to describe underperformance during cognitive testing generally, and the term *effort test failure* (ETF) will be used to refer to instances where insufficient effort was exerted in the direction of competent performance on a specific test. Psychometric measures that are used to identify suboptimal effort are known as effort tests and are in a category of measures that evaluate the validity of symptoms, collectively known as symptom validity tests (SVTs; Pankratz, 1979).

Effort tests typically require a low level of effort and cognitive ability to complete them adequately (Heilbronner et al., 2009) and they are substantially (but not wholly) robust to the influence of neurological injury or disease; effort tests masquerade as difficult measures of cognition (usually memory). ETF on any specific effort measure is determined by failure to achieve a test score above or below a cut-score that has been shown by previous research to discriminate between those affording adequate effort and those affording suboptimal effort. Typically, cut-scores are set to an acceptable specificity of at least 90% (Babikian & Boone, 2007), meaning that false positive reporting (reporting poor effort when effort was normal) may occur on average in one of ten cases. A measure with imperfect sensitivity may fail to detect suboptimal effort when, in fact, the examinee’s effort was not a determined attempt (false negative). Most effort measures have sensitivity in the range of 40-

70% (Boone, 2011), indicating that about one-third to one-half of cases of insufficient effort may not be detected by any one individual measure.

Some measures of effort are independent or standalone tests, usually relying on a forced-choice paradigm (e.g., Test of Memory Malingering, TOMM; Tombaugh, 1996). Others are built into neuropsychological tests that are already in use and employ empirically-based markers for the detection of low effort. Such measures are often described as *embedded measures* (e.g., the Logical Memory recognition trial from the Advanced Clinical Solutions of the Wechsler Memory Scale (WMS) – Fourth Edition; Pearson Education, 2008). There is a growing call to employ both standalone and embedded measures of effort throughout a testing session (AACN, 2007; Meyers & Volbrecht, 2003).

Base rates of ETF

ETF has been shown to be a relatively common phenomenon. Mittenberg, Patton, Canyock, and Condit (2002) undertook a survey of American Board of Clinical Neuropsychology diplomats and found that estimates of probable malingering and symptom exaggeration during neuropsychological evaluation ranged from 8% of general medical cases to 38.5% in personal injury litigants alleging mild traumatic brain injury (mTBI). Larrabee (2003) reviewed 11 studies including a total of 1363 mTBI litigants and found that on average 40% of the sample showed a suboptimal effort. Miller, Boyd, Cohn, Wilson, & McFarland (2006) found that greater than 50% of Social Security disability applicants demonstrated suboptimal effort while Chaftez (2008) found that in adult disability applicants, some 68% of claimants failed at least one SVT and 46% failed two or more SVTs. Ardolf, Denney, and Houston (2007) examined base rates in criminal defendants and found that 90% afforded suboptimal effort on one measure and 71% afforded suboptimal effort using two or more indicators.

Professional position on effort testing

The understanding that suboptimal test taker effort is relatively common and potentially (if undetected) leads to incorrect diagnosis and to misallocation of public funds (Chaftez, Abrahams, & Kohlmaier, 2007) has led the professional bodies of clinical neuropsychologists to increasingly endorse the role of effort testing as an important component of neuropsychological assessments. Professional bodies in the USA including the National Academy of Neuropsychology (NAN) (Bush et al., 2005) and the American Academy of Clinical Neuropsychology (AACN) (AACN, 2007; Heilbronner, Sweet, Morgan, Larrabee, & Millis, 2009), in the United Kingdom (British Psychological Society, 2009), and New Zealand (New Zealand Psychologists Board, 2013) have published position papers on the use of effort testing, each of which endorses the use of effort measures.

The NAN statement (Bush et al., 2005) strongly advocates the use of effort testing. The position statement notes that symptom exaggeration occurs in a sizeable minority of neuropsychological examinees with greater prevalence in forensic contexts. An adequate assessment of response validity is described as “essential” (p. 426) and it is stated that “The clinician should be prepared to justify a decision not to assess symptom validity as part of a neuropsychological evaluation,” (p. 41). The Board of Directors of the AACN (2007) emphasised the importance of assessing effort in all evaluations and the Heilbronner et al. (2009) statement reiterates both the NAN and AACN (2007) positions, stating that the “assessment of effort and genuine reporting of symptoms is important in all evaluations” (p. 1121). The statement indicates that the “use of psychometric indicators is the most valid approach to identifying neuropsychological response validity,” (p. 1106). The authors note that, “stand-alone effort measures and embedded validity indicators should both be employed,” (p. 1106). The British Psychological Society position statement concurs and directs that “Effort tests should be given routinely as part of clinical assessment of cognitive function," (p. 1). The New Zealand statement is less definitive about the routine use of effort measures but notes that when effort is to be formally assessed, multiple effort measures should be employed. Emphasis is given to the concept that there are reasons other than malingering why low effort may be applied (e.g., severe psychiatric disorder) and that other explanations for ETF need to be examined.

Impact of test taker effort

Test-taking effort has been shown to have a substantial influence on neuropsychological test performance. Green, Rohling, Lees-Haley and Allen (2001) showed that in a diverse clinical sample (*N* = 904) of compensation-seekers, 53% of the variance that was evident in a neuropsychological test battery was explained by test-taking effort (Word Memory test; WMT; Green, Iverson, & Allen, 1999). Stevens, Friedel, Mehen, and Merten (2008) examined cognitive test performances in a diverse sample of 233 adults undertaking neurological, psychiatric and psychological examinations in a German compensation-seeking or litigating context. They found that performance on effort measures (WMT and Medical Symptom Validity Test (MSVT; Green, 2004) correlated significantly with performance on all neuropsychological measures (Wechsler Adult Intelligence Scale-Revised, WAIS-R, subtests, WMS-R subtests, Trail Making Test, and the Attentional Network Test (Gauggel and Böcker, 2003). Those authors reported that effort accounted for up to 35% of the variance in cognitive domains in the entire sample and in a sub-group (*n* = 42) of those with demonstrated substantial brain injury (as defined by brain imaging abnormalities). The authors reported that after controlling for effort there was no significant effect that could be attributed to injury, a finding similar to that of Green et al. (2001).

Other similar findings have been described in the literature pertaining to the traumatic brain injury (TBI) population. Green et al. (2001) examined the impact of low effort on neuropsychological testing using a subsample of 470 compensation-seeking TBI patients.

Brain injury severity variables (Glasgow Coma Scale, loss of consciousness, positive CT or MRI findings, and posttraumatic amnesia) each accounted for at most 1% of the variance seen in neuropsychological outcome, while effort accounted for in excess of 50% of variance (Green et al., 2001). Constantinou, Bauer, Ashendorf, Fisher, and McCaffery (2005), examined the performance of a sample of litigating mTBI patients (*N* = 69) on the Halstead-Reitan Neuropsychological Battery for Adults (HNRB-A; Reitan &Wolfson, 1993). On a composite measure of the HRNB-A, 47% of the variance was accounted for by effort (Trial 2 of the TOMM).

These findings have been challenged (Bowden, Shores & Mathias, 2006). Bowden et al., (2006) argued that the conclusions represent circularity of logic and arguably, the effort measures might be assessing cognition rather than effort. Additionally, Allen, Bigler, Larsen, Goodrich-Hunsaker and Hopkins (2007) examined functional magnetic resonance data from four healthy participants completing a portion of the WMT. A reliable activation pattern was seen across all participants and was restricted to cortical areas associated with task difficulty, memory load, concentration and other forms of cognitive effort (Allen et al., 2007). These findings have been cited by those who challenge the view that the cognitive demands of effort tests are negligible (Allen et al., 2007).

A number of studies have provided strong evidence that, except for patients at the extremes of cognitive impairment, effort as assessed by symptom validity tests and cognition are largely independent and that brain damage does not account for findings of ETF. Binder, Kelly, Villanueva, & Winslow (2003) reported that the neuropsychological performance of compensation-seeking patients with mTBI who failed the Portland Digit Recognition Test (PDRT; Binder & Willis, 1991; *n* = 34) was indistinguishable from that of a well-motivated, non-compensation-seeking, moderate-to-severe TBI (MS-TBI) sample (*n* = 60) across a range of cognitive measures including WAIS-R Verbal Intelligence Quotient (VIQ), WAIS-R

Full Scale Intelligence Quotient (FSIQ), Rey Auditory Verbal Learning Test (RAVLT) measures, and Finger Tapping.

Green and Flaro (2003) found that the WMT performance of diverse groups of children with neurological, developmental and psychiatric disorders (Total $N = 135$) was not significantly different to a motivated group of independently living adults seeking child custody (*p's* > .10). They also reported that there was no effect of verbal IQ on effort test performance ($p > .10$), and the disabled children performed better on the WMT than a litigating sample of mTBI patients ($N = 197$) ($p < .001$). Green, Flaro, Brockhaus, and Montijo (2012) showed that only 5.3% of 380 children with developmental disability failed the WMT and only 4.9% of a subsample of 265 of the children failed subtests of the MSVT. Methodologically, it is not clear that the samples from these two studies (Green & Flaro, 2003 and Green et al., 2012) were independent and that they represent independent support of the findings. Additionally, although the samples of children were diagnosed with developmental disability, the average general mental ability of the group was normal (VIQ: $M = 92.6$, $SD = 15.46$) and as such the relationship between effort and general cognitive level may not have been as stridently tested in that sample as was implied.

Further support for the position that effort and cognition are dissociable comes from Ord, Greve, Bianchini & Aguerrevere (2010) who investigated the impact of TBI severity and effort in individuals administered the Wisconsin Card Sorting Test (WCST; Heaton, Chelune, Talley, Kay, & Curtiss, 1993) in samples of patients with mTBI ($n = 109$) and MS-TBI ($n = 67$). Effort during testing, as determined by performance on the PDRT, WMT, TOMM and Reliable Digit Span (RDS; Greiffenstein, Gola, & Baker, 1994) had a much larger effect on WCST ($d = 0.42$) than mTBI ($d = 0.05$) or MS-TBI ($d = 0.09$). West, Curtis, Greve and Bianchini (2011) came to a similar determination investigating the effect of effort using the PDRT and RDS on WMS-III scores in a sample of TBI patients ($N = 132$). When effort was controlled, a dose-response relationship was seen between injury severity and WMS-III scores, such that the mTBI patients applying good effort did not differ from the normative sample (Cohen's d = 0.07) while MS-TBI was associated with a moderate effect on WMS-III scores (Cohen's d = -0.52) and effort had a larger effect on WMS-III scores than injury severity (average Cohen's d = -1.27).

Finally, Fox (2011), employing a range of 25 commonly used neuropsychological measures in a diverse sample of neurological patients (N = 220) found the WMT and Computerized Test of Attention and Memory (CTAM; Fox, 1999) performances were not associated with brain damage as determined by clear radiological evidence or unequivocal evidence from neurological examination (r_{pb} = -.06 for the WMT; r^{pb} = .07 for the CTAM).

Methodological issues in the literature include the diverse samples often employed, possible sample duplication across some studies, and a preponderance of studies employing the WMT. Those issues may somewhat limit the generalizability of findings in this field. Positively, samples sizes have tended to be large and as such small effects have likely not been obscured by reduced power. When all studies are considered in conjunction, findings strongly indicate that effort has a significant impact on neuropsychological test performances, typically greater than the effect of even significant and objective brain damage.

Methods of effort testing

The first documented efforts toward psychometrically assessing neuropsychological test taker effort were those taken by André Rey (Frederick, 2003; Rey 1941, 1964). One of those measures, the Fifteen Item Test (FIT; Rey, 1964) adopted a strategy of presenting 15 easily memorised items with a recall trial. The task appears difficult but in fact is simple, taps immediate memory and attention (not learning), and because of item redundancy participants need only recall three or four concepts to perform well (Strauss, Sherman, & Spreen, 2006).

The FIT has been found to have some limitations, specifically that it is prone to false positives in those with very low IQ and serious brain trauma (Goldberg & Miller, 1986; Hays, Emmons & Stallings, 2000; Vallabhajosula & Van Gorp, 2001). Furthermore, the FIT has been shown to be relatively low in sensitivity (Boone, Salazar, Lu, Warner-Chacon, & Razani, 2002; Reznek, 2005; Vickery, Berry, Inman, Harris, & Orey, 2001). Despite having limitations, the test is rapidly administered and scored, has good specificity in the brain injured population (Millis & Kler, 1995; Inman & Berry, 2002, Reznek, 2005; Vickery et al., 2001) and now, almost fifty years since the development of the test, the FIT remains in widespread clinical usage (Slick, Tan, Strauss & Hultsch, 2004)

Slick and colleagues (2004) showed that the FIT ranked equal top with the TOMM among effort tests often employed by effort experts, while Sullivan, Lange and Dawes (2006) showed that the FIT ranked as the most frequently employed effort measure among Australian neuropsychologists. Sharland and Gfeller (2007) found that only the TOMM outranked the FIT in terms of effort measures often employed by a sample of NAN neuropsychology practitioners.

Published cut-scores that indicate ETF on the FIT vary from a low of <8 to a high of <12 (Strauss, Sherman, & Spreen, 2006), but the cut score most widely employed in the literature is <9 (Nitch & Glassmire, 2007). Sensitivity is typically slightly less than 50% and specificity levels are typically above 90% (Boone et al., 2002; Nitch & Glassmire, 2007; Vickery et al., 2001).

In the 1960s Grosz and Zimmerman (Grosz & Zimmerman, 1965; Zimmerman & Grosz, 1966) developed the use of a forced-choice paradigm for the assessment of functional blindness and in the 1970's Pankratz and colleagues (Pankratz, 1979; Pankratz, Fausti, & Peed, 1975) extended this paradigm to other functional pseudo-neurological disorders including reported memory impairment (Binder & Pankratz, 1987). The forced-choice approach, subsequently refined by Warrington (1984) and by Hiscock and Hiscock (1989), has become a successful effort testing model and has been adopted by developers of the PDRT, the TOMM, the Victoria Symptom Validity Test (VSVT; Slick, Hopp, Strauss, & Spellacy, 1996), the Computerized Assessment of Response Bias (CARB; Allen, Conder, Green & Cox, 1997), the 21-Item test (Iverson, 1998), the WMT, and the Word Choice Test from the Advanced Clinical Solutions of the WMS-Fourth Edition (Pearson Education, 2008) – tests that are today amongst the most sensitive, specific and widely employed of the effort measures (Slick et al., 2004; Sharland & Gfeller, 2007; Sullivan, Lange & Dawes, 2006; Vallabhajosula & Van Gorp, 2001).

The TOMM is one of the earliest adaptations of the Hiscock and Hiscock forced-choice paradigm. Slick et al. (2004) showed that the TOMM ranked at the top of a list of effort tests employed by effort experts, and Sullivan et al. (2006) showed that the TOMM ranked as the second most frequently employed effort measure among Australian neuropsychologists. Sharland and Gfeller (2007) found that the TOMM ranked highest among effort measures used by NAN neuropsychologists.

Although the TOMM has been shown to be sensitive to suboptimal effort with specificity rates of greater than 90% using published cut-scores (Rees, Tombaugh, & Boulay, 2001; Tombaugh 1996; 1997), the test has been criticised for lacking sensitivity (Tan, Slick, Strauss, & Hultsch, 2002). Subsequent research by Greve, Bianchini, and Doane (2006) found that published TOMM cut-offs were unnecessarily conservative. They found that by raising the cut score of 45 correct on Trial 2 or the retention trial to <48, sensitivity was increased to >.50 with no loss of specificity (>.90).

There has been widespread acceptance among clinicians that it is unsafe to rely on any single approach or instrument toward the detection of suboptimal effort during neuropsychological assessment (Lu, Rogers, & Boone, 2007). As such multiple directions of research have been followed toward developing measures of effort. Standalone forced-choice measures have been increasingly augmented by the use of embedded effort measures (Schutte & Axelrod, 2013); these are effort indicators that are embedded in neuropsychological tests already in common use. Embedded measures include patterns of performance on neuropsychological measures and regression algorithms using regression equations and discriminant function (Mittenberg, Theroux-Fichera, Zielinski, & Heilbronner, 1995; Mittenberg et al., 2001; Sherman, Boone, Lu, & Razani, 2002; Wolfe et al., 2010).

Embedded measures from the WAIS and WMS have been researched, including the RDS measure of effort, a metric extracted from Digit Span, originally devised by Greiffenstein, Gola, and Baker (1994). RDS is among the most widely researched and validated embedded metrics of effort, having been further examined and supported as a valid measure of effort using the Wechsler scales by Greiffenstein, Gola, and Baker (1995), Meyers & Volbrecht (1998), Mathias, Greve, Bianchini, Houston, & Crouch (2002), Larrabee (2003), Etherton, Bianchini, Ciota, & Greve (2005), Axelrod, Fictenberg, Millis, & Wertheimer (2006), including the fourth edition of the Wechsler scales (Miller et al., 2011; Young, Sawyer, Roper & Baughman, 2012).

RDS is among the most frequently employed effort measures. Embedded effort measures were not canvassed in the survey by Slick et al. (2004) but Sharland and Gfeller's (2007) survey identified that the RDS was ranked second of a range of embedded effort measures cited as *always used* by NAN neuropsychologists, ranking behind the California Verbal Learning Test (CVLT). That survey also found that the RDS was ranked third of all effort measures rated *always used*, ranking behind TOMM (ranked 1^{st}) and the CVLT (ranked 2^{nd}).

Cut-scores of <7 and <8 indicating low effort have been proposed for RDS. Suhr and Barrash (2007) reviewed RDS studies and found that "the vast majority" (p. 162) demonstrated a specificity of 90% or greater at a cut-score of <8, in diverse clinical samples. One study employing a cut-score <8 (Heinly, Greve, Bianchini, Love, & Brennan, 2006) found that the RDS had specificity <90% (83%) in an mTBI sample, but that same study showed 91% specificity for an MS-TBI sample. Two other studies (Larrabee, 2003; Mathias et al., 2002) found that the RDS was associated with >90% specificity in TBI samples.

Use of multiple effort indicators

Reliance on performance on a single measure of test taker effort to indicate low effort has been described as problematic due to imperfect specificity and sensitivity of effort measures (Victor, Boone, Serpa, Buehler, & Ziegler 2009). Imperfect (<100%) specificity and sensitivity of an effort measure means that any individual effort test may fail to detect low effort that in fact exists, or may indicate low effort when in fact effort was normal. The psychometric limitations of effort measures, in addition to the advantage of assessing effort at multiple time points during a testing session, have led many experts to recommend the use of multiple effort measures (Boone, 2007, 2009; Bush et al., 2005; Larrabee, 2003; Schutte & Axelrod, 2013; Vickery et al., 2004; Victor et al., 2009) to determine suboptimal effort. Failure on two or more well-validated measures of effort has been shown to be associated with >90% specificity for suboptimal effort (Chaftez, 2011; Victor et al., 2009), and failure on three or more measures has been shown to be associated with almost 100% specificity (Chaftez, 2011; Larrabee, 2003; Victor et al., 2009).

Symptom Exaggeration

Symptom exaggeration is a construct that is related but independent and separate of ETF. Symptom exaggeration relates to the *over-production* or reporting of neuropsychological symptoms rather than the under-performance of effort seen in cases of ETF. Symptom exaggeration has traditionally been assessed via psychological measures that require self-reporting of symptoms (Heilbronner et al., 2009), for example the Minnesota

Multiphasic Personality Inventory – Revised Form (MMPI-2-RF; Ben-Porath & Tellegen, 2008).

Miller (1961a, 1961b) writing on *accident neurosis,* first described the floridity of symptom reports seen in the TBI population, a picture that he described as "Gross dramatization of symptoms" (p. 922). He described patients reporting symptoms using language including "'terrible,' 'terrific,' or 'agonising'" (p. 922), also the behaviours associated with this population including "… groaning and quivering … a flaccid grip easily strengthened by distraction or encouragement; or by the patient's slumping forward with head in hands during the consultation" (p. 922). Miller described the incidence of accident neurosis in a sample of 200 TBI patients, finding that 31% of patients without any radiographic evidence of skull fracture were found to display "gross psychoneurosis" (p. 920), while only 9% of patients with simple fracture and 8% of patients with compound skull fractures, displayed this presentation. Miller's descriptions of symptom floridity went largely unexamined in the neuropsychological literature until the 1990's when symptom over-reporting in the TBI population began to be examined more systematically.

Gass (1991) and Gass and Russell (1991) examined Minnesota Multiphasic Personality Inventory (MMPI; Hathaway & McKinley, 1951) responses in samples (N = 75 and 58 respectively) of TBI patients and reported that items that were weighted toward neurological complaints constituted a factor that spanned the traditional MMPI clinical scales. Sample selection procedures and sample independence were not made clear by those authors but the results suggested that neurological symptom reporting could be identified by the MMPI. Youngjohn, Davis, and Wolf (1997), found evidence of an effect of litigation on symptom reporting on the MMPI-2 (Hathaway & McKinley, 1989) and a paradoxical effect of severity of TBI. Like Gass and Russell, symptom reporting on the MMPI-2 was elevated across a number of scales in a sample of litigating severe TBI participants (n = 18) relative to non-litigating severe TBI participants. This phenomena was exaggerated in those participants who had sustained mTBI (*n* = 30). Paniak et al. (2002) undertook a longitudinal study to compare rates of symptom reporting in a sample of non-compensation-seeking adults with mTBI (*n* = 50) with a sample of compensation-seeking adults with mTBI (*n* = 18) over the course of one-year post-injury. Symptom incidence and severity was higher in the compensation-seeking sample at intake, 3-months, and 12-months post-injury. The generalizability of findings from these studies may have been hampered by the small sample sizes but tentatively, results have indicated that symptom-reporting was elevated in compensation-seeking TBI samples, that symptom-reporting increased with the passage of time post-injury and paradoxically, that symptom-reporting was elevated in those who had sustained mTBI compared with those who had sustained MS-TBI.

Employing a large sample (*N* = 759), Greiffenstein and Baker (2006) specifically examined a hypothesis stemming from the work of Miller (1961a, 1961b), namely that TBI claimants with late post-concussion syndrome would display elevated symptom reporting. The authors included an examination of effort (TOMM, the Rey Word Recognition List (Frederick, 2003), and a grip strength task) in a compensation-seeking sample, however, traditional or empirical cut-scores to define ETF were not employed. Therefore it is not possible to translate the clinical implications of the findings or to compare results across studies. Nevertheless, florid displays of symptoms were seen in those displaying ETF. None of the patients that reported zero postconcussive symptoms were classified with ‘possible simulation’ across the three measures, while of those reporting 9-10 symptoms 53.3% were classified with ‘possible simulation’ (Greiffenstein & Baker, 2006).

Tsanadis and colleagues (2008) undertook a partial replication of the study by Paniak et al. (2002) using a larger sample (*N* = 158) and two measures of effort. Symptom reports on the four indices of the Postconcussive Symptoms Questionnaire (PCSQ; Axelrod & Lees-

Hayley, 2002) were compared in two unbalanced groups that included mTBI participants displaying ETF (*n* = 25) and a group of MS-TBI patients (*n* = 133). Across the four indices of the PCSQ the poor effort mTBI group consistently reported more symptoms with greater severity than did the MS-TBI group (Cohen's *d* range: 0.47 – 1.04). Both groups included unequal numbers of compensation-seeking and non-compensation-seeking patients and it was not clear from the description of the sample that the MS-TBI group gave an entirely optimal effort. Those factors may have inadvertently diluted effects by potentially including patients giving suboptimal effort in the MS-TBI group.

Lange, Iverson, Brooks, & Rennison (2010) reported similar findings, in a homogenous sample of compensation-seeking mTBI patients (*N* = 63) utilizing a different measure of symptom-report (Post-Concussion Scale (Lovell et al., 2006)). Participants failing the TOMM (*n* = 15) reported significantly more symptoms than did participants passing TOMM (*n* = 48) (Cohen's *d* = 0.79). Finally, these findings were supported by Lange et al. (2013) who reported that in a sample of US military service member TBI patients, those reporting sufficient symptoms to be classified positive for postconcussive disorder (*n* = 65) were significantly more likely to fail the WMT than those classified negative for postconcussive disorder (*n* = 60) (OR = 8.07, 95% CI [3.08-21.83]). That sample was quite highly selected on the basis of being military personnel and participants were selected from a larger pool for having completed personality testing. As such the sample may not be representative of the civilian population of TBI patients.

Although the literature reviewed above has tended to employ relatively small comparison groups and valid and suboptimal effort groups may have not always been clearly delineated, findings across studies are consistent and without exception indicate that symptom-reporting is elevated in compensation-seeking TBI populations. Findings also show that there is a relationship between rates of symptom-reporting and ETF such that those participants found to display ETF report higher levels of symptomatology than those that pass effort measures.

Nonverbal behaviours, symptom exaggeration and effort

Although the symptom-reporting literature described above has addressed aspects of Miller's (1961a; 1961b) hypotheses about the floridity of symptom exaggeration, the bulk of studies that have been conducted to date have examined self-reporting of symptoms or endorsing symptoms using questionnaires, checklists and inventories. Of relevance in the current context is the fact that Miller's lectures also focused on florid behavioural manifestations of symptom exaggeration.

Behavioural displays of illness aside from self-reports of symptoms are assessed in medical examinations. In the chronic pain population abnormal behaviours (Waddell, McCulloch, Kummel, & Venner, 1980; Waddell, Somerville, Henderson, & Newton, 1992) are examined and when displayed by any patient, are thought to reflect a process of symptom exaggeration – "a magnified or more emphatic presentation of the severity of their problem" (Waddell, Pilowsky, & Bond, 1989, p. 50). These behaviours, described as *Waddell signs* include the patient displaying disproportionate facial expressions of pain, also muscle tension and tremor (Waddell et al., 1980). In the neurological examination *giveaway* weakness (Gould, Miller, Goldberg, & Benson, 1986) and *Hoover's sign* (Ziv, Djaldetti, Zoldan, Avraham, & Melamed, 1998) present as behavioural markers of nonorganic or functional impairment (Gould et al., 1986). In the case of giveaway weakness, sudden cessation of isotonic contraction is evident during motor testing and in the case of Hoover's sign, a supposedly paralysed limb can be seen to exert downward stabilising force when the healthy contralateral limb is raised.

There has been a recent call to examine and track aspects of behaviour during neuropsychological testing and to relate those behaviours to effort test performance

(Denning, 2013). Nonverbal behaviours have been fruitfully investigated as markers of deception in the broader social psychology literature. Ekman and O'Sullivan (1991) described their investigations into the ability of professional groups to detect lying. Professionals (Secret Service agents) who paid attention to nonverbal cues of lying (including displaying a strained voice, avoiding eye contact) had significantly increased accuracy in detecting lying over those other professionals who relied on verbal indicators. This work led to further studies of behavioural markers of deception and lying (Ekman, O'Sullivan & Frank, 1999; Frank & Ekman, 1997; Frank & Ekman, 2004). Ekman and O'Sullivan (2006) reviewed in detail the nonverbal aspects of behaviour associated with deception and with malingering. The authors described four aspects of facial expressions reliably associated with deception, lying and concealment of emotion – the morphology of the expression, the timing, duration and speed of onset of the facial expression, the symmetry of the expression, and smoothness of the trajectory of the expression. Notably, non-malingered facial expressions engaged all related facial muscle groups while malingered facial expressions involved a constrained set of muscle groups. Genuine facial expressions tended to be symmetrical, developed slowly and were prolonged, and appeared smooth without a jagged or stepped trajectory. Ekman and O'Sullivan (2006) emphasized that malingering can be discernible via abnormal overt behaviours.

To date the relationship between displays of abnormal behaviours and ETF has been barely examined in the TBI population. Some research has examined abnormal (non-organic) jerky movements and their relationship to effort and ETF in other neurological and pseudo-neurological populations. Heintz et al. (2013) examined ETF displayed on the Amsterdam Short Term Memory Test (ASTM; Schmand & Lindeboom, 2005), in a sample of participants displaying jerky movements thought not to have any neurological cause (*n* = 26). The authors reported a higher rate of ETF in that group compared to those with an established diagnosis of Gilles de la Tourette syndrome (n = 16; p = 0.02) or healthy controls (n = 22; p = 0.04). That study included only a small, rarely encountered clinical sample and employed only one measure of effort, leading to the risk of false-positive and false-negative ETF findings. As such, the findings while tentative do suggest a relationship between pseudo-neurological abnormal behaviours and ETF. The study also pointed to the need for further investigations into the abnormal non-organic behaviours that can be seen in clinical neuropsychology populations.

Effort studies have also examined the performance of patients displaying psychogenic nonepileptic seizure-like behaviours (PNES). PNES is associated with the display of florid abnormal behaviours including shaking, writhing, jerking, or rocking, or an episode of unresponsiveness in the absence of any electrographic ictal discharge on EEG that would suggest an epileptic event was occurring (Williamson, Holsman, Chaytor, Miller, & Drane, 2012). A number of studies have shown that individuals displaying PNES perform poorly on effort measures (Binder, Salinsky & Smith, 1994; Binder, Kindermann, Heaton, & Salinsky, 1998; Cragar, Berry, Fakhoury, Cibula, & Schmitt, 2006; Drane et al., 2006; Hill, Ryan, Kennedy, & Malamut, 2003; Williamson, Drane, Stroup, Miller, & Holmes, 2003; Williamson, Holsman, Chaytor, Miller, & Drane, 2012), although not all have reported that finding (Strutt, Hill, Scott, Uber-Zak, & Fogel, 2011). Rates of single-measure ETF in PNES samples vary from 3% (Strutt et al., 2011) to 64% (Williamson, Drane, Stroup, Miller, & Holmes, 2003). Methodologies vary across studies with most employing only one measure of effort (Binder, Salinsky & Smith, 1994; Binder et al., 1998; Drane et al., 2006; Locke et al., 2006; Strutt et al., 2011; Williamson et al., 2012). Cragar and colleagues (2006) employed multiple effort measures and reported that 19% of their PNES sample failed two or more effort measures. Unfortunately, with only one exception (Williamson et al., 2012), the compensation-seeking status of participants in those studies was not documented, however a number of studies incorporated veterans affairs patients (Binder, Salinsky & Smith, 1994; Binder et al., 1998), who may have been compensation-seeking. Sample variability, including the variability of access to secondary gains, may partially account for the wide disparity in findings in respect of incidence of ETF in these samples (Williamson et al., 2012).

In one of the few studies of abnormal behaviours seen in cases of TBI, Cottingham and Boone (2010) reported on the display of atypical speech behaviours in a case following mTBI. The patient displayed an abrupt, hypernasal, and halting pattern of speech that was variable and inconsistent and did not follow “any known neurobehavioral pattern of actual speech/language dysfunction” (p. 1017). During cognitive testing the patient displayed ETF on three separate effort indicators. Similarly, Binder, Spector, and Youngjohn (2012) have reported on a small series (N = 3) of cases of very atypical and non-organic speech-pattern abnormalities following mTBI, evaluated in the context of personal injury lawsuits. Each of the three cases displayed ETF across a large range of effort measures including standalone and embedded measures. Axelrod (2009) reported on a non-litigating and non-compensation-seeking case of abnormal speech behaviours following mTBI. In this case ETF was evident on a number of measures including below-chance performance on the Warrington Recognition Memory Test (WRMT; Warrington, 1984). Although case studies form a low level of evidence (Phillips et al., 2001) and therefore no strong conclusions can be taken from the few TBI case studies that have been published, each has suggested that abnormal illness behaviours and ETF might be associated in TBI patients.

In a recent and novel study, Denning (2013) examined the behaviours of TBI patients (N = 151) completing the TOMM. Denning hypothesised that tracking atypical behavioural responses in addition to effort test performance may increase the sensitivity of effort measures. Denning found that by combining behavioural responses (pointing and naming items in contravention of the test instructions) with performance on the first 10 items of the TOMM (TOMMe10), sensitivity of the TOMM was increased 7% over sensitivity seen with the traditional scoring of the TOMM. Denning's research employed only the TOMM and was specifically focussed on atypical behaviours displayed during the administration of that measure. Thus, the research had a very specific focus, aimed at increasing the sensitivity of the TOMM. Whether a broad range of abnormal behaviours can be expected during administration of the TOMM and other measures of effort remains unclear.

Taken as a whole, there have been suggestions from the literature that certain abnormal behaviours in clinical populations may be detectable and associated with ETF. Although findings remain very preliminary they raise the possibility of a novel and as yet largely unexplored behavioural tool that may assist in the detection of symptom exaggeration and which might be predictive of ETF. Further preliminary investigations into the abnormal behaviours that may have some association with ETF using a TBI clinical sample and validated measures of ETF are indicated.

Differentiating ETF and symptom exaggeration from malingering

The Diagnostic and Statistical Manual of Mental Disorders (5th Ed.) (APA, 2013) defines malingering as the intentional production of false or grossly exaggerated physical or psychological symptoms, motivated by external incentives. In the context of the neuropsychological evaluation, the physical or psychological symptoms produced by the malingerer may include exaggerated symptom reporting and/or ETF, both of which present a false or exaggerated picture of impairment and disability (Bush et al., 2005).

The presence of secondary gains, usually financial incentives, is known to be associated with poor outcome following brain injury (Belanger, Curtiss, Demery, Lebowitz, & Vanderploeg, 2005; Binder & Rohling, 1996; Carroll et al., 2004) and studies of ETF in the TBI population have demonstrated that ETF is associated with the availability of financial incentives. Moreover, it has been reliably demonstrated that there is a dose-response relationship between financial incentives and performance on effort measures (e.g., Bianchini, Curtis, & Greve, 2006). That literature is reviewed more extensively in Chapter 3.

Faust, Ahern, Bridges and Yonce (2012) emphasise, in-keeping with the DSM conceptualisation, that the determination of malingering requires not just misrepresentation of one's health status and the presence of secondary gains, but also determination of intentionality. The misrepresentation of one's health status is reflected in suboptimal effort during testing and in symptom exaggeration (Faust et al., 2012). Determining the intentionality of the misrepresentation is more challenging because the demand of that test is to appraise the internal state of mind of an examinee (Delis & Wetter, 2007). In most cases it is difficult if not impossible to ascertain the internal state of an examinee in any objective manner (Delis & Wetter, 2007; Shapiro & Teasell, 2004). Modern symptom validity tests that adopt the forced-choice model afford the opportunity to gain some information in the rare cases where an individual scores below chance levels. Below chance performance on a forced choice measure indicates that the examinee almost certainly knew the correct answer but deliberately or consciously chose the incorrect answer (Boone, 2007). In such cases intentionality is established by a probabilistic statistical test – the performance falls below chance levels using probability statistics for the binomial distribution (Grote & Hook, 2007).

Slick, Sherman, and Iverson, (1999) employed this strategy in their classification criteria for malingered neurocognitive dysfunction, suggesting that malingering could be defined as the volitional exaggeration or fabrication of cognitive dysfunction for secondary gain. They further indicated that definite response bias, as defined by below chance performance on a forced choice measure, was an essential criterion of malingered neurocognitive dysfunction. Others (Boone, 2007; Martin, Bolter, Todd, Gouvier, & Niccols, 1993) have noted that below chance performance is a rare occurrence and only relatively unsophisticated examinees perform so poorly on forced-choice effort measures. Accordingly, Boone (2007) and Larrabee et al. (2007, 2008) have argued that the Slick, Sherman, and Iverson (1999) criterion for the definite malingering of neurocognitive dysfunction that relates to the presence of definite response bias should be extended to include failure on several well-validated effort indices. In keeping with that recommendation, Slick and Sherman (2013) have recently broadened the criteria for malingered neurocognitive dysfunction to include evidence of "one or more very strong indicators of exaggeration/fabrication of neuropsychological problems or deficits" (p. 63), where those indicators represent below chance performance on forced choice measures or high *posterior probability* (≥0.95) on one or more well-validated psychometric indices that performance is below actual ability level. Thus, under certain conditions, poor performance on effort measures serves to indicate both the misrepresentation of health status and the intentionality of that misrepresentation.

The final test of malingering, as established by Slick and Sherman (2013), is whether the behaviours (including ETF) cannot be accounted for by psychiatric, neurological or developmental factors. Slick and Sherman identify that such psychological constructs as conversion disorder, dissociative amnesia, factitious disorder, fabricating symptoms for psychosocial (attention) gains, cogniform condition, neurocognitive hypochondriasis, stereotype threat and oppositional-defiant presentations may all be associated with symptom exaggeration and ETF but not be indicative of malingering. Barker, Horner, & Bachman (2010) also note that ETF might be seen in circumstances other than malingering including lack of interest, opposition to testing, fatigue, lack of understanding of the purposes of testing by the patient and motivation to maintain a "sick role." Iverson (2010) suggested that psychological variables including a sense of entitlement, justification, neediness, anger, frustration, greed, reinforced behaviour patterning, depressive negativistic thinking, personality characteristics and disorders, nocebo effect (negative expectations about sickness that are causally linked to symptoms), misattribution, and ‘good old days’ bias in thinking, could be associated with symptom exaggeration and ETF. Clearly, ETF represents an essential but not sufficient criterion for malingering and other causes of ETF are possible.

Conclusion

Effort testing has a long history of use in neuropsychological assessment as a means of validating, or invalidating as the case may be, the results of cognitive assessment. Effort testing has been shown to be valid and reliable and has been stridently espoused by professional neuropsychology bodies because it has been repeatedly shown that affording an insufficient effort during neuropsychological testing is very common, particularly so in cases where secondary gains are evident. Effort tests can be standalone measures or embedded in neuropsychological tests and there is an increasing call to measure effort using multiple measures and not rely on any one test performance in the classification of low effort. ETF is seen in cases of malingering but malingering cases are only a subset of those displaying low effort and there are other, non-malingering variables, thought to be related to ETF.

Symptom exaggeration is a construct that is related to, but independent of, ETF. Symptom exaggeration in the form of florid reports of symptomatology is known to be related to ETF, as are some other abnormal illness-related behaviours including non-neurological speech patterns, abnormal jerky movements, and psychogenic seizure-like behaviours. The overt behaviours that are related to ETF in TBI patients have yet to be examined even at a cursory level.

There is a need to identify the diverse variables that might be associated with ETF in order to expand understandings of effort and ETF during neuropsychological testing. An improved understanding of the variables that are associated with ETF will allow the development of a comprehensive model of effort during neurocognitive testing.

References

AACN Board of Directors. (2007). American Academy of Clinical Neuropsychology (AACN) practice guidelines for neuropsychological assessment and consultation. *The Clinical Neuropsychologist, 21*(2), 209-231.

Allen, L.M., Conder, R.L., Green, P., & Cox, D.R. (1997). *CARB' 97 Manual for the Computerized Assessment of Response Bias (CARB)*. Durham, NC: CogniSyst, Inc.

Allen, M. D., Bigler, E. D., Larsen, J., Goodrich-Hunsaker, N. J., & Hopkins, R. O. (2007). Functional neuroimaging evidence for high cognitive effort on the Word Memory Test in the absence of external incentives. *Brain Injury, 21*, 1425-1428.

American Psychiatric Association. (2013). *DSM-V*. Washington DC: American Psychiatric Association.

Ardolf, B. R., Denney, R. L., & Houston, C. M. (2007). Base rates of negative response bias and malingered neurocognitive dysfunction among criminal defendants referred for neuropsychological evaluation. *The Clinical Neuropsychologist, 21*(6), 899-916.

Axelrod, B. N. (2009). Fabrication of psychiatric symptoms: Somatoform and psychotic disorders. In J. J. Morgan, & J. J. Sweet (Eds.), *Neuropsychology of malingering casebook* (pp. 180–194). New York: Psychology Press

Axelrod, B. N., & Lees-Haley, P. (2002). Construct validity of the PCSQ as related to the MMPI-2. Archives of Clinical Neuropsychology, 17, 343–350.

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.

Babikian, T., & Boone, K. B. (2007). Intelligence tests as measures of effort. In K. B. Boone (Ed.), *Assessment of feigned cognitive impairment: A neuropsychological perspective* (pp. 103-127). New York: Guilford Publications, Inc.

Barker, M. D., Horner, M. D., & Bachman, D. L. (2010). Embedded indices of effort in the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS) in a geriatric sample. *The Clinical Neuropsychologist*, *24*, 1064-1077.

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*, 215–227.

Ben-Porath, Y. S., & Tellegen, A. (2008). *Minnesota Multiphasic Personality Inventory-2-Restructured Form (MMPI-2-RF): Manual for administration, scoring, and interpretation*. Minneapolis, MN: University of Minnesota Press.

Bianchini K. J., Curtis, K. L., & Greve, K. W. (2006.) Compensation and malingering in traumatic brain injury: A dose-response relationship? *The Clinical Neuropsychologist, 20*, 831-847.

Binder, L. M., & Pankratz, L. (1987). Neuropsychological evidence of a factitious memory complaint. *Journal of Clinical and Experimental Neuropsychology*, *9*, 167-171.

Binder, L. M., Spector, J., & Youngjohn, J. R. (2012). Psychogenic stuttering and other acquired nonorganic speech and language abnormalities. *Archives of Clinical Neuropsychology*, *27*, 557-568.

Binder, L. M., & Willis, S. C. (1991). Assessment of motivation after financially compensable minor head trauma. *Psychological Assessment*, *3*, 175-181.

Binder, L. M., Kelly, M. P., Villanueva, M. R., & Winslow, M. M. (2003). Motivation and neuropsychological test performance following mild head injury. *Journal of Clinical and Experimental Neuropsychology*, *25*, 420-430.

Binder, L. M., Kindermann, S. S., Heaton, R. K., & Salinsky, M. C. (1998). Neuropsychologic impairment in patients with nonepileptic seizures. *Archives of*

Clinical Neuropsychology, 13, 513-522.

Binder, L. M., & Rohling, M. L. (1996). Money matters: a meta-analytic review of the effects of financial incentives on recovery after closed-head injury. *American Journal of Psychiatry, 153*, 7-10.

Binder, L. M., Salinsky, M. C., & Smith, S. P. (1994). Psychological correlates of nonepileptic seizures. *Journal of Clinical and Experimental Neuropsychology, 16*, 524–530.

Boone, K. (2011). Somatoform Disorders, Factitious Disorder, and Malingering. In M. R. Schoenberg & J. G. Scott (Eds.), *The little black book of neuropsychology* (pp. 551-566). New York: Springer

Boone, K. B. (2007). A reconsideration of the Slick et al., (1999) criteria for malingered neurocognitive dysfunction. In K. B. Boone (Ed.), *Assessment of feigned cognitive impairment: A neuropsychological perspective* (pp. 29-49). New York: Guilford Publications, Inc.

Boone, K. B. (2009). The need for continuous and comprehensive sampling of effort/response bias during neuropsychological examinations. *The Clinical Neuropsychologist*, *23*, 729-741.

Boone, K. B., Salazar, X., Lu, P., Warner-Chacon, K., & Razani, J. (2002). The Rey 15-item recognition trial: A technique to enhance sensitivity of the Rey 15-item memorization test. *Journal of Clinical and Experimental Neuropsychology*, *24,* 561-573.

Bowden, S. C., Shores, E. A., & Mathias, J. L. (2006). Does effort suppress cognition after traumatic brain injury? A re-examination of the evidence for the Word Memory Test. *The Clinical Neuropsychologist*, *20*, 858-872.

British Psychological Society (2009). *Assessment of effort in clinical testing of cognitive functioning for adults.* Leister: British Psychological Society.

Bush, S. S., Ruff, R. M., Tröster, A. I., Barth, J. T., Koffler, S. P., Pliskin, N. H., Reynolds, C. R., & Silver, C. H. (2005). Symptom validity assessment: Practice issues and medical necessity: NAN Policy & Planning Committee. *Archives of Clinical Neuropsychology*, *20,* 419-426.

Carroll, L., Cassidy, J. D., Peloso, P., Borg, J., Von Holst, H., Holm, L., Paniak, C., & Pepin, M. (2004). Prognosis for mild traumatic brain injury: results of the WHO Collaborating Centre Task Force on Mild Traumatic Brain Injury. *Journal of Rehabilitation Medicine*, *36*, 84-105.

Chafetz, M. D. (2008). Malingering on the social security disability consultative exam: Predictors and base rates. *The Clinical Neuropsychologist*, *22*, 529-546.

Chafetz, M. (2011). Reducing the probability of false positives in malingering detection of Social Security disability claimants. *The Clinical Neuropsychologist*, *25*, 1239-1252.

Chafetz, M. D., Abrahams, J. P., & Kohlmaier, J. (2007). Malingering on the Social Security Disability consultative exam: A new rating scale. *Archives of Clinical Neuropsychology*, *22*, 1–14.

Chafetz, M. D., Prentkowski, E., & Rao, A. (2011). To work or not to work: Motivation (not low IQ) determines symptom validity test findings. *Archives of Clinical Neuropsychology*, *26*, 306-313.

Constantinou, M., Bauer, L., Ashendorf, L., Fisher, J. M., & McCaffrey, R. J. (2005). Is poor performance on recognition memory effort measures indicative of generalized poor performance on neuropsychological tests? *Archives of Clinical Neuropsychology*, *20*, 191-198.

Cottingham, M. E., & Boone, K. B. (2010). Non-credible language deficits following mild traumatic brain injury. *The Clinical Neuropsychologist*, *25*, 1006–1025.

Cragar, D. E., Berry, D. T., Fakhoury, T. A., Cibula, J. E., & Schmitt, F. A. (2006).

Performance of patients with epilepsy or psychogenic non-epileptic seizures on four measures of effort. *The Clinical Neuropsychologist*, *20*, 552–566.

Delis, D. C., & Wetter, S. R. (2007). Cogniform disorder and cogniform condition: Proposed diagnoses for excessive cognitive symptoms. *Archives of Clinical Neuropsychology*, *22*, 589-604.

Denning, J. H. (in press). Combining the Test of Memory Malingering Trial 1 with behavioral responses improves the detection of effort test failure. *Applied Neuropsychology*.

Drane, D. L., Williamson, D. J., Stroup, E. S., Holmes, M. D., Jung, M., Koerner, E., Chaylor, N., Wilensky, A. J., & Miller, J. W. (2006). Cognitive impairment is not equal in patients with epileptic and psychogenic nonepileptic seizures. *Epilepsia*, *47,* 1879-1886.

Ekman, P., & O'Sullivan, M. (1991). Who can catch a liar? *American Psychologist*, *46*(9), 913-920.

Ekman, P., & O'Sullivan, M. (2006). From flawed self-assessment to blatant whoppers: the utility of voluntary and involuntary behavior in detecting deception. *Behavioral Sciences & the Law*, *24,* 673-686.

Ekman, P., O'Sullivan, M., & Frank, M. G. (1999). A few can catch a liar. *Psychological Science*, *10,* 263-266.

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*, 101-106.

Faust, D., Ahern, D. C., Bridges, A. J., & Yonce, L. J. (2012). Assessment of Malingering and Falsification: Pushing the Boundaries of Knowledge in Research and Clinical Practice. In C. R. Reynolds & A. M. Horton Jr. (Eds.), *Detection of Malingering* *during Head Injury Litigation* (2nd ed.) (pp. 1-46). New York: Springer.

Fox, D. D. (2009). Manual for the Computerized Test of Attention and Memory (8th ed.). Glendale, CA: D. D. Fox.

Fox, D. D. (2011). Symptom validity test failure indicates invalidity of neuropsychological tests. *The Clinical Neuropsychologist*, *25*(3), 488-495.

Frank, M. G., & Ekman, P. (1997). The ability to detect deceit generalizes across different types of high-stake lies. *Journal of personality and social psychology*, *72*, 1429.

Frank, M. G., & Ekman, P. (2004). Appearing truthful generalizes across different deception situations. *Journal of Personality and Social Psychology,86*, 486-495.

Frederick, R. I. (2003). A review of Rey's strategies for detecting malingered neuropsychological impairment. *Journal of Forensic Neuropsychology*, *2*(3-4), 1-25.

Gass, C. S. (1991). MMPI-2 interpretation and closed head injury: A correction factor. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, *3*, 27-31.

Gass, C. S., & Russell, E. W. (1991). MMPI profiles of closed head trauma patients: Impact of neurologic complaints. *Journal of Clinical Psychology*, *47*, 253-260.

Gauggel, S., Böcker, M., 2003. *Der Aufmerksamkeits-Netzwerk-Test*. Germany: Neuropsychologisches Zentrum Kassel.

Goldberg, J. O., & Miller, H. R. (1986). Performance of psychiatric inpatients and intellectually deficient individuals on a task that assesses the validity of memory complaints. *Journal of Clinical Psychology*, *42*, 792-795.

Gould, R., Miller, B.L., Goldberg, M.A., & Benson, D.F. (1986). The validity of hysterical signs and symptoms. *Journal of Nervous and Mental Disease, 174*, 593-597.

Green P. (2004). *Manual for the Medical Symptom Validity Test*. Edmonton, Alberta, Canada: Green's Publishing.

Green, P., & Flaro, L. (2003). Word Memory Test performance in children. *Child Neuropsychology*, *9*, 189-207.

Green, P., Flaro, L., Brockhaus, R., & Montijo, J. (2012). Performance on the WMT, MSVT, and NV-MSVT in children with developmental disabilities and in adults wit mild traumatic brain injury. In C. R. Reynolds & A. M. Horton Jr. (Eds.), *Detection of Malingering during Head Injury Litigation* (2nd ed.) (pp. 201-219). New York: Springer.

Green, P., Iverson, G. L., & Allen, L. (1999). Detecting malingering in head injury litigation with the Word Memory Test. *Brain Injury, 13,* 813-819.

Green, P., Rohling, M. L., Lees-Haley, P. R., & Allen, L. M. A. (2001). Effort has a greater effect on test scores than severe brain injury in compensation claimants. *Brain Injury, 15*, 1045-1060.

Greiffenstein, M. F., & Baker, W. J. (2006). Miller was (mostly) right: Head injury severity inversely related to simulation. *Legal and Criminological Psychology*, *11*, 131-145.

Greiffenstein, M. F., Baker, W. J., & Gola, T. (1994). Validation of malingered amnesia measures with a large clinical sample. *Psychological Assessment*, *6*, 218-224.

Greiffenstein, M. F., Gola, T., & Baker, W. J. (1995). MMPI-2 validity scales versus domain specific measures in detection of factitious traumatic brain injury. *The Clinical Neuropsychologist*, *9*, 230-240.

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*, 1176–1190.

Grosz, H. J., & Zimmerman, J. (1965). Experimental analysis of hysterical blindness: A follow-up report and new experimental data. *Archives of General Psychiatry*, *13*, 255.

Grote, C. L., & Hook, J. N. (2007). Forced-choice recognition tests of malingering. In G. J.

Larrabee (Ed.), *Assessment of malingered neuropsychological deficits* (pp. 44–79). New York, NY: Oxford University Press.

Hathaway, S. R., & McKinley, J. C. (1951). *Minnesota Multiphasic Personality Inventory manual (Revised).* San Antonio, TX: Psychological Corporation.

Hathaway, S. R., & McKinley, J. C. (1989). *MMPI-2: Minnesota Multiphasic Personality Inventory-2: manual for administration and scoring*. Minneapolis, MN: University of Minnesota Press.

Hays, J. R., Emmons, J., & Stallings, G. (2000). Dementia and mental retardation markers on the Rey 15-item visual memory test. *Psychological Reports*, *86*, 179-182.

Heaton, R. K., Chelune, G. J., Talley, J. L., Kay, G. G., & Curtiss, G. (1993). *Wisconsin Card Sorting Test manual* (Rev. ed.). Odessa, FL: Psychological Assessment Resources.

Heilbronner, R. L., Sweet, J. J., Morgan, J. E., Larrabee, G. J., Millis, S. R., & Conference Participants 1. (2009). American Academy of Clinical Neuropsychology Consensus Conference Statement on the neuropsychological assessment of effort, response bias, and malingering. *The Clinical Neuropsychologist*, *23*, 1093-1129.

Heinly, M. T., Greve, K. W., Bianchini, K. J., Love, J. M., & Brennan, A. (2006). WAIS Digit Span-based indicators of malingered neurocognitive dysfunction: Classification accuracy in traumatic brain injury. *Assessment, 12,* 429-444.

Heintz, C. E., van Tricht, M. J., van der Salm, S. M., van Rootselaar, A. F., Cath, D., Schmand, B., & Tijssen, M. A. (2013). Neuropsychological profile of psychogenic jerky movement disorders: importance of evaluating non-credible cognitive performance and psychopathology. *Journal of Neurology, Neurosurgery & Psychiatry, 84,* 862-867.

Hill, S. K., Ryan, L. M., Kennedy, C. H., & Malamut, B. L. (2003). The relationship between measures of declarative memory and the Test of Memory Malingering in patients with and without temporal lobe dysfunction. *Journal of Forensic Neuropsychology*, *3*, 1-18.

Hiscock, M., & Hiscock, C. K. (1989). Refining the forced-choice method for the detection of malingering. *Journal of Clinical and Experimental Neuropsychology, 11*, 967-974.

Inman, T. H., & Berry, D. T. (2002). Cross-validation of indicators of malingering: A comparison of nine neuropsychological tests, four tests of malingering, and behavioral observations. *Archives of Clinical Neuropsychology*, *17*, 1-23.

Iverson, G. L. (1998). *21-Item Test research manual*. Vancouver, B.C.: Author.

Iverson, G. L. (2010). Detecting exaggeration, poor effort, and malingering in neuropsychology. In A. MacNeill Horton & L. C. Hartlage (Eds.), *The handbook of forensic neuropsychology* (2nd ed.) (pp. 91-136). New York: Springer Publishing Company.

Lange, R. T., Brickell, T., French, L. M., Ivins, B., Bhagwat, A., Pancholi, S., & Iverson, G. L. (2013). Risk factors for postconcussion symptom reporting after traumatic brain injury in US military service members. *Journal of Neurotrauma*, *30*, 237-246.

Lange, R. T., Iverson, G. L., Brooks, B. L., & Rennison, V. 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,* 961–972.

Larrabee, G. J. (2003). Detection of malingering using atypical performance patterns on standard neuropsychological tests. *The Clinical Neuropsychologist*, *17*(3), 410-425.

Larrabee, G. J. (2008). Aggregation across multiple indicators improves the detection of malingering: Relationship to likelihood ratios. *The Clinical Neuropsychologist*, *22*, 666-679.

Larrabee, G. J., Greiffenstein, M. F., Greve, K. W., & Bianchini, K. J. (2007). Refining diagnostic criteria for malingering. In G. J. Larrabee (Ed.), *Assessment of malingered neuropsychological deficits* (pp. 334-371). Oxford: Oxford University Press.

Lezak, M. D., Howieson, D. B., Loring, D. W., Hannay, H. J., & Fischer, J. S. (2004). *Neuropsychological assessment* (4th ed.). New York: Oxford University Press.

Locke, D. E., Berry, D. T., Fakhoury, T. A., & Schmitt, F. A. (2006). Relationship of indicators of neuropathology, psychopathology, and effort to neuropsychological results in patients with epilepsy or psychogenic non-epileptic seizures. *Journal of Clinical and Experimental Neuropsychology*, *28*, 325-340.

Lovell, M. R., Iverson, G. L., Collins, M. W., Podell, K., Johnston, K. M., Pardini, D., Norwig, J., & Maroon, J. C. (2006). Measurement of symptoms following sports-related concussion: Reliability and normative data for the post-concussion scale. *Applied Neuropsychology*, *13*, 166–174.

Lu, P., Rogers, S. A., & Boone, K. B. (2007). Use of standard memory tests to detect suspect effort. In K. B. Boone (Ed.), *Assessment of feigned impairment: A neuropsychological perspective* (pp. 128-151). New York: Guilford Press.

Martin, R. C., Bolter, J. F., Todd, M. E., Gouvier, W. D., & Niccolls, R. (1993). Effects of sophistication and motivation on the detection of malingered memory performance using a computerized forced-choice task. *Journal of Clinical and Experimental Neuropsychology*, *15*, 867-880.

Mathias, C. W., Greve, K. W., Bianchini, K. J., Houston, R. J., & Crouch, J. A. (2002). Detecting malingered neurocognitive dysfunction using the reliable digit span in traumatic brain injury. *Assessment*, *9*, 301-308.

Meyers, J. E., & Volbrecht, M. (1998). Validation of reliable digits for detection of malingering. *Assessment*, *5*, 303-307.

Meyers, J. E., & Volbrecht, M. E. (2003). A validation of multiple malingering detection methods in a large clinical sample. *Archives of Clinical Neuropsychology*, *18*, 261-276.

Miller, H. (1961a). Accident neurosis: Lecture I. *British Medical Journal, 1*(5230), 919-925.

Miller, H. (1961b). Accident neurosis: Lecture II. *British Medical Journal, 1*(5231), 992-998.

Miller, J. B., Millis, S. R., Rapport, L. J., Bashem, J. R., Hanks, R. A., & Axelrod, B. N. (2011). Detection of insufficient effort using the Advanced Clinical Solutions for the Wechsler Memory Scale. *The Clinical Neuropsychologist*, *25*, 160-172.

Miller, L. S., Boyd, M. C., Cohn, A., Wilson, J. S., & McFarland, M. (2006). *Prevalence of sub-optimal effort in disability applicants.* Poster presented at the meeting of the International Neuropsychological Society, Boston. MA.

Millis, S. R., & Kler, S. (1995). Limitations of the Rey Fifteen-Item Test in the detection of malingering. *The Clinical Neuropsychologist, 9,* 241-244.

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,* 1094-1102.

Mittenberg, W., Theroux, S., Aguila-Puentes, G., Bianchini, K., Greve, K., & Rayls, K. (2001). Identification of malingered head injury on the Wechsler Adult Intelligence Scale. *The Clinical Neuropsychologist*, *15*, 440.

Mittenberg, W., Theroux-Fichera, S., Zielinski, R. E., & Heilbronner, R. L. (1995). Identification of malingered head injury on the Wechsler Adult Intelligence Scale: Revised. *Professional Psychology, Research and Practice*, *26*, 491-498.

New Zealand Psychologists Board (2013). *Guidelines on the use of psychometric tests February 2013*. Wellington: New Zealand Psychologists Board.

Nitch, S. R., & Glassmire, D. M. (2007). Non-forced-choice measures to detect noncredible cognitive performance. In K. B. Boone (Ed.), *Assessment of feigned cognitive impairment: A neuropsychological perspective* (pp. 78-86). New York: The Guilford Press.

Ord, J. S., Greve, K. W., Bianchini, K. J., & Aguerrevere, L. E. (2010). Executive dysfunction in traumatic brain injury: The effects of injury severity and effort on the Wisconsin Card Sorting Test. *Journal of Clinical and Experimental Neuropsychology, 32,* 132-140.

Paniak, C., Reynolds, S., Toller-Lobe, G., Melnyk, A., Nagy, J., & Schmidt, D. (2002). A longitudinal study of the relationship between financial compensation and symptoms after treated mild traumatic brain injury. *Journal of Clinical and Experimental Neuropsychology*, *24,* 187-193.

Pankratz, L. (1979). Symptom validity testing and symptom retraining: procedures for the assessment and treatment of functional sensory deficits. *Journal of Consulting and Clinical Psychology*, *47*, 409-410.

Pankratz, L., Fausti, S. A., & Peed, S. (1975). A forced-choice technique to evaluate deafness in the hysterical or malingering patient. *Journal of Consulting and Clinical Psychology*, *43*, 421-422.

Pearson Education. (2008). *Wechsler Memory Scale, Fourth Edition: Clinical features of the new edition*. New York: Pearson Education.

Phillips, B., Ball, C., Sackett, D., Badenoch, D., Straus, S., Haynes, B., & Dawes, M. (2001). *Oxford Centre for Evidence-Based Medicine levels of evidence (May 2001)*. Oxford: Centre for Evidence-Based Medicine. Available at: http://www.cebm.net/levels_of_evidence.asp.

Rees, L. M., Tombaugh, T. N., & Boulay, L. (2001). Depression and the Test of Memory Malingering. *Archives of Clinical Neuropsychology, 16*, 501-506.

Reitan, R. M., & Wolfson, D. (1993). *The Halstead–Reitan Neuropsychological Test Battery: Theory and clinical interpretation* (2nd ed.). Tucson, AR: Neuropsychology Press.

Rey, A. (1941). Psychological examination of traumatic encephalopathy. *Archives de Psychologie, 28*, 286–340.

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

Reznek, L. (2005). The Rey 15-item memory test for malingering: A meta-analysis. *Brain Injury*, *19*, 539-543.

Schmand, B., & Lindeboom, J. (2005). *Amsterdam Short Term Memory Test*. Leiden: PITS.

Schutte, C., & Axelrod, B. N. (2013). Use of embedded cognitive symptom validity measures in mild traumatic brain injury cases. In D. A. Carone & S. S. Bush (Eds.), *Mild traumatic brain injury: Symptom validity assessment and malingering* (pp. 159-181). New York: Springer

Shapiro, A. P., & Teasell, R. W. (2004). Behavioural interventions in the rehabilitation of acute v. chronic non-organic (conversion/factitious) motor disorders. *The British Journal of Psychiatry*, *185*, 140-146.

Sharland, M. J., & Gfeller, J. D. (2007). A survey of neuropsychologists' beliefs and practices with respect to the assessment of effort. *Archives of Clinical Neuropsychology*, *22*, 213-223.

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*, 242-250.

Slick, D. J., Hopp, G., Strauss, E., & Spellacy, F. J. (1996). Victoria Symptom Validity Test: Efficiency for detecting feigned memory impairment and relationship to neuropsychological tests and MMPI-2 validity scales. *Journal of Clinical and Experimental Neuropsychology*, *18*, 911-922.

Slick, D. J., & Sherman, E. M. S. (2013). Differential diagnosis of malingering. In D. A. Carone & S. S. Bush (Eds.), *Mild traumatic brain injury: Symptom validity assessment and malingering* (pp. 57-72). New York: Springer.

Slick, D. J., Sherman, E. M., & Iverson, G. L. (1999). Diagnostic criteria for malingered neurocognitive dysfunction: Proposed standards for clinical practice and research. *The Clinical Neuropsychologist*, *13*, 545-561.

Slick, D. J., Tan, J. E., Strauss, E. H., & Hultsch, D. F. (2004). Detecting malingering: A survey of experts' practices. *Archives of Clinical Neuropsychology*, *19*(4), 465-473.

Stevens, A., Friedel, E., Mehren, G., & Merten, T. (2008). Malingering and uncooperativeness in psychiatric and psychological assessment: Prevalence and effects in a German sample of claimants. *Psychiatry Research*, *157*, 191-200.

Strauss, E. H., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary* (3rd ed.). Oxford: Oxford University Press.

Strutt, A. M., Hill, S. W., Scott, B. M., Uber-Zak, L., & Fogel, T. G. (2011). Motivation, psychopathology, locus of control, and quality of life in women with epileptic and nonepileptic seizures. *Epilepsy & Behavior*, *22*, 279-284.

Suhr, J. A., & Barrash, J. (2007). Performance on standard attention, memory, and psychomotor speed tasks as indicators of malingering. In G. J. Larrabee (Ed.), *Assessment of malingered neuropsychological deficits* (pp. 131-170). Oxford: Oxford University Press.

Sullivan, K., Lange, R. T., & Dawes, S. (2006). Methods of detecting malingering and estimated symptom exaggeration base rates in Australia. *Journal of Forensic*

Neuropsychology, *4*, 49-70.

Tan, J. E., Slick, D. J., Strauss, E., & Hultsch, D. F. (2002). How'd they do it? Malingering strategies on symptom validity tests. *The Clinical Neuropsychologist*, *16*, 495-505.

Tombaugh, T. (1996). *Test of Memory Malingering (TOMM).* North Tonawanda, NY: Multi-Health Systems.

Tombaugh, T. N. (1997). The Test of Memory Malingering (TOMM): Normative data from cognitively intact and cognitively impaired individuals.*Psychological Assessment*, *9,* 260-268.

Tsanadis, J., Montoya, E., Hanks, R.A., Millis, S.R., Fichtenberg, N.L., & Axelrod, B.N. (2008). Brain injury severity, litigation status, and self-report of postconcussive symptoms. *The Clinical Neuropsychologist*, *22*, 1080–1092.

Vallabhajosula, B., & Van Gorp, W. G. (2001). Post-Daubert admissibility of scientific evidence on malingering of cognitive deficits. *Journal of the American Academy of Psychiatry and the Law Online, 29,* 207-215.

Vickery, C. D., Berry, D. T. R., Dearth, C. S., Vagnini, V. L., Baser, R. E., Cragar, D. E., & Orey, S. A. (2004). Head injury and the ability to feign neuropsychological deficits. *Archives of Clinical Neuropsychology, 19*, 37–48.

Vickery, C. D., Berry, D. T., 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*, 45-73.

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*, 297-313.

Waddell, G., McCulloch, J. A., Kummel, E. D., & Venner, R. M. (1980). Nonorganic physical signs in low-back pain. *Spine*, *5*, 117-125.

Waddell, G., Pilowsky, I., & Bond, M. R. (1989). Clinical assessment and interpretation of abnormal illness behaviour in low back pain. *Pain, 39*, 41-53.

Waddell, G., Somerville, D., Henderson, I., Newton, M. (1992). Objective clinical-evaluation of physical impairment in chronic low-back-pain. *Spine, 17*, 617-628.

Warrington, E. K. (1984). *Recognition Memory Test*. California: Western Psychological Services.

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*, 114-125.

Williamson, D. J., Drane, D. L., Stroup, E. S., Miller, J. W., & Holmes, M. D. (2003). Most patients with psychogenic nonepileptic seizures do not exert valid effort on neurocognitive testing. *Epilepsia*, *44*(suppl 9), 13.

Williamson, D. J., Holsman, M., Chaytor, N., Miller, J. W., & Drane, D. L. (2012). Abuse, not financial incentive, predicts non-credible cognitive performance in patients with psychogenic non-epileptic seizures. *The Clinical Neuropsychologist*, *26*, 588-598.

Wolfe, P. L., Millis, S. R., Hanks, R., Fichtenberg, N., Larrabee, G. J., & Sweet, J. J. (2010). Effort indicators within the California Verbal Learning Test-II (CVLT-II). *The Clinical Neuropsychologist*, *24*, 153-168.

Young, J. C., Sawyer, R. J., Roper, B. L., & Baughman, B. C. (2012). Expansion and re-examination of Digit Span effort indices on the WAIS-IV. *The Clinical Neuropsychologist*, *26*, 147-159.

Youngjohn, J. R., Davis, D., & Wolf, I. (1997). Head injury and the MMPI-2: Paradoxical severity effects and the influence of litigation. *Psychological Assessment*, *9*, 177-184.

Zimmerman, J., & Grosz, H. J. (1966). “Visual” performance of a functionally blind person. *Behaviour Research and Therapy*, *4*(1), 119-134.

Ziv, I., Djaldetti, R., Zoldan, Y., Avraham, M., & Melamed, E. (1998). Diagnosis of "non-organic" limb paresis by a novel objective motor assessment: the quantitative Hoover's test. *Journal of Neurology, 245*, 797-802.

CHAPTER 3: VARIABLES INFLUENCING EFFORT TEST PERFORMANCE

To date there has been little formal investigation of the variables that are predictive of effort test failure (ETF). The effort research literature has largely focussed on the phenomenon of malingering, a condition which represents only a subset of those displaying ETF during cognitive testing but it has been suggested that there are a range of non-malingering variables that might influence performance on effort testing (Barker, Horner, & Bachman, 2010; Iverson, 2010; Slick & Sherman, 2013). This chapter comprises a review of the literature on predictors of ETF and a critical examination of the range of known predictors of ETF including secondary gain variables, organic and injury-related variables, non-organic psychological variables, and demographic variables. The objectives of the empirical papers included in the current thesis are described and the specific aims of each of the three empirical studies and the systemtic review are detailed.

Compensation-Seeking and Effort

In 1995 Rohling, Binder, and Langhinrichsen-Rohling published a meta-analysis of the results of 32 studies covering 3,802 pain patients and 3,849 controls. That review provided compelling evidence that compensation seeking/receiving was related to increased pain reports and reduced treatment efficacy in pain patients. In the following year, the same team of researchers (Binder & Rohling, 1996) again employed meta-analytic methodology to investigate the role of financial incentives to maintain symptoms and disability after TBI. Across 17 studies and 2,353 participants, greater neuropsychological abnormality and disability was evident in brain injured patients that possessed financial incentives to be unwell. Financial compensation was found to have a moderate effect size on a range of outcome variables including symptom occurrence and duration.

Those studies were among the first to demonstrate that symptom-reporting appeared to be at least partially under the contingency control of external financial and secondary gains. In 2006, Bianchini, Curtis and Greve reported a dose-response relationship between financial incentives and performance on effort tests. In a sample of 332 TBI patients, divided into three incentive groups – no incentive, low incentive under State law, high incentive under Federal law – effort (as determined by performance on five effort indicators) was found to be dependent on the level of financial incentive.

Other studies have subsequently shown that compensation-seeking participants perform more poorly than non-compensation-seeking participants on a variety of effort measures, including studies employing the PDRT (Binder, 1993), the TOMM (Rees, Tombaugh, Gansler, & Moczynski, 1998), RDS (Meyers & Volbrecht, 1998), the Memory Assessment Scales (Ross, Krukowski, Putnam, & Adams, 2003), the VSVT (Doss, Chelune & Naugle, 1999; Grote et al., 2006), the WMS-III Rarely Missed Index (Lange, Sullivan, & Anderson, 2005), the WMT (Flaro, Green, & Robertson, 2007), a digit symbol recognition trial (Kim et al., 2010a) and the WRMT (Kim et al., 2010b).

Flaro, Green and Robertson (2007) sought to further examine the role of external incentives in influencing ETF. Employing the WMT, ETF was contrasted in two clinical groups: 774 adults with mixed severity TBI seeking worker's compensation and 118 low-cognitive-function adults undertaking parenting fitness assessments in the process of seeking custody of their children. WMT failure was 23 times higher in those with mTBI (n = 577) than in those low-cognitive-functioning adults seeking custody of their children and the authors concluded that the study emphasised the importance of external incentive in determining effort test performance.

In contrast to the literature reviewed above, a number of studies have found no or only weak associations between effort and compensation-seeking. Ross, Putnam, and Adams (2006) examined the relationship between compensation-seeking status and effort in a sample of 369 TBI patients referred for neuropsychological assessment. The percentage of patients displaying ETF did not vary as a function of compensation-seeking status. Calculating an effect size from their published data, the difference between the compensation-seeking and non-compensation-seeking groups was trivial (Cohen's d = 0.05). The authors suggested that the failure to find an association between ETF and compensation-seeking was potentially a product of dichotomizing the compensation-seeking variable, however, dichotomizing that variable has been the standard approach used in the literature and appears unlikely to be an adequate explanation for findings. Formal power analysis employing PASS12 (Hintze, 2013) indicates that power was adequate to detect a medium-sized effect but not a small effect at $\alpha = .05$. Only one, relatively uncommonly employed measure of effort (Recognition Memory Test; Warrington, 1984) was employed and the percentage of participants failing that effort measure was low (20%). The finding could indicate that insensitivity to suboptimal effort may have been problematic in that study and may have biased results.

Stulemeijer, Andriessen, Brauer, Vos and van der Werf (2007) conducted a retrospective study of 110 mTBI patients from a prospective cohort (N = 618) and found that litigation status was not associated with performance on the ASTM (p = 0.89). The sample size in that study was smaller than that reported in Ross, Putnam and Adams (2006) but total cases failing the ASTM was closer to typical ETF findings at 27% and power to detect a medium effect at $\alpha = .05$ was adequate (power = .88). The ASTM has been shown to be more sensitive to suboptimal effort than most effort measures and equally sensitive to the WMT (Miller, 2010), suggesting that low sensitivity is not likely to have been an explanation of the null findings.

Similarly, Fox (2011) investigated the relationship between compensation-seeking and ETF using the WMT and the CTAM in a diverse sample of neurological referrals (N = 220). No relationship between financial incentive and ETF on both effort measures was reported ($r_{pb} = -.04$ for the WMT; $r_{pb} = -.07$ for the CTAM). Failure rates were 34% for the WMT and 35% for the CTAM, suggesting that sensitivity was not problematic in that study and the large sample suggests that power was adequate to detect a difference. Sample recruitment procedures were poorly documented which makes it difficult to determine whether the sample was representative or the findings generalizable.

Finally, Williamson, Holsman, Chaytor, Miller, and Drane (2012) examined effort on the WMT in a sample of participants (N = 103) displaying PNES. Compensation-seeking patients did not fail the WMT at a greater rate than non-compensation-seeking patients (χ^2 = .1, p = .75). Calculating Cohen's *d* from the published data, the difference between the groups was trivial (d = .07). A formal analysis of power indicates that power was adequate to detect a medium effect at α = .05 (power = .86). Of the total sample 35% failed WMT, further suggesting that inadequate sensitivity did not account for the null findings. Although there were no substantial methodological flaws in that study, it is not clear that findings from the sample would generalize to effort testing in the TBI population.

Moore and Donders (2004) investigated predictors of ETF in a sample of 132 TBI patients using an alternate statistical methodology – logistic regression. When compensation-seeking and a premorbid psychiatric history were included in a logistic regression equation compensation-seeking was a significant predictor of invalid effort (OR = 3.48, $p < .05$), however, a premorbid psychiatric history was the stronger predictor (OR = 3.72, $p < 0.05$). Donders and Boonstra (2007) also employed logistic regression with the effort test results from 87 TBI patients to examine predictors (age, time assessed post-injury, gender, premorbid psychiatric history, personal abuse history and experiencing prolonged coma) of ETF. When included in the analysis along with other variables, compensation-seeking did not predict suboptimal effort ($p > .10$), however, a psychiatric history and the *absence* of coma were predictive of ETF. The study employed a relatively large number of variables relative to sample size, raising the possibility of over-fitting of the data. Nevertheless, when these findings are seen alongside the other somewhat sample-dependent findings described above and the previous logistic regression study by Moore and Donders (2004), they raise the possibility that compensation-seeking might gain its predictive association with ETF from some other (potentially more important) covariate.

The literature in this field indicates that although financial compensation-seeking is likely to influence effort during cognitive testing, compensation-seeking does not wholly account for the variance seen in effort. When compensation-seeking and/or litigation are included in predictive studies alongside other predictive variables, its power to predict ETF appears to decrease. There is a need to investigate the predictive power of compensation-seeking status using multivariable logistic regression analysis and employing a larger sample of TBI participants.

Variables associated with ETF False Positive findings

Mental retardation

Although effort tests typically have a high degree of specificity, specificity rates are always below 100% and the risk of false positive findings (i.e., findings of suboptimal effort when effort was normal) is present. Repeated studies have shown that although performance on effort measures is relatively independent of performance on tests of cognitive functioning, effort test performance may not be wholly immune to the effects of very low levels of cognitive functioning. An early but small (*N* = 16) study (Goldberg & Miller, 1986) showed that 38% of participants with mental retardation (MR) failed the FIT. Subsequent studies employing a range of effort measures support the finding that ETF is associated with MR and that most tests lack specificity in this population (Boone et al., 2002; Hurley & Deal, 2006; Iverson & Franzen, 1994; Lu, Boone, Cozolino, & Mitchell, 2003; Marshall & Happe, 2007; Ray, 2012; Victor, Boone, & De La Rossa-Trujillo, 2005). Given that the majority of studies have shown that false positives are unacceptably high in the MR population undertaking effort testing, it is important to exclude patients with MR from studies of effort in other clinical populations.

Dementia

Findings in respect to dementia and ETF have consistently shown that few effort measures maintain adequate specificity when employed with people affected by dementia (Dean, Victor, Boone, Philpott, & Hess, 2009). Dean and colleagues noted that free-standing forced-choice measures such as the WMT and TOMM were associated with high rates of false positives, as were the Digit Memory Test, the VSVT, and the MSVT. Non-forced-choice freestanding measures including the Dot Counting Test and the FIT also had unacceptably low specificities and embedded measures also demonstrated low specificity. Since that 2009 review, Kiewel, Wisdom, Bradshaw, Pastorek, and Srutt (2012) further examined the use of Digit Span indices in a sample of patients with Alzheimer's disease (*N* = 142) and reported that only the Vocabulary–Digit Span index demonstrated adequate specificity through the range of dementia severities. Bortnik, Horner, and Bachman (2013) examined four other commonly used effort measures including the TOMM and the FIT, and reported that none had adequate sensitivities or specificities as to be useful in cases of dementia. Those results highlight the importance of excluding patients with known or suspected dementia in studies of ETF.

Traumatic Brain Injury Severity and Effort

TBI severity has been examined as a predictor of suboptimal effort. As noted in Chapter 2, the literature has compellingly demonstrated a paradoxical finding in respect of TBI severity, such that TBI patients with mTBI tend to report greater symptom frequency and severity than those with more severe TBIs (Paniak et al., 2002; Tsanadis et al., 2008). The effort literature generally evidences a parallel phenomenon in respect of the relationship between TBI severity and test taker effort.

Green, Iverson, and Allen (1999) contrasted the performance of a sample of compensation-seeking mTBI patients (n = 234) with a sample of compensation-seeking patients with MS-TBI (n = 64) on the WMT and CARB. An inverse relationship between injury severity and effort was reported: participants with less severe injuries performed significantly worse on the effort measures than those with more severe injuries. Effect sizes for injury severity on effort were not reported but calculating effect sizes from the author's published data shows that Cohen's d ranged from 0.43 for immediate recognition to 0.49 for delayed recognition. The effect size of effort on the CARB was large (Cohen's d = 0.63).

Green and colleagues (Green, Rohling, Lees-Haley, & Allen, 2001) subsequently investigated the comparative impact of suboptimal effort and moderate-to-severe TBI on cognitive test performance. Of a compensation-seeking mTBI group of 276 patients, 34% failed the WMT as opposed to a failure rate of 18% in 90 TBI patients with MS-TBI. Calculations from the published data reveal the effect size of mTBI on effort to be large (r = .61).

The finding of an inverse relationship between TBI severity and effort has been replicated many times. Moss, Jones, Fokias, & Quinn (2003) found that while TBI severity was negatively correlated with a range of neuropsychological test outcomes in a compensation-seeking sample passing the TOMM (n = 54) (Range r_s: -0.16, -0.39, p's <.05), the relationship was not evident in those failing the TOMM (n = 24) (Range r_s: 0.02, .28, p's >.10).

Carone (2008) found that children with moderate-to-severe brain damage outperformed adults with mTBI on the MSVT. Specific statistics were not reported. Calculating from their published data, the difference in MSVT fail rates between the groups was significant (χ^2 (1, N = 105) = 4.59, p < .05, d = .43). The adult mTBI sample had proportionately more compensation-seekers than the child group and that sampling bias may have confounded findings.

Donders and Boonstra (2007) found that having *no* history of sustained prolonged coma following TBI (*N* = 87) increased the risk of suboptimal effort (on the CVLT-II or WMT) four-fold. Compensation-seeking status was controlled during logistic regression analyses but precise statistics were not reported. Finally, Green (2011; 2013) and others (Guidotti Breting & Sweet, 2013; Green & Iverson, 2011; Greiffenstein & Baker, 2006; Tsanadis et al., 2008) have reported that rate of failure on SVTs was greater for those compensation-seeking patients with a history of mTBI than those with more severe TBIs, although the Tsanadis et al. (2008) study did not control well for compensation-seeking status and findings may have been confounded in that case.

Given that the inverse relationship between TBI severity and ETF appears to be a stable finding in the effort literature, TBI severity may be a useful indicator with which to test any multivariable predictive model of ETF. Replicating the findings would serve to strengthen the validity of the resulting predictive model. To date, only Donders and Boonstra (2007) have attempted to include both TBI severity and compensation-seeking in a regression study and their findings suggested that the inverse severity-ETF relationship was evident even after controlling for compensation-seeking. The study included a relatively small sample relative to the number of variables considered and there is the risk that the data was inadvertently over-fitted. The finding therefore deserves replication with a larger sample of both compensation-seeking and non-compensation-seeking participants.

Education and Effort

In one of the few studies to specifically investigate the impact of education on effort, Tombaugh (1997) stated that education did not account for a significant amount of variance seen in TOMM scores in a sample of healthy community volunteers (*N* = 405), reporting that age and education (combined) accounted for less than 2% of the variance on Trial 2 and Retention. The methodology of multiple regression was employed in possible contravention of the assumption of normality of distribution of the dependent variable (TOMM scores) and Tombaugh acknowledged that the restricted range of TOMM scores may have served to reduce the contribution of age and education. Tombaugh (1997) conducted a second study using the same statistical methodology and examining TOMM results in a non-compensation-seeking mixed sample of neurological patients (N = 138). Age and education (combined) accounted for less than 11% of the variance on Trial 2 and Retention. Tombaugh concluded that education did not significantly affect TOMM performance. Similarly, Rees et al. (1998) found that in a very small sample of TBI patients (N = 18), education contributed 12% of the variance in TOMM (Trial 2) scores. It was concluded that education did not significantly impact on TOMM performance. Arguably, in both of those studies methodological problems may have inadvertently reduced the statistical effects of education on ETF and the conclusions of Tombaugh (1997) and Rees and colleagues (1998) may not have been warranted.

When specific findings in relation to education are reported they tend to indicate that lower education is associated with ETF (e.g., Babikian, Boone, Lu & Arnold, 2006; Back et al., 1996; Baker, Donders, & Thompson, 2000; Davis, McHugh, Bagley, Axelrod & Hanks, 2011; Duff et al., 2011; Greve, Etherton, Ord, Bianchini, & Curtis, 2009; Mahdavi, Mokari, & Amiri, 2011; Salazar, Lu, Wen, & Boone, 2007; Stevens, Friedel, Mehren, & Merten, 2008; Strutt, Scott, Lozano, Tieu, & Peery, 2012; Stulemeijer, Andriessen, Brauer, Vos, & Van Der Werf, 2007). However, the inverse relationship between education and ETF has not been a consistent finding, with other researchers reporting that education is unrelated to ETF (Armistead-Jehle, 2010; Gunner, Miele, Lynch, & McCaffrey, 2012; Rees et al., 1998; Tan, Slick, Strauss, & Hultsch, 2002; Tombaugh, 1997; Young, Caron, Baughman, & Sawyer, 2012) or only weakly associated (Powell, Locke, Smigielski, & McCrea, 2011).

In those studies that have reported no statistically significant effect of education on effort there has been a strong trend for education to be lower in the group failing SVTs, albeit to a non-statistically-significant level (e.g., Armistead-Jehle, 2010; Axelrod, Fichtenberg, Millis, & Wetheimer, 2006; Ord, Boettcher, Greve, & Bianchini, 2010; Gunner et al., 2012; Powell et al., 2011; Rees et al., 1998; Young et al., 2012). In fact, in no instance has education been higher in the sample displaying ETF than in the sample displaying valid effort.

Additionally, studies examining education have tended to employ only one measure of effort and relatively small sample sizes. Power analysis utilising G*Power (Buchner, 1997) revealed that when undertaking a t-test of mean differences, a total sample of N = 82 is required at power of .80 and an α of .05 to detect a medium effect (d = .30). Only two of the above studies (Ord et al., 2010; Young et al., 2012) employed a sufficient sample size to detect a medium effect and none had a sufficient sample size to detect a small effect. It may be that in cases where education is invariably lower in samples affording ETF than samples affording valid effort, insufficient sample size accounts for the failure to detect a statistically significant difference. Examining the relationship between effort and education in a larger sample size appears to be indicated.

Some studies that have employed a statistical methodology of correlating effort test performance with education have shown no significant correlation between the two variables (Arnold et al., 2005; Constantinou & McCaffrey, 2003; Curtis, Greve, Bianchini, & Brennan, 2006; Larrabee, 2003; Meyers & Volbrecht, 1998; Meyers, Volbrecht, Axelrod, & Reinsch-Boothby, 2011; Teichner & Wagner, 2004). Others such as that conducted by Babikian, Boone, Lu, and Arnold (2006) have revealed positive correlations between Digit Span effort scores and education. Boone, Lu, and Wen (2005) found correlations between education and Rey Auditory Verbal Learning Test (RAVLT; Lezak, 1995) effort measures and Constantinou & McCaffrey (2003) found positive correlations between FIT scores and education in a sample of 128 healthy children but not between TOMM and education. Nelson et al. (2003) found evidence of positive correlations between education and FIT as well as two other embedded measures and Duff et al. (2011) found significant correlations between education and effort indices of the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS; Randolph, 1998) such that lower education was associated with ETF.

Correlation studies typically achieve most valid results when the variables of interest have wide variance and correlations can appear spuriously low in cases where variability is restricted (Bland & Altman, 2011). Effort test scores tend to have a restricted range (Tombaugh, 1997) and therefore the statistical technique of correlation may fail to detect a relationship that actually exists because of the restricted range of scores seen on effort measures. Accordingly, it is concluded that the matter requires further examination using an alternate statistical methodology.

The results of logistic regression (a statistical technique that avoids the challenges faced in meeting the assumptions of both correlation and multiple regression), employing education as a predictor of suboptimal effort have been reported only once (Baker, Donders, & Thompson, 2000). Baker and colleagues found that low levels of education were a risk factor for false-positive identification of suboptimal effort in a TBI sample (N = 134) using embedded measures from the CVLT (Wald χ^2 (1, N = 134) = 8.64, $p < .01$).

Stable findings have not been found in respect to the relationship between education and ETF, however, the results of previous research suggest that a relationship might exist and that the relationship may be inverse – low levels of education may be predictive of ETF. Previous studies have been affected by the use of small samples and the use of data analytical techniques potentially in contravention of statistical assumptions; this may have inadvertently led to null findings and type II errors. It may be the case that more educated participants afford better effort because they have less motivation to malinger neurocognitive dysfunction. Alternatively, it may be the case that more educated participants are evading detection of suboptimal effort (Rapport, Farchione, Coleman, & Axelrod, 1998). If so, the use of multiple effort measures may be a useful strategy to minimize the risk of non-detection of suboptimal effort (Nelson et al., 2003). Using a large sample and powerful multivariable analytic techniques may improve detection of any effect of education on ETF that has previously not been detected.

Psychological Variables and Effort

Affective disturbance and effort.

The results of research into the role that affective variables play in effort testing have conflicted. Early studies, largely conducted in the context of validating effort measures, found that affective disturbance including depressed and anxious affect had no appreciable impact on effort test performance (e.g., Ashendorf, Constantinou, & McCaffrey, 2004; Rees, Tombaugh, & Boulay, 2001). In a narrative review of a number of embedded and standalone effort measures and affective disturbance Goldberg, Back-Madruga and Boone (2007) concluded, "Data from these studies were consistent in showing no impact of depression, including increasing severity of depression and depression subtypes, on 12 separate effort indicators" (p. 305). Appraisal of the inclusion criteria of Goldberg and colleagues indicates that only one WMT study was included in their review of ETF and depression. In fact, the effort literature in which the WMT has been employed includes a number of reports of a significant impact of affective disturbance on effort (Bauer, 2007; Brooks, Johnson-Greene, Lattie, & Ference, 2012; Gorissen, Sanz, & Schmand, 2005; Green 2009; Green, Rohling, Lees-Haley, & Allen, 2001; Lange, Pancholi, Bhagwat, Anderson-Barnes, & French, 2012; Rohling, Green, Allen, & Iverson, 2002). However, similar findings of an effect of negative mood on effort have been noted using other effort measures since publication of the narrative review by Goldberg et al. (2007) including the TOMM (Bauer, 2007; Brooks et al., 2012; Tsanadis et al., 2008), RDS (Brooks et al., 2012); the ASTM (Heintz et al., 2013; Stulemeijer, Andriessen, Brauer, Vos, & Van der Werf, 2007), the VSVT (Silver, 2012), the Recognition Memory Test (Millis, 1992; Tsanadis et al., 2008) and the MSVT (Stevens et al., 2008).

The majority of studies conducted to date have comprised examination of the relationship between depressed affect and effort and only very few studies have specifically examined the relationship between anxious affect and effort (Ashendorf, Constantinou, & McCaffrey, 2004; Lange et al., 2012; Locke, Smigielski, Powell, & Stevens, 2008; O'Bryant, Finlay, & O'Jile, 2007, Sumanti, Boone, Savodnik, & Gorsuch, 2006). Ashendorf, Constantinou, & McCaffrey (2004) reported finding no effect of depression or anxiety on TOMM performance in a highly selected sample of 197 community dwelling older adults, in whom participants with major psychological conditions had been previously screened *out*, as had participants receiving treatment with antidepressants or therapy. A small subsample (n = 23) of the group reported some anxiety symptomatology on the State Trait Anxiety Inventory (STAI; Spielberger, 1983; State Anxiety M = 48.26, SD = 2.63) and although that group had lower TOMM scores than five other comparison groups, the differences between the groups were statistically non-significant. Calculating from their published data, comparing the State Anxious group with the Non-State Anxious group there was a clinically significant and medium-sized effect of anxiety on TOMM Trial 2 (d = 0.53). Comparing the Trait Anxious group with the Non-Trait Anxious group, a small to medium effect of trait anxiety on effort was evident (d = .33). The highly selected sample employed in that study (i.e., participants were screened out for significant psychological disturbance) was unusual given the purposes of the study and the reasons for screening out those with psychological disturbance was not explained. That exclusion criterion potentially introduced substantial sampling bias and likely diluted any effects of psychological disturbance on effort.

O'Bryant, Finlay, & O'Jile, (2007) also examined the relationship between TOMM scores and the STAI in a diverse sample of neurological patients (N = 67) referred for neuropsychological assessment. It was found that the STAI State and Trait measures were significantly correlated with scores on Trial 1 of the TOMM (r = -.25 and r = -.37, respectively, *p's* < .05) but not Trial 2 or the Retention trial (*r's* < .25, *p's* > .05). Significantly, mean scores on both TOMM Trial 2 and the TOMM Retention trial were very low: M = 46.58 (SD = 7.44), and M = 46.11 (SD = 8.71) respectively. Both were below recent cut-scores of <47 for the TOMM (Greve, Bianchini, & Doane, 2006). It is evident that the average level of effort was very low in that depressed and anxious sample.

Sumanti and colleagues (2006) examined Personality Assessment Inventory (PAI; Morey, 1991) performance and effort test (FIT and Dot Counting (Boone et al., 2002) performances in a sample of compensation-seeking consecutive psychiatric referrals (N = 233). Significant correlations between the PAI Anxiety scale and both the FIT and Dot Counting tests (r = -.23, p < .001 and r = .22, p < .01 respectively) and the Anxiety Related Disorders subscale and both the FIT and Dot Counting tests (r = -.20, p < .01 and r = .17, p < .05 respectively) were reported. Those findings were supported by Lange et al. (2012), who also employed the PAI with a different set of effort measures (WMT and four embedded measures) in a sample of 143 TBI patients (compensation-seeking status not reported). Those displaying ETF were found to score higher on both the PAI Anxiety (d = 0.33) and Anxiety-Related Disorders (d = 0.34) scales than those not displaying ETF.

Finally, Locke et al. (2008) examined TOMM failure in a sample of 87 TBI patients. Anxiety was assessed using the Beck Anxiety Inventory (Beck & Steer, 1990) and results showed that although anxiety scores were higher in the suboptimal effort group than the optimal effort group, the difference was statistically and clinically non-significant (Cohen's d

= -.04). The sample was described as treatment-seeking but compensation-seekers were not screened out and (on the basis of the referral sources including an attorney, a worker's compensation nurse and a probation officer) were likely included in the sample.

At present the literature remains divided on the role that affective disturbance has on effort. Early papers indicated that depressed and anxious affect had no role on effort but recent studies have been more consistently reporting effects of negative affect on effort. Methodological issues (e.g., small comparison groups, highly selected samples) and choice of statistical procedures (e.g., correlation with variables with restricted ranges of data) may have inadvertently biased results from early studies. The relationship that affective symptoms have with ETF independent of compensation-seeking is currently unclear. Further research comprising both a large sample whose compensation-seeking status has been better delineated and more powerful multivariable statistical procedures may be informative.

Psychotic disturbance and effort.

While there remains uncertainty over the impact of affective disturbance on effort testing, the literature has presented fairly consistent findings in respect to the impact of psychotic symptomatology on effort. Gorrisen, Sanz, and Schmand (2005) examined WMT performance in 64 patients diagnosed with schizophrenia; most participants (72%) failed the WMT. Duncan (2005) found that 8% of their sample of 50 participants with psychotic illness failed the TOMM. Dandachi-Fitzgerald, Ponds, Peters, and Merckelbach (2011) examined cognitive underperformance in a mixed psychiatric sample and found in their small sample of psychotic patients (*n* = 8) one quarter failed the ASTM. Other researchers (Donders & Kirkwood, 2013; Goldberg, Back-Madruga & Boone, 2007) have reviewed the literature on psychosis and effort and have concluded that there is a high rate of ETF in that population. Those authors recommend that effort tests be interpreted with caution in individuals affected by psychosis.

A dissenting note comes from Schroeder and Marshall (2011), who examined the performance of a sample of patients with psychotic disorders (*n* = 104) on seven embedded measures of effort. Of their sample, 26% of the patients with psychosis failed one or more effort measures but few (7%) failed two or more measures. Schroeder and Marshall concluded that previous research that relied solely on the WMT (e.g. Gorissen, Sanz and Schmand, 2005), which is known to have relatively high attentional demands (e.g., Batt, Shores, & Chekaluk, 2008), likely accounted for the difference in findings. Further analysis of the role that psychosis plays in effort testing is warranted, employing a range of effort measures.

Pain and effort.

Abnormally slowed recovery from TBI and mTBI has been associated with the presence of comorbid conditions including chronic pain (Iverson, 2005). Pain disorder as a predictor of ETF has come under considerable investigation in the effort literature. Results from those studies are generally consistent and show that in the clinically pain-disordered and compensation-seeking population, neurocognitive test taker effort tends to be compromised (Bianchini, Greve, & Glynn, 2005). Studies have included diverse samples of compensation-seeking chronic pain patients (Etherton, Bianchini, Heinly, & Greve, 2006; Gervais, Rohling, Green, & Ford, 2004; Greve, Ord, Curtis, Bianchini, & Brennan, 2008; Greve, Etherton, Ord, Bianchini, & Curtis, 2009; Greve et al., 2010; Meyers & Diep, 2000; Meyers & Volbrecht, 2003; Suhr, 2003; Suhr & Spickard, 2007) or more selected samples of chronic pain patients including compensation-seeking patients with Fibromyalgia (Brooks, Johnson-Greene, Lattie, & Ference, 2012; Gervais et al., 2001) and compensation-seeking patients with chronic regional pain syndrome Type 1 (Greiffenstein, Gervais, Baker, Artiola, & Smith, 2012). The finding that pain is associated with ETF has not been reported in clinical pain patients who are not compensation-seeking (Gervais et al., 2000; Meyers & Diep, 2000; Meyers &

Volbrecht, 2003). Significantly, additional studies into laboratory-induced pain and effort show that while people with clinical pain disorders tend to display reduced effort, those suffering from laboratory-induced pain do not (Etherton, Bianchini, Ciota, & Greve, 2005; Etherton, Bianchini, Heinly, & Greve, 2006). The implication of this finding is that other variables that are associated with clinical pain (e.g., the chronicity of pain, sleep deprivation) may be contributing substantial variance to effort test findings. Compensation-seeking may be one of those variables but other, as yet unexamined clinical variables, may also be contributing to the relationship.

To date there have been no studies of ETF in cases of TBI with comorbid pain disorder; indeed, in previous research into the relationship between ETF and chronic pain cases of TBI have been excluded from the patient cohorts under study. It remains unclear whether the association that has been identified between chronic pain and slowed recovery from TBI (Iverson, 2005) is due to an association between ETF and pain disorder.

Personal history variables and effort.

The personal history of an individual has been examined as a predictor of suboptimal effort in few studies. Moore and Donders (2004) reported that having a psychiatric history was predictive of suboptimal effort in a sample of 132 rehabilitation referrals. Using logistic regression analyses the authors reported that having a psychiatric history was associated with an almost four-fold (OR = 3.48, $p < .05$) risk of suboptimal effort during cognitive testing. Donders and Boonstra (2007) examined predictors of ETF in a sample of 87 TBI patients. A psychiatric history was associated with significantly increased risk of poor effort (χ^2 (1, N = 87) = 6.76, $p < .01$, OR = 3.74, 90% CI [1.58, 8.84]).

Employing the same TBI sample, Donders and Boonstra (2007) further examined whether having a personal abuse history, including incidents of physical and/or sexual abuse, was associated with ETF. It was found that having an abuse history was significantly predictive of ETF (χ^2 (1, N = 87) = 4.05, p < .05, OR = 3.37, 90% CI [1.20, 9.55]). Williamson et al. (2012) reported that in a sample of 103 patients with PNES, a history of physical, sexual or emotional abuse was associated with ETF (χ^2 = 7.3, p < .01).

The limited evidence provides some indication that personal variables including having a history of psychiatric illness and having a history of physical, sexual or emotional abuse are predictive of suboptimal effort. An abuse history is known to be a strong predictor of developing psychiatric illness in adulthood (Arnow, 2004). It remains unclear to what extent the variables of abuse, current psychiatric illness and historic psychiatric illness individually contribute to reduced test taker effort. Employing a multivariable model of ETF that includes these personal history variables along with other psychological variables (and in particular, affective disturbance, psychotic illness and pain disorder) would allow for some appraisal of the relative contributions of each variable to suboptimal effort.

Cultural Variables and Effort

A substantial cognitive research literature exists that establishes culture-dependent psychometric test performance, such that majority (typically White, English-speaking) cultures tend to outperform both minority cultures and those for whom English is their second language (ESL) (Terrell, Terrell, & Taylor, 1980; Terrell & Terrell, 1983; Steele & Aronson 1995; Chan, 1997; Chan, Schmitt, DeShon, Clause, & Delbridge, 1997; Nabors, Evans & Strickland, 2000; Kennepohl, Shore, Nabors & Hanks, 2004; Walker, Batchelor & Shores, 2009). Given that there is a relationship between culture and performance on neuropsychological and cognitive tests, it is conceivable that a similar relationship between culture and performance on measures of effort exists. This potential relationship has received limited attention.

Vilar-López et al. (2007) examined the ability of three specific cognitive symptom exaggeration measures (VSVT, TOMM, and *b* test) to discriminate between three small

Spanish subsamples – a litigating sample of TBI patients (*n* = 14), a nonlitigating sample of TBI patients (*n* = 12) and a group of analogue malingerers (*n* = 35) . Results showed that in comparison to findings from English-speaking populations, the tests were not effective in discriminating between the two groups of non-litigating and litigating TBI patients. In a follow-up study, Vilar-López, Gomez-Rio, Caracuel-Romero, Elvira, and Perez-Garcia (2008a) reported that in Spanish participants with mTBI, the VSVT and *b* test discriminated a non-compensation-seeking sample (*n* = 30) from a probable malingering sample (*n* = 10) but the FIT lacked discriminability. In a similar study using the same sample Vilar-Lopez and co-workers (2008b) examined the ability of the TOMM and the Dot Counting test to discriminate the groups. The TOMM was able to discriminate the groups but Dot Counting was less effective and large numbers of false positive and false negatives were noted.

Salazar, Lu, Wen, and Boone, (2007) examined the impact of ethnicity and ESL on performance on a range of nine independent embedded effort measures in a diverse non-compensation-seeking sample of neuropsychological referrals (*N* = 168). Caucasians (*n* = 85) were reported to have performed significantly better on Digit Span embedded effort measures than Hispanics (*n* = 32), and Caucasians scored higher than African Americans (*n* = 32) on RAVLT and Rey-Osterrieth embedded measures (specific statistics not reported). Two independent Digit Span effort scores were found to be related to the age at which English was learned (r_s = -.248, p = .001 and r_s = -.29, p = .0001). Years resident in the United States and years educated in the United States were examined as variables of acculturation and were reported to not relate to any effort scores (specific statistics were not, however, reported).

Yang et al. (2012) examined the use of Digit Span-based effort measures in a Taiwanese Chinese TBI sample (*N* = 132). The authors reported that the normative digit span performance differed between the US and Taiwanese standardization samples such that the Chinese-speaking sample out-perform the English-speaking on both digits forward (d = .86) and digits backward (d = .31). The findings suggested that a higher RDS cut-score was needed for the Chinese-speaking population and the Vocabulary minus Digit Span score (VDS; Curtis, Greve, & Bianchini, 2009) may not be effective in discriminating those affording good effort from those affording suboptimal effort in the Chinese-speaking population. Benuto and Leany (2013) raised similar but inverse concerns about the validity of use of the RDS index with Hispanic patients. Notably, the specificity of the measure had not been examined in Hispanics but the authors identified that Hispanics typically performed more poorly on Digit Span than Caucasians and indicated that, in their view, current research did not support the use of RDS in Spanish speaking populations (Benuto & Leano, 2013).

Finally, Burton, Vilar-López, and Puente (2012) examined performance on three effort measures (TOMM, FIT, Dot Counting) in Spanish-speaking US citizens, including 29 control participants, 28 capital murder forensic participants and 25 personal injury/compensation-seeking forensic litigation participants. Somewhat unexpectedly, the tests did not discriminate the healthy control participants from those with a high motivation to give suboptimal effort (those charged with capital murder) (p*'s* > .05) but the FIT and TOMM did discriminate between the compensation-seeking forensic group and the control sample (p*'s* < .05). The findings suggested that the tests may not be operating in the Spanish-speaking population in the same manner that they do in the English-speaking population.

Although definitive conclusions cannot yet be drawn, available literature does suggest that cultural variables may impact on effort test performance such that those of minority cultures may underperform relative to the majority culture on effort measures. The extent to which English as a second language contributes to suboptimal effort as opposed to other acculturation variables remains unclear. Further research is required to clarify the nature of the relationships between ethnicity and acculturation and effort.

Conclusion

ETF is known to be associated with a number of variables. The availability of secondary gains, in the form of financial gain, appears to be associated with ETF, although research findings in relation to financial benefits are less definitive when the effects of financial benefits have been examined in the context of other predictors. TBI severity is consistently, inversely related to ETF such that those with mTBI tend to perform more poorly on effort measures than those with severe TBIs. Research shows that very severe cognitive impairment in the form of dementia and mental retardation likely impacts on effort test performance. Although the literature is not entirely consistent in respect to the relationship between education and effort, research findings trend toward showing that lower education is associated with reduced effort test performance. Psychiatric variables appear to be associated with effort test performance. Psychosis has fairly consistently been found to significantly affect performance on tests of effort, with psychotic individuals scoring more poorly than healthy controls. The literature on the influence of negative affect on effort remains divided with some, predominantly early researchers, finding no effect of depressed or anxious mood on effort, while other more recent studies have shown an effect of negative affect on effort. The presence of chronic pain has been consistently associated with ETF but the effects of that variable have not yet been examined in a TBI sample. Cultural variables including English as a second language and acculturation have been shown to impact on effort tests such that samples with low levels of acculturation and non-English speakers typically perform more poorly on effort measures than majority culture, English-speaking samples. This is an emerging literature and further research is needed to clarify the relationship between effort and acculturation variables.

There is a need to clarify and identify statistical predictors of suboptimal effort. At the outset, understanding what statistical correlates of suboptimal effort exist will help to direct the course of future research by pointing to possible causal variables that underpin suboptimal effort and will promote the development of experimental studies that attempt to manipulate cognitive test effort and confirm hypothesised causal relationships between variables and effort.

The majority of studies conducted to date have examined predictors of suboptimal effort using group comparison methodologies with relatively small samples. More sophisticated multivariable statistical methodologies have seldom been utilised in this literature and in some instances, correlation and multiple regression analytic techniques have been employed potentially in contravention of the assumptions of those statistical approaches. Multivariable techniques offer the opportunity to evaluate a variety of predictors concurrently and to examine the relative predictive value of specific variables, while avoiding the statistical pitfalls of traditional univariate analytical techniques.

References

Armistead-Jehle, P. (2010). Symptom validity test performance in US veterans referred for evaluation of mild TBI. *Applied Neuropsychology, 17,* 52-59.

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,* 105-120.

Arnow, B. A. (2004). Relationships between childhood maltreatment, adult health and psychiatric outcomes, and medical utilization. *The Journal of Clinical Psychiatry*, *65*, 10.

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,* 125-130.

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*, 513-523.

Babikian, T., Boone, K. B., Lu, P., & Arnold, G. (2006). Sensitivity and specificity of various digit span scores in the detection of suspect effort. *The Clinical Neuropsychologist*, *20,* 145-159.

Back, C., Boone, K. B., Edwards, C., Parks, C., Burgoyne, K., & Silver, B. (1996). The performance of schizophrenics on three cognitive tests of malingering, Rey 15-Item memory test, Rey Dot Counting, and Hiscock Forced-Choice method. *Assessment*, *3*, 449-457.

Baker, R., Donders, J., & Thompson, E. (2000). Assessment of incomplete effort with the California Verbal Learning Test. *Applied Neuropsychology*, *7,* 111-114.

Barker, M. D., Horner, M. D., & Bachman, D. L. (2010). Embedded indices of effort in the

Repeatable Battery for the Assessment of Neuropsychological Status (RBANS) in a geriatric sample. *The Clinical Neuropsychologist*, *24*, 1064-1077.

Batt, K., Shores, E. A., & Chekaluk, E. (2008). The effect of distraction on the Word Memory Test and Test of Memory Malingering performance in patients with severe brain injury. Journal *of the International Neuropsychological Society, 14*, 1074–1080.

Bauer, L. (2007). Is there a relationship between personality pathology and performance on symptom validity tests? Unpublished doctoral dissertation, State University of New York, New York, USA.

Beck, A., & Steer, R. A. (1990). *Beck Anxiety Inventory,* San Antonio, TX: The Psychological Corporation.

Benuto, L. T., & Leany, B. D. (2013). Assessing effort and malingering with the Hispanic client. In L. T. Benuto (Ed.). *Guide to Psychological Assessment with Hispanics,* (pp. 99-112). New York: Springer.

Bianchini, K. J., Curtis, K. L., & Greve, K. W. (2006). Compensation and malingering in traumatic brain injury: a dose-response relationship?. *The Clinical Neuropsychologist*, *20*, 831-847.

Bianchini, K. J., Greve, K. W., & Glynn, G. (2005). On the diagnosis of malingered pain-related disability: lessons from cognitive malingering research. *The Spine Journal*, *5,* 404.

Binder, L. M. (1993). Assessment of malingering after mild head trauma with the Portland Digit Recognition Test. *Journal of Clinical and Experimental Neuropsychology*, *15,* 170-182.

Binder, L. M., & Rohling, M. L. (1996). Money matters: a meta-analytic review of the effects of financial incentives on recovery after closed-head injury. *American Journal of Psychiatry*, *153,* 7-10.

Bland, J. M., & Altman, D. G. (2011). Correlation in restricted ranges of data. *BMJ*, 342: d556.

Boone, K. B., Lu, P., & Wen, J. (2005). Comparison of various RAVLT scores in the detection of noncredible memory performance. *Archives of Clinical Neuropsychology*, *20*, 301-319.

Boone, K. B., Lu, P., Back, C., King, C., Lee, A., Philpott, L., Shamieh, E., & Warner-Chacon, K. (2002). Sensitivity and specificity of the Rey Dot Counting Test in patients with suspect effort and various clinical samples. *Archives of Clinical Neuropsychology, 17*, 625–642.

Bortnik, K. E., Horner, M. D., & Bachman, D. L. (2013). Performance on standard indexes of effort among patients with dementia. *Applied Neuropsychology*, *20,* 233-242.

Brooks, L., Johnson-Green, D., Lattie, E., & Ference, T. (2012). The relationship between performances on neuropsychological symptom validity testing and the MCMI-III in patients with fibromyalgia. *The Clinical Neuropsychologist*, *26,* 816-831.

Buchner, A., Erdfelder, E., & Faul, F. (1997). How to Use G*Power [WWW document]. URL http://www.psycho.uniduesseldorf.de/aap/projects/ gpower/how_to_use_gpower.html

Burton, V., Vilar-López, R., & Puente, A. E. (2012). Measuring Effort in Neuropsychological Evaluations of Forensic Cases of Spanish Speakers. *Archives of Clinical Neuropsychology*, *27*, 262-267.

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*, 960-971.

Chan, D. (1997). Racial subgroup differences in predictive validity perceptions on personality and cognitive ability tests. *Journal of Applied Psychology, 82,* 311-320.

Chan, D., Schmitt, N., DeShon, R.P., Clause, S.C., & Delbridge, K. (1997). Reactions to cognitive ability tests: the relationships between race, test performance, face validity perceptions, and test-taking motivation. *Journal of Applied Psychology, 82,* 300-310.

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

Curtis, K. L., Greve, K. W., & Bianchini, K. J. (2009). The Wechsler Adult Intelligence Scale-III and malingering in traumatic brain injury: Classification accuracy in known groups. *Assessment, 16,* 401–414.

Curtis, K. L., Greve, K. W., Bianchini, K. J., & Brennan, A. (2006). California Verbal Learning Test Indicators of Malingered Neurocognitive Dysfunction Sensitivity and Specificity in Traumatic Brain Injury. *Assessment, 13,* 46-61.

Dandachi-Fitzgerald, B., Ponds, R. W., Peters, M. J., & Merckelbach, H. (2011). Cognitive underperformance and symptom over-reporting in a mixed psychiatric sample. *The Clinical Neuropsychologist, 25,* 812-828.

Davis, J. J., McHugh, T. S., Bagley, A. D., Axelrod, B. N., & Hanks, R. A. (2011). Cross-Validation of Picture Completion Effort Indices in Personal Injury Litigants and Disability Claimants. *Archives of Clinical Neuropsychology, 26,* 768-773.

Dean, A. C., Victor, T. L., Boone, K. B., & Arnold, G. (2008). The relationship of IQ to effort test performance. *The Clinical Neuropsychologist, 22,* 705-722.

Dean, A. C., Victor, T. L., Boone, K. B., Philpott, L. M., & Hess, R. A. (2009). Dementia and effort test performance. *The Clinical Neuropsychologist, 23,* 133-152.

Donders, J., & Boonstra, T. (2007). Correlates of invalid neuropsychological test performance after traumatic brain injury. *Brain Injury, 21,* 319-326.

Donders, J., & Kirkwood, M. W. (2013). Symptom validity assessment with special populations. In D. C. Carone & S. S. Bush (Eds.), *Mild traumatic brain Injury: Symptom validity assessment and malingering* (pp. 399-410). New York: Springer Publishing Company.

Doss, R. C., Chelune, G. J., & Naugle, R. I. (1999). Victoria Symptom Validity Test. *Journal of Forensic Neuropsychology*, *1*, 5-20.

Duff, K., Spering, C. C., O'Bryant, S. E., Beglinger, L. J., Moser, D. J., Bayless, J. D., Culp, K. R., Mold, J. W., Adams, R.L., & Scott, J. G. (2011). The RBANS Effort Index: Base rates in geriatric samples. *Applied Neuropsychology*, *18*, 11-17.

Duncan, A. (2005). The impact of cognitive and psychiatric impairment of psychotic disorders on the Test of Memory Malingering (TOMM). *Assessment*, *12*, 123-129.

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*, 101-106.

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*, 101-106.

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*, 1218-1237.

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*, 1218-1237.

Flaro, L., Green, P., & Robertson, E. (2007). Word Memory Test failure 23 times higher in mild brain injury than in parents seeking custody: The power of external incentives. *Brain Injury*, *21*, 373-383.

Fox, D. D. (2011). Symptom validity test failure indicates invalidity of neuropsychological tests. *The Clinical Neuropsychologist*, *25*, 488-495.

Fox, D., & Lees-Haley, P. (1995, June). *Preliminary validation of a new test of malingered memory problems: The CTAM*. Presented at the Computerized Behavioral Testing of Humans in Neurotoxicology Research. Portland, OR.

Gervais, R. O., Rohling, M. L., Green, P., & Ford, W. (2004). A comparison of WMT, CARB, and TOMM failure rates in non-head injury disability claimants.*Archives of Clinical Neuropsychology*, *19*, 475-487.

Gervais, R. O., Russell, A. S., Green, P., Allen, L. M., Ferrari, R., & Pieschl, S. D. (2001). Effort testing in patients with fibromyalgia and disability incentives. *The Journal of Rheumatology*, *28*, 1892-1899.

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 (Ed.), *Assessment of feigned cognitive impairment* (pp. 281-309). New York, NY: The Guilford Press.

Goldberg, J. O., & Miller, H. R. (1986). Performance of psychiatric inpatients and intellectually deficient individuals on a task that assesses the validity of memory complaints. *Journal of Clinical Psychology*, *42*, 792-795.

Gorissen, M., Sanz, J. C., & Schmand, B. (2005). Effort and cognition in schizophrenia patients. *Schizophrenia Research*, *78*, 199-208.

Green, P. (2009). Questioning common assumptions in depression. In J. E. Morgan & J. J. Sweet (Eds.), *Neuropsychology of malingering casebook*, (pp. 132-144). New York, NY: Psychology Press.

Green, P. (2011). Comparison between the Test of Memory Malingering (TOMM) and the Nonverbal Medical Symptom Validity Test (NV-MSVT) in adults with disability claims. *Applied Neuropsychology, 18*, 18-26.

Green, P. (2013). Noncredible explanations of noncredible performance on symptom validity tests. In D. A. Carone & S. S. Bush (Eds.), *Mild traumatic brain injury: Symptom validity assessment and malingering*, (pp. 73-99). New York: Springer.

Green, P., & Iverson, G. L. (2001). Validation of the computerized assessment of response bias in litigating patients with head injuries. *The Clinical Neuropsychologist*, *15*(4), 492-497.

Green, P., Iverson, G. L., & Allen, L. (1999). Detecting malingering in head injury litigation with the Word Memory Test. *Brain Injury, 13*(10), 813-819.

Green, P., Rohling, M. L., Lees-Haley, P. R., & Allen, L. M. (2001). Effort has a greater effect on test scores than severe brain injury in compensation claimants. *Brain Injury*, *15*(12), 1045-1060.

Greiffenstein, M. F., & Baker, W. J. (2006). Miller was (mostly) right: Head injury severity inversely related to simulation. *Legal and Criminological Psychology*, *11*, 131-145.

Greiffenstein, M., Gervais, R., Baker, W. J., Artiola, L., & Smith, H. (2013). Symptom Validity Testing in Medically Unexplained Pain: A Chronic Regional Pain Syndrome Type 1 Case Series. *The Clinical Neuropsychologist*, *27*(1), 138-147.

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*, 1176–1190.

Greve, K. W., Bianchini, K. J., Etherton, J. L., Meyers, J. E., Curtis, K. L., & Ord, J. S. (2010). The Reliable Digit Span test in chronic pain: Classification accuracy in detecting malingered pain-related disability. *The Clinical Neuropsychologist*, *24*, 137-152.

Greve, K. W., Etherton, J. L., Ord, J., Bianchini, K. J., & Curtis, K. L. (2009). Detecting malingered pain-related disability: Classification accuracy of the Test of Memory Malingering. *The Clinical Neuropsychologist*, *23*, 1250-1271.

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.

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*, 709-719.

Guidotti Breting, L. M., & Sweet, J. J. (2013). Freestanding cognitive symptom validity tests: Use and selection in mild traumatic brain injury. In D. C. Carone & S. S. Bush (Eds.), *Mild traumatic brain Injury: Symptom validity assessment and malingering* (pp. 145-157). New York: Springer Publishing Company.

Gunner, J. H., Miele, A. S., Lynch, J. K., & McCaffrey, R. J. (2012). The Albany consistency index for the Test of Memory Malingering. *Archives of Clinical Neuropsychology*, *27*, 1-9.

Heintz, C. E., van Tricht, M. J., van der Salm, S. M., van Rootselaar, A. F., Cath, D., Schmand, B., & Tijssen, M. A. (2013). Neuropsychological profile of psychogenic jerky movement disorders: importance of evaluating non-credible cognitive performance and psychopathology. *Journal of Neurology, Neurosurgery & Psychiatry, 84,* 862-867.

Henry, M., Merten, T., Wolf, S. A., & Harth, S. (2010). Nonverbal Medical Symptom Validity Test performance of elderly healthy adults and clinical neurology patients. *Journal of Clinical and Experimental Neuropsychology, 32,* 19-27.

Hintze, J. (2013). *PASS 12*. Kaysville, Utah: NCSS, LLC.

Hurley, K. E., & Deal, W. P. (2006). Assessment instruments measuring malingering used with individuals who have mental retardation: Potential problems and issues. *Mental Retardation*, *44*, 112–119.

Iverson, G. L. (2005). Outcome from mild traumatic brain injury. *Current Opinion in Psychiatry, 18*, 301-317.

Iverson, G. L. (2010). Detecting exaggeration, poor effort, and malingering in neuropsychology. In A. MacNeill Horton & L. C. Hartlage (Eds.), *The handbook of forensic neuropsychology* (2nd ed.) (pp. 91-136). New York: Springer Publishing Company.

Iverson, G. L., & Franzen, M. D. (1994). The Recognition Memory Test, Digit Span, and Knox Cube Test as markers of malingered memory impairment. *Assessment, 1,* 323–334.

Kennepohl, S., Shore, D., Nabors, N., & Hanks, R. (2004). African American acculturation and neuropsychological test performance following traumatic brain injury. *Journal of the International Neuropsychological Society, 10*, 566–577.

Kiewel, N. A., Wisdom, N. M., Bradshaw, M. R., Pastorek, N. J., & Strutt, A. M. (2012). A Retrospective Review of Digit Span-Related Effort Indicators in Probable Alzheimer's Disease Patients. *The Clinical Neuropsychologist*, *26*(6), 965-974.

Kim, M. S., Boone, K. B., Victor, T., Marion, S. D., Amano, S., Cottingham, M. E., Ziegler, E. A., & Zeller, M. A. (2010a). The Warrington Recognition Memory Test for words as a measure of response bias: Total score and response time cutoffs developed on “real world” credible and noncredible subjects. *Archives of Clinical Neuropsychology*, *25*, 60-70.

Kim, N., Boone, K. B., Victor, T., Lu, P., Keatinge, C., & Mitchell, C. (2010b). Sensitivity and specificity of a digit symbol recognition trial in the identification of response bias. *Archives of Clinical Neuropsychology, 25*, 420-428.

Lange, R. T., Pancholi, S., Bhagwat, A., Anderson-Barnes, V., & French, L. M. (2012). Influence of poor effort on neuropsychological test performance in US military personnel following mild traumatic brain injury. *Journal of Clinical and Experimental Neuropsychology*, *34*, 453-466.

Lange, R. T., Sullivan, K., & Anderson, D. (2005). Ecological validity of the WMS-III Rarely Missed Index in personal injury litigation. *Journal of Clinical and Experimental Neuropsychology*, *27*, 412-424.

Larrabee, G. J. (2003). Detection of malingering using atypical performance patterns on standard neuropsychological tests. *The Clinical Neuropsychologist*, *17*(3), 410-425.

Lezak, M.D. (1995). *Neuropsychological assessment* (3rd ed.). New York: Oxford University Press

Locke, D. E., Smigielski, J. S., Powell, M. R., & Stevens, S. R. (2008). Effort issues in post-acute outpatient acquired brain injury rehabilitation seekers. *Neurorehabilitation*, *23*, 273-281.

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. *Clinical Neuropsychologist, 17*, 426–440.

Marshall, P., & Happe, M. (2007). The performance of individuals with mental retardation on cognitive tests assessing effort and motivation. *The Clinical Neuropsychologist*, *21*, 826-840.

Meyers, J. E., & Diep, A. (2000). Assessment of malingering in chronic pain patients using neuropsychological tests. *Applied Neuropsychology, 7*, 133-139.

Meyers, J. E., & Volbrecht, M. E. (1998). Validation of reliable digits for detection of malingering. *Assessment*, *5*, 303-307.

Meyers, J. E., & Volbrecht, M. E. (2003). A validation of multiple malingering detection methods in a large clinical sample. *Archives of Clinical Neuropsychology*, *18*(3), 261-276.

Meyers, J. E., Volbrecht, M., Axelrod, B. N., & Reinsch-Boothby, L. (2011). Embedded symptom validity tests and overall neuropsychological test performance. *Archives of clinical neuropsychology*, *26*(1), 8-15.

Miller, D. C. (2012). Review of the Amsterdam Short-Term Memory Test. In The Fourteenth Mental Measurements Yearbook. Available from http://buros.unl.edu/buros/

Millis, S. (1992). The Recognition Memory Test in the detection of malingered and exaggerated memory deficits. *The Clinical Neuropsychologist, 6*, 406–414.

Moore, B. A., & Donders, J. (2004). Predictors of invalid neuropsychological test performance after traumatic brain injury. *Brain Injury*, *18*, 975-984.

Morey, L. C. (1991). *Personality Assessment Inventory: Professional manual*. Odessa, FL: Psychological Assessment Resources.

Moss, A., Jones, C., Fokias, D., & Quinn, D. (2003). The mediating effects of effort upon the relationship between head injury severity and cognitive functioning. *Brain Injury*, *17*(5), 377-387.

Nabors, N.A., Evans, J.D., & Strickland, T.L. (2000). Neuropsychological assessment and intervention with African Americans. In E. Fletcher-Janzen, T.L. Strickland, & C.R. Reynolds (Eds.), *Handbook of Cross-Cultural Neuropsychology* (pp. 31-42). New York, NY: Kluwer Academic/Plenum Publishers.

O'Bryant, S. E., Finlay, C. G., & O'Jile, J. R. (2007). TOMM Performances and self-reported symptoms of depression and anxiety. *Journal of Psychopathology and Behavioral Assessment*, *29*, 111-114.

Ord, J. S., Boettcher, A. C., Greve, K. W., & Bianchini, K. J. (2010). Detection of malingering in mild traumatic brain injury with the Conners' Continuous Performance Test–II. *Journal of Clinical and Experimental Neuropsychology*, *32*, 380-387.

Paniak, C., Reynolds, S., Toller-Lobe, G., Melnyk, A., Nagy, J., & Schmidt, D. (2002). A longitudinal study of the relationship between financial compensation and symptoms after treated mild traumatic brain injury. *Journal of Clinical and Experimental Neuropsychology*, *24*, 187-193.

Powell, M. R., Locke, D. E., Smigielski, J. S., & McCrea, M. (2011). Estimating the diagnostic value of the Trail Making Test for suboptimal effort in acquired brain injury rehabilitation patients. *The Clinical Neuropsychologist*, *25,* 108-118.

Randolph, C. (1998). *Repeatable Battery for the Assessment of Neuropsychological Status.* San Antonio, TX: The Psychological Corporation.

Rapport, L. J., Farchione, T. J., Coleman, R. D., & Axelrod, B. N. (1998). Effects of coaching on malingered motor function profiles. *Journal of Clinical and Experimental Neuropsychology*, *20*, 89-97.

Ray, C. L. (2012). Assessment of feigned cognitive impairment: A cautious approach to the use of the Test of Memory Malingering for individuals with intellectual disability. *Open Access Journal of Forensic Psychology, 4,* 24-50.

Rees, L. M., Tombaugh, T. N., & Boulay, L. (2001). Depression and the test of memory malingering. *Archives of Clinical Neuropsychology*, *16,* 501-506.

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*, 10-20.

Rohling, M. L., Binder, L. M., & Langhinrichsen-Rohling, J. (1995). Money matters: A meta-analytic review of the association between financial compensation and the experience and treatment of chronic pain. *Health Psychology*, *14*, 537-547.

Rohling, M. L., Green, P., Allen, L. M., & Iverson, G. L. (2002). Depressive symptoms and neurocognitive test scores in patients passing symptom validity tests. *Archives of Clinical Neuropsychology, 17*, 205-222.

Ross, S. R., Krukowski, R. A., Putnam, S. H., & Adams, K. M. (2003). The Memory Assessment Scales in the detection of incomplete effort in mild head injury. *The Clinical Neuropsychologist*, *17*, 581-591.

Ross, S. R., Putnam, S. H., & Adams, K. M. (2006). Psychological disturbance, incomplete effort, and compensation-seeking status as predictors of neuropsychological test performance in head injury. *Journal of Clinical and Experimental Neuropsychology*, *28*, 111-125.

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 (Ed.), *Assessment of feigned cognitive impairment* (pp. 405-427). New York, NY: The Guilford Press.

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*, 437-453.

Silver, J. M. (2012). Effort, exaggeration and malingering after concussion. *Journal of Neurology, Neurosurgery & Psychiatry*, *83*, 836-841.

Slick, D. J., & Sherman, E. M. S. (2013). Differential diagnosis of malingering. In D. A. Carone & S. S. Bush (Eds.), *Mild traumatic brain injury: Symptom validity assessment and malingering* (pp. 57-72). New York: Springer.

Spielberger, C. (1983). *State-Trait Anxiety Inventory (Form Y)*. Palo Alto, California: Mind Garden.

Steele, C.M. & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69*, 797–811.

Stevens, A., Friedel, E., Mehren, G., & Merten, T. (2008). Malingering and uncooperativeness in psychiatric and psychological assessment: Prevalence and effects in a German sample of claimants. *Psychiatry Research*, *157*, 191-200.

Strutt, A. M., Scott, B. M., Lozano, V. J., Tieu, P. G., & Peery, S. (2012). Assessing sub-optimal performance with the Test of Memory Malingering in Spanish speaking patients with TBI. *Brain Injury*, *26*, 853-863.

Stulemeijer, M., Andriessen, T. M. J. C., & Brauer, J. M. P., Vos, P. E., Van der Werf, S. (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*, 309-318.

Suhr, J. A. (2003). Neuropsychological impairment in fibromyalgia: Relation to depression, fatigue, and pain. *Journal of Psychosomatic Research*, *55*, 321–329.

Suhr, J., & Spickard, B. (2007). Including measures of effort in neuropsychological assessment of pain- and fatigue-related medical disorders. In K.B. Boone (Ed.), *Assessment of feigned cognitive impairment* (pp. 259-280). New York, NY: The Guilford Press.

Sumanti, M., Boone, K. B., Savodnik, I., & Gorsuch, R. (2006). Noncredible psychiatric and cognitive symptoms in a workers' compensation “stress” claim sample. *The Clinical Neuropsychologist*, *20,* 754-765.

Tan, J. E., Slick, D. J., Strauss, E., & Hultsch, D. F. (2002). How'd they do it? Malingering strategies on symptom validity tests. *The Clinical Neuropsychologist*, *16*, 495-505.

Teichner, G., & Wagner, M. T. (2004). The Test of Memory Malingering (TOMM): Normative data from cognitively intact, cognitively impaired, and elderly patients with dementia. *Archives of Clinical Neuropsychology*, *19*, 455-464.

Temple, R. O., McBride, A. M., Horner, M. D., & Taylor, R. M. (2003). Personality characteristics of patients showing suboptimal cognitive effort. *The Clinical Neuropsychologist*, *17*, 402-409.

Terrell, F., & Terrell, S.L. (1983). The relationship between race of examiner, cultural mistrust, and the intelligence test performance of black children. *Psychology in the Schools, 20,* 367-369.

Terrell, F., Terrell, S.L., & Taylor, J. (1980). Effects of race of examiner and type of reinforcement on the intelligence test performance of lower-class black children. *Psychology in the Schools, 17*, 270-272.

Tombaugh, T. N. (1997). The Test of Memory Malingering (TOMM): Normative data from cognitively intact and cognitively impaired individuals. *Psychological Assessment*, *9*, 260-268.

Tsanadis, J., Montoya, E., Hanks, R. A., Millis, S. R., Fichtenberg, N. L., & Axelrod, B. N. (2008). Brain injury severity, litigation status, and self-report of postconcussive symptoms. *The Clinical Neuropsychologist, 22*, 1080–1092.

Victor, T. L., Boone, K. B., & De La Rosa-Trujillo, B. A. (2005). *The use of effort tests in a mentally retarded population*. Poster presented at the National Academy of Neuropsychology Annual Convention, Tampa Bay, Florida, USA.

Vilar- López, R., Gomez-Rio, M., Caracuel-Romero, A., Llamas-Elvira, J., & Perez-Garcia, M. (2008b). Use of specific malingering measures in a Spanish sample. *Journal of clinical and experimental neuropsychology,30*, 710-722.

Vilar-López, R., Gómez-Río, M., Santiago-Ramajo, S., Rodríguez-Fernández, A., Puente, A.

E., & Pérez-García, M. (2008a). Malingering detection in a Spanish population with a known-groups design. *Archives of Clinical Neuropsychology, 23*, 365-377.

Vilar-López, R., Santiago-Ramajo, S., Gómez-Río, M., Verdejo-García, A., Llamas, J. M., & Pérez-García, M. (2007). Detection of malingering in a Spanish population using three specific malingering tests. *Archives of Clinical Neuropsychology, 22*, 379-388.

Walker, A.J., Batchelor, J., & Shores, A. (2009). Effects of education and cultural background on performance on WAIS-III, WMS-III, WAIS-R and WMS-R measures: Systematic review. *Australian Psychologist, 44*, 216-223.

Warrington, E. K. (1984). *Recognition Memory Test manual*. Berkshire, United Kingdom: NFER Nelson.

Williamson, D. J., Holsman, M., Chaytor, N., Miller, J. W., & Drane, D. L. (2012). Abuse, not financial incentive, predicts non-credible cognitive performance in patients with psychogenic non-epileptic seizures. *The Clinical Neuropsychologist*, *26*, 588-598.

Yang, C. C., Kao, C. J., Cheng, T. W., Yang, C. C., Wang, W. H., Yu, R. L., Hsu, Y., & Hua, M. S. (2012). Cross-cultural Effect on Suboptimal Effort Detection: An Example of the Digit Span Subtest of the WAIS-III in Taiwan. *Archives of Clinical Neuropsychology*, *27*, 869-878.

Young, J. C., Caron, J. E., Baughman, B. C., & Sawyer, R. J. (2012). Detection of Suboptimal Effort with Symbol Span: Development of a New Embedded Index. *Archives of Clinical Neuropsychology*, *27*, 159-164.

CHAPTER 4: CURRENT RESEARCH AND OBJECTIVES

There is a need to clarify and identify statistical predictors of suboptimal effort. The majority of studies conducted to date have examined predictors of suboptimal effort using group comparison methodologies with relatively small samples and have examined one subset of those who might afford suboptimal effort during neuropsychological evaluation – those malingering illness. More sophisticated multivariable statistical methodologies offer the opportunity to evaluate a variety of predictors concurrently and to examine the relative predictive value of specific variables, while avoiding the statistical pitfalls of traditional univariate analytical techniques. The opportunity exists to begin to develop a predictive model of suboptimal effort; a model that incorporates a range of variables including those previously hypothesized and examined and other as-yet unexamined predictive variables.

The Research Comprising the Current Thesis

Objectives

The research reported in the current thesis was designed in order to meet the following objectives:

1. To clarify the relative importance of previously identified predictors of ETF (compensation-seeking, TBI severity, education) when those variables are incorporated into a multivariable predictive model employing a large sample.
2. To examine the relationship between other, previously unexamined variables (abnormal nonverbal behaviours, workplace accident) and ETF.
3. To examine the role that psychological (self-reported mood disorder, psychosis, pain disorder, substance use disorder) and personal history variables (psychiatric history, personal abuse history, substance abuse history) have in predicting ETF.
4. To examine the role of ethnicity and acculturation variables (being foreign born, English as a second language, age at which English was learned, years educated in

New Zealand, and number of years resident in New Zealand) in predicting ETF.

5. To systematically examine the literature that describes the relationship between negative affect and ETF in compensation-seeking samples.

To meet those objectives four studies were undertaken. The first study comprised a broad examination of the predictors of ETF investigating previously examined predictors in a multivariable model alongside previously unexamined variables. In the second study, the predictive relationship between psychological and personal history variables and ETF was more closely examined. The third study comprised an investigation into the relationship between acculturation variables and ETF, specifically testing variables that have been identified by Salazar et al. (2007) in a multivariable model and with a new and ethnically diverse population from that previously studied. In the fourth study, depressive symptomatology and its relationship with ETF was examined via a systematic review of the literature. The specific aims and hypotheses of each study are detailed in the following chapters.

CHAPTER 5: EFFORT TEST FAILURE: TOWARDS A PREDICTIVE MODEL

Reference (see APPENDIX B)

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.

Abstract

Predictors of effort test failure were examined in an archival sample of 555 traumatically brain injured (TBI) adults.

Logistic regression models were used to examine whether compensation-seeking, injury-related, psychological, demographic, and cultural factors predicted effort test failure (ETF). ETF was significantly associated with compensation-seeking (OR = 3.51, 95% CI [1.25, 9.79]), low education (OR: .83 [.74, .94]), self-reported mood disorder (OR: 5.53 [3.10, 9.85]), exaggerated displays of behavior (OR: 5.84 [2.15, 15.84]), psychotic illness (OR: 12.86 [3.21, 51.44]), being foreign-born (OR: 5.10 [2.35, 11.06]), having sustained a workplace accident (OR: 4.60 [2.40, 8.81]), and mild traumatic brain injury severity compared with very severe traumatic brain injury severity (OR: 0.37 [0.13, 0.995]). ETF was associated with a broader range of statistical predictors than has previously been identified and the relative importance of psychological and behavioral predictors of ETF was evident in the logistic regression model. Variables that might potentially extend the model of ETF are identified for future research efforts.

An estimated 1.7 million traumatic brain injury (TBI) related emergency department visits occur each year in the United States, from which 275,000 individuals are hospitalized (Faul, Xu, Wald & Coronado, 2010). Following TBI many people undergo neuropsychological assessment, with those assessments frequently being conducted in order to determine whether or not cognitive impairments exist sufficient to entitle the individual to disability or insurance payments. It is recognized that a substantial proportion (up to 30-40%) of test examinees fail to put forth optimal effort during cognitive testing (Larrabee, 2007). Accordingly, there has been a growing interest in objectively assessing effort, and identifying risk factors for, or predictors of, low test-taking effort (Chaftez & Prentkowski, 2011; Donders & Moore, 2007; Moore & Donders, 2004; Dandachi-Fitzgerald, Ponds, Peters, & Merckelbach, 2011).

The provision of financial gains via compensation-seeking/litigation for disability has been repeatedly shown to predict effort test failure (ETF) and malingering (Bianchini, Curtis, & Greve, 2006, Henry et al., 2011; Paniak et al., 2002). ETF describes suboptimal performance on specific symptom validity tests (SVTs) and embedded cognitive measures that have been validated as tests of effort. It is important to note that ETF does not equate to neurocognitive malingering. Malingering is behavior that is evident in a subset of the group of individuals displaying ETF and, although ETF is essential for the classification of neurocognitive malingering, ETF alone is not sufficient for that classification (Larrabee, Greiffenstein, Greve, & Bianchini, 2007; Slick, Sherman, & Iverson, 1999). Better knowledge of ETF predictors can help explain why some patients present with this response style besides secondary gain.

With the exception of gender, which has very seldom been found to be predictive of ETF, demographic variables of education, age, and ethnicity have been variably associated with ETF. Findings in respect to education have trended toward lower education being predictive of ETF (e.g., Babikian, Boone, Lu & Arnold, 2006; Greve, Etherton, Bianchini, & Curtis, 2009; Mahdavi, Mokari, & Amiri, 2011; Stulemeijer, Andriessen, Brauer, Vos, & Van Der Werf, 2007) and even in those studies that have reported no effect of education there has been a tendency for education to be lower in the group failing SVTs (e.g., Armistead-Jehle, 2010; Ord, Boettcher, Greve, & Bianchini, 2010).

Advanced age is a known predictor of prolonged disability from work following illness (Flach, Krol, & Groothoff, 2008) and as such there are grounds for reasoning that advanced age might be a predictor of ETF. In fact, findings in respect to age have been inconsistent and somewhat sample-dependent. For example, Grote et al., (2000) and Donders and Boonstra (2007) found an age-related relationship with older people more likely to display ETF, while Tombaugh (1997) and others (e.g., Lange, Iverson, Brooks, & Rennison, 2010) have reported no significant relationship.

Similarly inconsistent findings have been found in respect to ethnicity. For example Meyers, Volbrecht, Axelrod and Reinsch-Boothby (2011) found no evidence of ethnicity being a moderator of ETF, while Salazar, Lu, Wen, and Boone, (2007) found a significant impact of ethnicity on ETF., Caucasians scored significantly higher than Hispanics on the embedded measures of Digit Span age corrected scaled score and Reliable Digit Span, and higher than African Americans on RAVLT and Rey-Osterrieth embedded measures. Salazar et al. found that in English as a second language groups effort scores were related to the age at which English was learned, but years living in the USA or years educated in the USA were not related to effort test scores. Foreign-born immigrant status, independent of ethnicity, has not been specifically investigated.

Psychological factors have been found to have a complex association with ETF. Although earlier researchers (e.g., Ashendorf, Constantinou, & McCaffrey, 2004; Rees, Tombaugh, & Boulay, 2001) found no relationship between depression and ETF in non-litigating psychiatric participants, other more recent studies including non-litigating psychiatric patients (e.g., Dandachi-Fitzgerald, Ponds, Peters, & Merckelbach, 2011; Gorissen, Sanz, & Schmand, 2005), compensation-seeking depressed participants (Green, 2009), non-litigating, non-compensation seeking TBI participants (Bierley et al., 2001), and compensation-seeking neurological participants (Armistead-Jehle, 2010; Stulemeijer et al., 2007; Rohling, Green, Allen, & Iverson, 2002; Suhr, Tranel, Wefel, & Barrash, 1997), have detected a relationship, but Schroeder & Marshall (2011) did not. Schroeder and Marshall (2011) have suggested that reliance on single SVT failure to indicate ETF might account for some of the inconsistent findings in this field and this will be further examined in the present study. Psychosis is a factor that has been shown to be associated with ETF (Goldberg, Back-Madruga, and Boone, 2007), although again, Schroeder and Marshall (2011) have questioned that finding.

There have been very few studies of the relationship that might exist between interview behaviors and test-taking effort. Greiffenstein and Baker (2006) drew on Miller's (1961a, 1961b) landmark lectures on *accident neurosis*, by investigating a behavioral component of test taker effort. They found that behavior, in the form of the *floridity* (frequency) of symptom reporting, was related to ETF such that higher symptom-floridity predicted ETF. A similar relationship between over-reporting and ETF was reported by Dandachi-Fitzgerald and co-workers (2011).

Injury-related variables have been examined as predictors. Specifically, many researchers have noted a paradoxical finding of better effort test performance among those with more severe brain injuries (e.g., Carone, 2008; Green, Iverson, & Allen, 1999; Green, Rohling, Lees-Haley, & Allen, 2001; West, Curtis, Greve, & Bianchini, 2011). Having a workplace accident has not been specifically investigated as a correlate of poor effort independent of compensation-seeking/litigation. Workplace injuries are associated with an increased risk of prolonged disability, blaming others for the injury, posttraumatic stress, and litigation (Mason, Wardrope, Turpin, & Rowlands, 2002); this might afford a potential context for ETF during neuropsychological evaluation.

Overall, the literature to-date has identified a number of variables that increase the probability or likelihood of ETF but with few exceptions (Donders & Boonstra, 2007; Moore & Donders, 2004) these variables have seldom been examined concurrently using multivariable statistical techniques. The aim of the current study was to examine predictors of ETF in an archival sample of consecutive referrals to a private clinical neuropsychology practice. Based on the literature it was anticipated that compensation-seeking (Bianchini, Curtis, & Greve, 2006), demographic and psychological variables would predict ETF. Specifically, it was hypothesised that age (Flach, Krol, & Groothoff, 2008), education (Babikian, Boone, Lu & Arnold, 2006), ethnicity and immigration status (Salazar, Lu, Wen, & Boone, 2007) would each be associated with ETF. In addition, a self-reported mood disorder (Green, 2009), displaying psychotic illness (Goldberg, Back-Madruga, and Boone, 2007), and displaying exaggerated symptomatic or behavioral floridity (Greiffenstein & Baker, 2006) were anticipated to predict ETF. Finally, it was hypothesized that injury variables including having sustained a workplace accident (Mason et al., 2002), and that ETF would be more likely in those with mTBI than more severe brain injuries (Green, Iverson, & Allen, 1999).

Method

Participants

An archival sample of 555 consecutive referrals to a private clinical neuropsychology practice in Auckland, New Zealand was examined. Participants included had sustained a TBI and were over the age of 16 years at the time of assessment. The data set was collected over the period 2001 to 2007 inclusive.

Participants were excluded on the basis of having a pre-existing history of mental retardation or dementia.

Compensation-seeking.

Most of the sample (*n* = 470; 84.7%) were seeking compensation continuance or seeking entitlement to compensation. Compensation was defined as worker's compensation income replacement (insurance) payments or disability social security benefits. None of the sample were engaged in litigation (litigation for damages is specifically precluded under New Zealand no-fault accident compensation legislation). The majority of the compensation-seekers (*n* = 422; 90%) were seeking continuance of worker's compensation payments while 46 (10%) were seeking continuance of social security benefits. A sizeable minority (*n* = 85; 15.3%) of the total sample were ineligible for compensation.

Demographic variables.

Men made up 72.8% of the sample. The mean age was 41 (*SD* = 12.31, range = 16-76) years. Ethnicity was as follows: European/White; (*n* = 418, 75.3%), Maori/Pacific Island (*n* = 105, 18.9%), Indo/Asian (*n* = 32, 5.8%). The foreign-born subgroup consisted of any non-New Zealand born individuals and comprised 18% (*n* = 99) of the sample. The sample had, on average, 11.8 (*SD* = 2.5, range = 2-21) years of education.

Injury variables

Severity of injury was classified on the basis of Glasgow Coma Scale (GCS) scores at the Emergency Department, duration of posttraumatic amnesia (PTA) assessed using the Westmead Post-traumatic Amnesia scale (Shores, Marosszeky, Sandanam, & Batchelor, 1986), and duration of loss of consciousness (LOC). When PTA data was unavailable, duration of PTA was assessed via clinical interview, which sought to establish the onset of continuous recall (Gronwall & Wrightson, 1980). PTA duration was estimated from clinical interview alone in 7.5% (*n* = 42) of cases where there had been no medical attention at the time of injury. In all such cases severity was estimated as mTBI. Comparing the group with mTBI injury classified from interview alone with the mTBI group classified from interview and documented medical data (i.e., GCS, PTA, LOC) (n = 225) revealed no differences in respect of age, gender, years of education, or classification with ETF.

Duration of loss of consciousness (LOC) was assessed in accordance with the guidelines of Ruff et al. (2009), specifically: that the duration of LOC should result from impact not other medical causes and that LOC was determined from collateral reports of witnesses present at the scene (e.g., paramedic) or from hospital medical records, not from self-report of the participant.

Severity of mild TBI (mTBI) was defined according to the WHO Collaborating Task Force mTBI diagnostic criteria (Carroll, Cassidy, Holm, Kraus, & Coronado, 2004). Complicated mTBI was differentiated from mTBI according to Williams, Levin and Eisenberg (1990). Moderate to severe TBIs were defined using the Teasdale and Jennett (1974) and Teasdale (1995) criteria.

TBI severity was classified as shown in Table 1. The mTBI-complicated and Moderate TBI groups were combined to ensure that parameter estimates were based on adequate numbers of cases and that there were no empty cells. This approach is supported by the findings of Kashluba, Hanks, Casey, and Millis (2008) who found that few differences in outcome are seen in cases of mTBI-complicated injuries and moderate TBI. Additionally, preliminary analyses that showed that the mTBI-complicated and Moderate TBI groups did not differ in respect of their predictive relationships with ETF. The Very Severe TBI and Extremely Severe TBI groups were similarly combined on the basis that preliminary analyses revealed that the two groups did not differ in respect of their predictive relationships with ETF.

Table 1

Injury Severity Criteria and Sample Characteristics

Descriptor	Criteria	*n* (%)
mTBI	LOC < 30 mins, PTA < 24 hours, GCS 13-15 at 30 mins	275 (49.5)
mTBI complicated	LOC <30 mins, PTA <24 hours, GCS 13-15 at 30 mins, visible (on CT brain imaging) intracranial abnormality not requiring surgery	38 (6.8)
Moderate	PTA 1-24 hours, GCS 9-12	54 (9.7)
Severe	PTA 1-7 days, GCS 3-8	79 (14.2)
Very severe	PTA 1-4 weeks, GCS 3-8	69 (12.4)
Extremely severe	PTA >4 weeks, GCS 3-8	40 (7.2)

Note. LOC = Loss of consciousness; mTBI = mild traumatic brain injury; PTA = posttraumatic amnesia; GCS = Glasgow Coma Scale.

Measures

All participants completed three measures of effort including one forced-choice symptom validity test. All participants completed the Test of Memory Malingering (TOMM; Tombaugh, 1996), the Fifteen Item Test (FIT; Rey, 1964), and Reliable Digit Span (RDS; Greiffenstein, Baker, & Gola, 1994).

Effort classification.

Effort was classified dichotomously – valid effort (VE) or ETF. Cut-offs employed for each measure were set to ensure >90% specificity following Baker, Donders, and Thompson (2000) and Boone, Salazar, Lu, Warner-Chacon, and Razani (2002). Consequently, the Greve, Bianchini, and Doane (2006) cut-off of <47 was employed for TOMM2 and TOMM Retention; An RDS cutoff of <8 was employed according to guidelines of Suhr and Barrash's

(2007) review; An FIT cutoff of <9 was used in accordance with the findings of Boone et al. (2002). ETF was operationalized as failure on any two measures in accord with the findings of Victor, Boone, Serpa, Buehler, & Ziegler (2009) or below chance performance on either TOMM2 or TOMM Retention (<18/50 at the 95% confidence interval using the binomial distribution). VE was classified using the requirement that participants pass all three effort measures at the cutoffs described above.

Psychological dimensions.

Psychological data were available from the archive records for each participant. All participants had undergone an approximately 60 minute semi-structured clinical interview with a licensed psychologist trained in clinical psychology (JW). Participants reporting a mood disorder and/or a psychotic disorder and who met the Diagnostic and Statistical Manual of Mental Disorders IV-TR diagnostic criteria (American Psychiatric Association, 2000) were coded as having a self-reported mood disorder or self-reported psychosis.

In a subsection of the sample exaggerated displays of behavior had been noted and were coded as Behavioral Floridity (BFlor). BFlor conceptually builds on the work of Miller (1961a, 1961b, 1972) and Greiffenstein and Baker (2006). Miller's original lectures (1961a, 1961b) described other aspects of abnormal behavior seen in this population. He described florid behaviors – "gross dramatization of symptoms," (p. 922) including extreme behaviors such as "groaning", "slumping forward with head in hands," "quivering," and using what might now be described as catastrophising language when symptom-reporting ("terrible," "agonizing"). Building on Miller's papers, BFlor is defined here as extreme displays of symptom-related behavior and no assessment of internal states including cognitive styles, beliefs, perceptions, etc., is implied.

For the purposes of this exploratory study, dichotomous coding of BFlor was made by clinical judgment of the principal author, a licensed clinical psychologist. A conservative approach was taken toward coding BFlor to ensure that only extreme and disproportionate displays of behavior comprised BFlor and particular care was taken to avoid mis-coding psychotic symptomatology, when evident, as BFlor. Examples of BFlor included lying on the floor with complaints of extreme tiredness following interview, requiring each and every question to be repeated, dry-retching while reporting extreme fatigue, atypical levels of symptom endorsement, extreme slowness of all movements, dramatic facial displays of tiredness, pathos, or marked affective blunting in the absence of affective/psychotic disturbance, atypical displays of language use, e.g., answering yes/no in German language despite the participant having never been fluent in the German language; missing the first spoken phoneme from each word. Because only the most extreme forms of behavior were considered to be BFlor, relatively few participants were coded as such (n = 36, 6.5%).

The participants with diagnosed psychotic illness (n = 15) were all male, seven of whom were diagnosed with a psychotic disorder due to a general medical condition (TBI) with symptom onset ranging from emergence from PTA to six years post-injury. Three were diagnosed with schizophrenia with age of onset ranging from late teens to early twenties and pre-dating their injuries. Two were diagnosed with delusional disorder with onset in late teens and early twenties and before injury. Two were diagnosed with schizoaffective disorder, one developing following immigration nine years before injury and one developing four months before TBI. One was differentially diagnosed with schizophrenia/substance-induced psychosis with symptom onset before injury. All participants but one were under psychiatric review, four were taking no psychotropic medications, the remainder taking atypical antipsychotics. Four participants were taking anticonvulsant medication for seizure disorders. Nine had been hospitalized for assessment and treatment, eight of which were involuntary committals. In seven participants injury was considered to be the precipitant for developing psychotic symptoms and in those, all injuries were classified as moderate to very severe. All but one of the participants was receiving compensation/compensation-seeking.

Procedure

This study was approved by the Human Research Ethics Committees of Macquarie University and the University of New England.

Results

Of the 555 participants, 111 (20.0%) were classified with ETF. Of these, seven participants (1%) scored below chance on the TOMM trials. A total of 352 participants were classified with VE.

Preliminary Analyses

Table 2 shows the results of exploratory bivariate statistical analyses. Participants were grouped according to ETF status and comparisons across the variables of interest were undertaken. Because the bivariate analysis was exploratory no adjustment for multiple comparisons was undertaken.

Variables that were noted to have small to moderate statistically significant positive relationships with ETF included age, being foreign born, ethnicity, having a workplace accident, compensation-seeking, having a diagnosed (self-reported) mood disorder, having a diagnosed (self-reported) psychotic illness, and displaying BFlor. Variables with a small to moderate inverse relationship with ETF included years of education, and TBI severity. All relationships with variables in bivariate analyses were statistically significant and as such all were included for further analysis.

Bivariate relationships between the predictors that were identified by the preliminary analysis were calculated and are displayed in Table 3.

Table 2

Characteristics of the Participants Grouped According to Effort

Variable	Good effort		ETF		*p*	Effect Size (*d*/*ϕ*)
	M	*SD*	*M*	*SD*		
Age	40.65	13.16	43.68	10.57	.01	.25
Education, years	12.09	2.46	11.19	2.52	.001	.36
	n	%	*n*	%		
Ethnicity					.004	.15
European/White	281	79.4	73	20.6		
Maori/Pacific Island	56	68.3	26	31.7		
Indo/Asian	15	55.6	12	44.4		
Immigration status					<.001	.27
NZ born	309	81.5	70	18.5		
Foreign born	43	51.2	41	48.8		
Self-reported Mental Illness						
Affective disorder					<.001	.36
No	269	86.8	41	13.2		
Yes	83	54.2	70	45.8		
Psychosis					<.001	.17
No	347	77.3	102	22.7		
Yes	5	35.7	9	64.3		
Behavioral Floridity					<.001	.34
No	345	79.9	87	20.1		
Yes	7	22.6	24	77.4		
TBI Severity					<.001	.20
mTBI	164	68.9	74	31.1		
Moderate	61	78.2	17	21.8		
Severe	52	78.8	14	21.2		
Very Severe	75	92.6	6	7.4		
Place of accident					<.001	.28
Non-workplace	312	81.5	71	18.5		
Workplace	40	50.0	40	50.0		
Comp.seek/Continuance					<.001	.17
No	68	93.2	5	6.8		
Yes	284	72.8	106	27.2		

Note. *p*-values are from independent *t*-tests or Chi-Square tests. Effect sizes for *t*-tests are Cohen's *d* and Cramer's Phi (*ϕ*) for Chi-Square tests. Indo/Asian = participants reporting themselves of Indian or Asian ethnicity; NZ = participants reporting themselves New Zealand born; mTBI = mild traumatic brain injury ; BFlor = behavioural floridity; Comp. seeking/Continuance = compensation seeking or seeking continuance of compensation,

Table 3

Bivariate relationships between main predictors

	Age	Yrs ed	Foreign born	Mood disorder	Psychotic illness	BFlor	TBI Severity	Wrkplc Acc	Comp-seek
Yrs ed	-.04								
Foreign born	.04	.15**							
Mood disorder	.11*	.01	.33**						
Psychotic illness	-.03	-.12**	.26	-.34					
BFlor	-.01	-.03	.43*	.63***	-1.00**				
TBI severity	-.06	.04	-.05	-.36***	.25	-.61***			
Wkplc Acc	.15***	-.07	.09	.31**	-1.00***	.51*	-.48***		
Comp-seek	.08	-.06	.27	.01	.44	.74***	.10	.37	
Ethn[1]	.13	.19***	-.06	-.05	.34	.34	.03	-.06	.19

Note. Relationships among dichotomous variables are expressed as Gamma coefficients. Relationships between continuous/ordinal variables and dichotomous variables are expressed as Pearson product-moment correlation coefficients/point-biserial correlation coefficients. [1] Relationships between Ethnicity (categorical data) and Age, Years education, and Severity are expressed as r ($\sqrt{R^2}$) following one-way ANOVAs employing the continuous variables as dependent variables and Ethnicity as facto.r. Yrs ed = years of education; BFlor = behavioral floridity; TBI = traumatic brain injury; Wrkpl Acc = workplace accident; Comp-seek = compensation-seeking or seeking continuance of compensation; Ethn = ethnicity.

* $p < .05$. ** $p < .01$. *** $p < .001$

Most relationships between the statistical predictors were relatively weak, ranging from .003 to -.21. Having a mood disorder was noted to correlate with BFlor (Gamma = .63, $p < .001$) and with having sustained a workplace accident (Gamma = .31, $p < .01$), and was negatively correlated with TBI severity (Gamma = -.36, $p < .001$). BFlor was also negatively correlated with TBI severity (Gamma = -.61, $p < .001$) and was positively correlated with having sustained a workplace accident (Gamma = .51, $p < .05$) and with being compensation-seeking (Gamma = .74, $p < .001$). Compensation-seeking was not significantly correlated with any other variable.

Predictive Model

Logistic regression analyses with ETF as the dependent variable were undertaken. Years of education was included as a continuous variable. Dichotomous variables included immigrant status, presence of a self-reported mood disorder and self-reported psychotic illness, displays of BFlor, having a workplace accident, and compensation-seeking. Injury severity and ethnicity were analyzed as categorical data.

The fit of the model was examined in three ways. First, examination of residuals revealed 11 cases with large standardized residuals. Those cases were eliminated and the same results were obtained and for this reason all cases were included in the analyses. Second, the fit of possible alternative models which might have been appropriate in the light of the relatively skewed binary dependent were examined. The fit of a model with a complementary log-log link was trivially worse than the original model with a log link (AIC = 360.3, BIC = 418.2 versus AIC = 359.3 and BIC = 417.2 for the original model). A model with a log-log link fitted markedly less well (AIC = 369.1, and BIC = 427.0). The original model was therefore retained. Third, multicollinearity was examined using a method described by Belsley, Kuh & Welsch (1980). A program written for Stata (Statacorp, 2011) by Hardin (1995) provided condition numbers and variance decomposition proportions based on the singular value decomposition of the ***X*** matrix of independent variables. One of the condition numbers (19.5), while lower than the highest criterion of 30 suggested by Belsley, Kuh & Welsch (1980), was higher than the next-lowest criterion, 15. Examination of the variance decomposition proportions showed that, as might be expected, age and years of education showed reasonably high collinearity. However, tests of reduced models showed (1) that age was never significant, adjusted or unadjusted for years of education, and years of education was significant whether or not it was adjusted for age, and (2) that the significance of other effects were very similar whether both or either of the variables were included in the model. The original model, which included both age and years of education, was therefore retained.

As presented in Table 4 a number of statistical predictors of ETF were positively identified as making unique contributions (i.e., with all other variables held constant) in the final logistic regression model. Unadjusted odds ratios are provided in Table 4 for comparison with the adjusted odds ratios.

Demographic variables.

Age was significantly related to ETF in bivariate analysis but age did not make a significant contribution when adjusted for the effect of the other variables in the model. Years of education was a significant predictor of ETF such that a one-year increase in education reduced the odds of ETF by .83 (or by 17%).

Compensation-seeking.

Compensation-seeking was a significant statistical predictor of ETF in both preliminary bivariate analysis and when adjusted for the effect of the other variables in the model: The odds of ETF in the compensation-seeking sample are 3.5 times those for the non-compensation-seeking sample.

Ethnicity.

Unadjusted odds ratios suggested that ethnicity was predictive of ETF, but ethnicity was not significantly related to ETF when included in the final model that included being foreign born, which was a more powerful predictor of ETF. In comparison with being locally born, being foreign born increased the odds of ETF by five times.

Mental state variables.

Having a self-reported mood disorder, psychotic illness, and the display of BFlor were predictive of ETF, both when considering unadjusted odds ratios and in the final model. Self-reporting mood symptoms increased the odds of ETF by five times relative to those without a formally diagnosed mood disorder, and having a psychotic disorder increased the odds of ETF by about 13 times. The odds of ETF were almost six times higher in those displaying BFlor than those not displaying BFlor.

Injury-related variables.

Severity of TBI was a significant predictor of ETF when considering the unadjusted odds such that individuals who had sustained severe or very severe injuries were less likely to display ETF than those with milder injuries, but TBI severity was not a significant predictor when adjusted for the effect of other variables in the model. Having sustained a workplace accident was predictive of ETF when examined in a bivariate logistic regression model and also in the multivariable model. Having a workplace accident increased the odds of ETF by about 4.5 times relative to those not having a workplace accident.

Goodness of fit.

Goodness of fit of the final logistic regression model including all predictors was assessed via the Hosmer and Lemeshow test (χ^2 (8, N = 463) = 12.19, p = .14). The area under the receiver operating characteristic curve for the model was .87 (95% CI [.82, .91]), indicating excellent discrimination in predicting those with ETF and those without (Hosmer & Lemeshow, 2000) and Nagelkerke R^2 = .48.

Table 4

Logistic Regression Analysis of Effort Test Failure as a Function of Demographic and Injury Predictors

Variable	Adjusted OR (95% CI)	*p* value	Unadjusted OR (95% CI)	*p* value
Age	1.01 [0.98, 1.03]	.36	1.01 [1.002, 1.03]	.02
Years education	.83 [0.74, 0.94]	.005	.85 [0.77, 0.94]	.001
Foreign born	5.10 [2.35, 11.06]	<.001	3.90 [2.43, 6.26]	<.001
Ethnicity[1]		.19		.005
Pac.Isl vs. Euro/W	1.85 [0.92, 3.69]	.08	1.78 [1.05, 3.04]	.03
IndoAs. Vs. Euro/W	.90 [0.26, 3.09]	.86	3.07 [1.38, 6.86]	.006
Mood disorder	5.53 [3.10, 9.85]	<.001	5.53 [3.50, 8.74]	<.001
Psychotic illness	12.86 [3.21, 51.44]	<.001	6.12 [2.00, 18.67]	.001
BFlor	5.84 [2.15, 15.84]	.001	13.59 [5.67, 32.58]	<.001
TBI Severity[2]		.25		.001
Moderate vs Mild	.87 [0.40, 1.86]	.71	.61 [0.33, 1.12]	.11
Severe vs Mild	.41 [0.30, 1.63]	.41	.59 [0.31, 1.14]	.12
V.Severe vs Mild	.37 [0.13, 0.995]	.04	.17 [0.07, 0.42]	<.001
Workplace Accident	4.60 [2.40, 8.81]	<.001	4.39 [2.64, 7.30]	<.001
Compensation-seeking	3.51 [1.25, 9.79]	.01	5.07 [1.99, 12.93]	.001
Constant	.09	.03		

Note. BFlor = behavioral floridity; TBI = traumatic brain injury; Pac.Isl = Pacific Island; Euro/W = European/White; IndoAs = Indo-Asian; V.Severe = very severe TBI. [1] Ethnicity was dummy coded with European/White as the reference group. [2] Severity was dummy coded with mild traumatic brain injury as the reference group.

Discussion

The aim of this study was to examine potential correlates of ETF and to gain an impression of the relative importance of predictors of ETF toward identifying a model of ETF in a traumatic brain injury sample. It was hypothesized that compensation-seeking, demographic, psychological/behavioral and injury variables would predict ETF. The predictive relationships identified are statistical in nature not causal.

This study found that 20% of the total sample displayed ETF. This is a relatively low proportion compared with the previous literature. This finding is likely due to this study including a higher proportion of participants with severe injuries than is typically seen in this literature and that ETF is more likely in those with mild injuries than severe injuries (West, Curtis, Greve, & Bianchini, 2011). Also, the proportion of compensation-seeking participants was relatively lower than that which is typically seen and there is a known relationship between compensation-seeking and ETF (Bianchini, Curtis, & Greve, 2006). In the compensation-seeking participants with mTBI 44% of participants failed one effort measure and 30% failed two or more. Additionally, although the measures employed here are among the most frequently employed measures of effort, other sensitive measures (e.g., Word Memory Test (Green, 2003)) were not included and this may have contributed to the factors described above and account for this apparently low proportion of ETF. Finally, this study has required failure on more than one effort measure for classification of ETF. Thirty-seven percent of the sample failed one effort measure – a rate that is similar to that seen in other effort studies in civil litigation settings (Larrabee, 2003a). However, Victor and colleagues (2009) showed that sensitivity to ETF appears to fall when increasing the number of effort measures employed, thus reducing the apparent proportion of those displaying ETF, but specificity and positive predictive power increases, making findings safer and more stable.

This study confirmed that compensation-seeking was significantly predictive of ETF. This finding was consistent with a number of previous studies that have found a significant main effect of compensation-seeking on effort and symptom exaggeration (e.g., Bianchini, Curtis, & Greve, 2006; Henry et al., 2011; Paniak et al., 2002). A number of previous researchers have independently reported that when incorporated into a predictive model with other variables, compensation-seeking is not significantly predictive of ETF (Donders & Boonstra, 2007; Ross, Putnam, & Adams, 2006; Stulemeijer et al., 2007) or only modestly predictive (Moore & Donders, 2004). Those previous findings in combination with the present results suggest that holding other potentially predictive variables constant by using a multivariable research strategy can help to clarify the predictive power of one variable (e.g., compensation-seeking). As new predictors become identified sample-dependent findings will become less problematic. Furthermore this emphasizes that the factors that underpin ETF may be multiple and that the search for other predictors may be a fruitful exercise.

The study found that having sustained a workplace accident was a predictor of ETF independent of compensation-seeking status. Having a workplace accident has previously been reported to be associated with adjustment difficulties and affective disturbance (Mason et al., 2002) but in this sample having a workplace accident was only weakly related to developing a self-reported mood disorder. This suggests that there are non-affective drivers of ETF in those who have suffered workplace accidents. Occupational variables such as low worker satisfaction, work monotony, work stress, and low levels of autonomy/control are associated with prolonged disability following injury and illness (Dragano & Schneider, 2011; Krokstad, Johnsen & Westin, 2002; Kuoppala, Lamminpaa, Vaanen-Tomppo, & Hinkka, 2011) and their relationship with ETF may prove worthy of further examination.

In contrast with some previous studies that failed to find a relationship between psychiatric disturbance and ETF (Schroeder & Marshall, 2011), this study found a significant relationship between self-reported mood disorder and ETF, and between displaying psychotic illness and ETF. This study supports previous findings with clinical samples, including non-compensation-seeking samples that have found a relationship between psychiatric disturbance and ETF (e.g., Dandachi-Fitzgerald, et al., 2011; Gorissen, 2005; Rohling et al., 2002; Stulemeijer et al., 2007). Schroeder and Marshall (2011) found no significant ETF in a non-compensation-seeking psychiatric sample, and suggested that previous findings of low effort in psychiatric samples might partially be an artifact of reliance on one SVT to diagnose poor effort. This thesis is not supported by the present study, which required failure on two or more of three specific and embedded measures.

Reporting of mood symptoms and psychosis in this study might be considered an aspect of heightened symptom endorsement, which is known to be related to ETF (Greiffenstein & Baker, 2006). Boone and Lu (1999), Larrabee (2003b), Mooney, Speed and Sheppard (2005), Armistead-Jehle (2010), and Jones, Ingram, and Ben-Porath (2012) found heightened psychological symptom reporting but no significant over-reporting in those displaying ETF. Repeated factor-analytic studies using a variety of personality measures have found that emotional over-reporting and ETF represent independent constructs (Jones & Ingram, 2011; Ruocco et al., 2005; Nelson, Sweet, Berry, Bryant, & Granacher, 2007; Whiteside, Dunbar-Mayer, & Waters, 2009). These findings are further supported by the data of Sumanti, Boone, Savodnik, and Gorsuch (2006), Demakis, Gervais, and Rohling, (2008) and Dandachi-Fitzgerald et al., (2011). Thus, while some individuals might over-report both cognitive and psychological symptoms, the findings of Nelson et al. (2007) and others indicate that this is unlikely to occur throughout a large sample such as in the present study. In respect of those displaying psychotic illness in the current study, clinical file review shows that the psychotic individuals typically had long and well-documented histories of psychosis that pre-dated their injuries and/or they had sustained severe and unambiguous brain injuries and several had come under involuntary committal to receive treatment – simple fabrication of their psychotic symptoms appeared unlikely.

In the present study the display of exaggerated behavior (BFlor) was closely related to ETF. This finding indicates that some patients signal their likelihood to display ETF via exaggerated illness-related behaviors in the session and supports previously related research on the behaviors that are associated with ETF (Dandachi-Fitzgerald et al., 2011; Greiffenstein & Baker, 2006). This is an exploratory study only and has not attempted to tightly operationalize the abnormal behaviors of interest, but it is hoped that this study might spur future research, allowing better operational definition of the abnormal behaviors seen in this population. Ekman and co-workers have commented comprehensively on the facial behavioral displays seen in lying and in malingering specifically (Ekman & O'Sullivan, 2006) and this study supports their position that symptom exaggeration is detectable through overt behavior.

BFlor was found to be closely related to compensation-seeking status and to self-reporting a mood disorder. The shared variance between these variables raises the possibility of a common factor of symptom over-reporting being present in this group. A measure of psychological symptom over-reporting (e.g., MMPI-2 Fp) was not included in this study and future studies examining BFlor may well investigate this possibility.

Ethnicity has been examined relatively infrequently as a predictor of ETF. Most previous studies have found no effect of ethnicity (Armistead-Jehle, 2010; Inman et al., 1998; Lange et al., 2010; Meyers et al., 2011), although Salazar et al., (2007) reported a significant effect of ethnicity and suggested separate effort test cut-off scores for different ethnicities. The current study found that with the exception of being foreign born, ethnicity was not predictive of cognitive symptom exaggeration. These findings support previous results of no impact of ethnicity on effort and suggest that ETF is more attributable to the foreign born status of an individual than their race. There is considerable social psychology experimental and survey evidence that shows that as social distance increases and group identification decreases self-interested behavior increases and ethical/fair behavior becomes less likely (e.g., Hoffman, McCabe, Shachat, & Smith, 1996; Wenzel, 2004). These findings of reduced test-taking effort in foreign-born participants may be a reflection of this broader social psychology phenomenon.

Previous studies have shown that advancing age is associated with prolonged disability from work following illness/injury (Bruusgaard, Smeby, & Claussen, 2010; Flach, Krol, & Groothoff, 2008), and it was hypothesized that age would be associated with ETF. A relationship between age and ETF was evident when age was examined independently such that each year of age increased the odds of ETF by 1%; however age was not a significant predictor in the multivariable model.

As noted above, education was found to be inversely related to ETF, such that a one-year increase in years of education decreased the odds of ETF by 17% and this supports the majority of previous findings in this area that show that lower education is associated with greater risk of ETF (Babikian, Boone, Lu & Arnold, 2006; Greve, Etherton, Bianchini, & Curtis, 2009; Mahdavi, Mokari, & Amiri, 2011; Stulemeijer et al., 2007).

TBI severity has been previously reported to be related to ETF (Green & Iverson, 2001, Greiffenstein & Baker, 2006; West, Curtis, Greve, & Bianchini, 2011). The present study supported previous findings and showed that those with the most severe injuries are less likely than those with mTBI to display ETF. It is worth noting however, that severity of injury was one of the weakest statistical predictors of effort in the entire model and that some 11% of those with severe or extremely severe injuries displayed ETF. As such, effort testing should not be reserved only for those with mild injuries.

A limitation of this study was the archival convenience sample. This methodology risks introducing sampling bias, in this case towards the compensation-seeking population. As such caution should be adopted in extending the conclusions of this study beyond that population. Also, this sample may over-represent more severely injured and chronically disabled individuals. These findings, however, may extend generalizability beyond the malingering population of mTBI that has traditionally been examined.

Second, because New Zealand precludes litigation for damages following injury, this study cannot assess the relative importance of litigation versus compensation-seeking in predicting ETF. It may be that the adversarial nature of litigation is a factor that increases the risk of ETF. It may also be that the larger, more immediate, secondary gain that is achieved by successful litigation has relatively greater potential for increasing ETF than the temporally distant secondary gains that are seen in worker's compensation and social security claims. Such a relationship would be consistent with the findings of the behavioural economics literature on temporal discounting of rewards (e.g., Killeen, 2009). An analysis of this distinction between different forms of secondary gains would be a useful addition to the preliminary model of ETF described here.

Another limitation was the use of clinical interview alone to judge self-report of mood disorder and the presence of BFlor, without the use of formal psychometrics. Additonally, BFlor was coded retrospectively on the basis of commentary in a patient's file rather than direct observation. These methods potentially introduced bias and error variance and negatively affect the reliability and generalizability of findings. That noted, DSM diagnostic criteria were strictly adhered to in determining self-reported mood disorders, and there is no accepted operational definition of BFlor, nor are there psychometrically sound behavioral measures of illness display. Behavioral observation techniques are required to further this direction of study, and would add a new category of tools for the detection of symptom exaggeration.

Although this study employed a large sample, logistic regression techniques are demanding of a high ratio of cases to variables and this is calculated based on the smaller of the two groups (Harrell, 2001). This study employed unbalanced groups (VE versus ETF), somewhat intrinsic to the subject matter, and consequently the case-to-variable ratio is somewhat low. Diagnostic procedures have not detected any sign of marked over-fitting but a degree of over-fitting is a possibility, particularly given the relatively low numbers of participants displaying BFlor and psychotic symptomatology.

Although the predictive model was statistically and clinically significant and showed excellent discriminative power, the model did not account for all the observed variance in effort in this sample. This means that there are substantial, to-date unexamined, predictors of effort that are not being detected by this model or by the extant literature. As noted above, workplace and occupational variables present as potentially important variables for future investigation. The pain and health psychology literature has examined such variables as catastrophisation (Sullivan, Bishop, & Pivik, 1995), self-efficacy (O'Leary, 1985), external health locus of control (Torres et al, 2009), fear of pain and fear of re-injury (Waddell, 1993), as predictors of disability – similar variables may also prove important predictors of effort in the neuropsychological domain. Low resilience (Seligman & Csikszentmihalyi, 2000) has also been posited as a driver of disability following brain injury (White, Driver, & Warren, 2008).

The findings of the present study indicate that compensation-related, injury-related, demographic, psychological and behavioral factors are statistical predictors or correlates of ETF. The picture of ETF appears more complex than has been seen previously. There is a need to investigate other variables that are potentially associated with low test taker effort.

References

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev. ed.). Washington, DC: Author.

Armistead-Jehle, P. (2010). Symptom validity test performance in U.S. Veterans referred for evaluation of Mild TBI. *Applied Neuropsychology, 17*, 52-59.

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.

Babikian, T., Boone, K. B., Lu, P., & Arnold, G. (2006). Sensitivity and specificity of various Digit Span scores in the detection of suspect effort. *The Clinical Neuropsychologist, 20*, 145-159.

Belsley, D. A., Kuh, E., & Welsch, R. E. (1980). *Regression diagnostics: identifying influential data and sources of collinearity*. New York: Wiley.

Bianchini K. J., Curtis, K. L., & Greve, K. W. (2006.) Compensation and malingering in traumatic brain injury: A dose-response relationship? *The Clinical Neuropsychologist, 20*, 831-847.

Bierley, R. A., Drake, A. I., Ahmed, A., Date, E. S., Rosner, M., Warden, D., Salazar, A. M., & The Defence and Veterans Head injury Program (DVHIP) Study Group (2001). Biased responding: A case series demonstrating a relationship between somatic symptoms and impaired recognition memory performance for traumatic brain injured individuals. *Brain Injury, 15*, 697–714.

Boone, K. B., & Lu, P. H. (1999). Impact of somatoform symptomatology on credibility of cognitive performance. *The Clinical Neuropsychologist, 13*, 414–419.

Boone, K. B., Salazar, X., Lu, P., Warner-Chacon, K., & Razani, J. (2002). The Rey 15-Item Recognition Trial: A technique to enhance sensitivity of the Rey 15-Item

Memorization Test. *Journal of Clinical and Experimental Neuropsychology, 24*, 561–573.

Bruusgaard, D., Smeby, L., & Claussen, B. (2010). Education and disability pension: A stronger association than previously found. *Scandinavian Journal of Public Health, 38*, 686–690.

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*, 960–971.

Carroll, L. J., Cassidy, D., Holm, L., Kraus, J., & Coronado, V. G. (2004). Methodological issues and research recommendations for mild traumatic brain injury: The WHO Collaborating Centre Task Force on Mild Traumatic Brain Injury. *Journal of Rehabilitation Medicine, 43*(Suppl.), 113–125.

Chaftez, M., & Prentkowski, E. (2011). A case of malingering by proxy in a social security disability psychological consultative examination. *Applied Neuropsychology, 18*, 143-149.

Dandachi-Fitzgerald, B., Pronds, R. W. H. M., Peters, M. J. V, & Merckelbach, H. (2011). Cognitive underperformance and symptom over-reporting in a mixed psychiatric sample. *The Clinical Neuropsychologist, 25*, 812-828.

Demakis, G. J., Gervais, R. O., & Rohling, M. L. (2008). The effect of failure on cognitive and psychological symptom validity tests in litigants with symptoms of post-traumatic stress disorder. *The Clinical Neuropsychologist, 22*, 879-895.

Donders, J., & Boonstra, T. (2007). Correlates of invalid neuropsychological test performance after traumatic brain injury. *Brain Injury, 21*, 319–326.

Dragano, N., & Schneider, L. (2011). Work Related Psychosocial Factors and the Risk of Early Disability Pensioning: A Contribution to Assessing the Need for Rehabilitation.

Rehabilitation, 50, 28-36.

Ekman, P., & O'Sullivan, M. (2006). From flawed self-assessment to blatant whoppers: The utility of voluntary and involuntary behavior in detecting deception. *Behavioral Sciences and the Law, 24,* 673–686.

Faul, M., Xu, L., Wald, M.M., & Coronado, V.G. (2010). *Traumatic brain injury in the United States: Emergency department visits, hospitalizations and deaths 2002-2006.* Atlanta, GA: Centers for Disease Control and Prevention.

Flach, P. A., Krol, B., & Groothoff, J. W. (2008). Determinants of sick-leave duration: A tool for managers? *Scandinavian Journal of Public Health, 36*, 713–719.

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 (Ed.), *Assessment of feigned cognitive impairment* (pp. 281-309). New York: The Guilford Press.

Gorissen, M., Sanz, J. C., & Schmand, B. (2005). Effort and cognition in schizophrenia patients. *Schizophrenia Research, 78*, 199-208.

Green, P. (2003). Green's Word Memory Test for windows user's manual. Edmonton, Alberta, Canada: Green's Publishing.

Green, P. (2009). Questioning common assumptions about depression. In J. E. Morgan & J. J. Sweet (Eds.), *Neuropsychology of malingering casebook* (pp. 132-144). New York: Psychology Press.

Green, P., & Iverson, G. L. (2001). Validation of the Computerized Assessment of Response Bias in litigating patients with head injuries. *The Clinical Neuropsychologist, 15*, 492-497.

Green, P., Iverson, G. L., & Allen, L. M. (1999). Detecting malingering in head injury litigation with the Word Memory Test. *Brain Injury, 13,* 813-819.

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*, 1045–1060.

Greiffenstein, M. F., & Baker, W. J. (2006). Miller was (mostly) right: Head injury severity inversely related to simulation. *Legal and Criminological Psychology, 11*, 131-145.

Greiffenstein, M. F., Baker, W. J., & Gola, T. (1994). Validation of malingered amnesia measures with a large clinical sample. *Psychological Assessment, 6*, 218-224.

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*, 1176–1190.

Greve, K.W., Etherton, J.L., Ord, J., Bianchini, K.J., & Curtis, K.L. (2009). Detecting malingered pain-related disability: Classification accuracy of the Test of Memory malingering. *The Clinical Neuropsychologist, 23*, 1250–1271.

Gronwall, D., & Wrightson, P. (1980). Duration of post-traumatic amnesia after mild head injury. *Journal of Clinical Neuropsychology, 2*, 51-60.

Grote, C. L., Kooker, E. K., Garron, D. C., Nyehuis, 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*, 709-719.

Hardin, J. W. (1995). Variance inflation factors and variance-decomposition proportions. *Stata Technical Bulletin* (STB 24). Retrieved from http://www.stata.com/bookstore/individual-stata-technical-bulletin-issues/

Harrell, F. E. (2001). Regression modeling strategies: With applications to linear models, logistic regression, and survival analysis. New York: Springer-Verlag.

Henry, G. K., Heilbronner, R. L., Mittenberg, W., Enders, C., Stevens, A., & Dux, M. (2011). Noncredible performance individuals with external incentives: Empirical derivation and cross-validation of the Psychosocial Distress Scale (PDS). *Applied Neuropsychology, 18,* 47-53.

Hoffman, E., McCabe, K., Shachat, K., & Smith, V. (1996). Social distance and other regarding behavior in dictator games. *American Economic Review, 86*, 653–660.

Hosmer, D. W., and Lemeshow, S. (2000). *Applied logistic regression* (2nd ed.). John Wiley & Sons, Inc: New York.

Inman, T. H., Vickery, C. D., Berry, D. T. R., Lamb, D. G., Edwards, C. L., & Smith, G. T. (1998). Development and initial validation of a new procedure for evaluating adequacy of effort given during neuropsychological testing: The Letter Memory Test. *Psychological Assessment, 10,* 128–139.

Jones, A., & Ingram, M. V. (2011). A Comparison of Selected MMPI-2 and MMPI-2-RF Validity Scales in Assessing Effort on Cognitive Tests in a Military Sample. *The Clinical Neuropsychologist*, *25*, 1207-1227.

Jones, A., Ingram, M. V., & Ben-Porath, Y. (2012). Scores on the MMPI-2 RF scales as a function of increasing levels of failure on cognitive symptom validity tests in a military sample. *The Clinical Neuropsychologist*, *26,* 790-815.

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,* 904–911.

Killeen, P. R. (2009). An addictive-utility model of delay discounting. *Psychological Review, 116*, 602-619.

Krokstad, S., Johnsen, R., & Westin, S. (2002). Social determinants of disability pension: a 10-year follow-up of 62 000 people in a Norwegian county population. *International Journal of Epidemiology, 31*, 1183–1191.

Kuoppala, J., Lamminpaa, A., Vaananen-Tomppo, I., & Hinkka, K. (2011). Employee well-being and sick leave, occupational accident, and disability pension a cohort study of civil servants. *Journal of Occupational and Environmental Medicine, 53*, 633-640**.**

Lange, R. T., Iverson, G. L., Brooks, B. L., & Rennison, V. 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,* 961–972.

Larrabee, G. J. (2003a). Detection of malingering using atypical performance patterns on standard neuropsychological tests. *The Clinical Neuropsychologist, 17,* 410-425.

Larrabee, G.J. (2003b). Exaggerated MMPI-2 symptom report in personal injury litigants with malingered neurocognitive deficit. *Archives of Neuropsychology, 18,* 673-686.

Larrabee, G. J. (2007). Malingering, research designs, and base rates. In G.J Larrabee (Ed.), *Assessment of malingered neuropsychological deficits* (pp. 3-13). Oxford: Oxford University Press.

Larrabee, G. J. Greiffenstein, M.F., Greve, K.W., & Bianchini, K. (2007). Refining diagnostic criteria for malingering. In G.J Larrabee (Ed.), *Assessment of malingered neuropsychological deficits* (pp. 334-371). Oxford: Oxford University Press.

Mahdavi, M-E, Mokari, N., & Amiri, Z. (2011). Educational level and pseudohypacusis in medico-lega compensation claims: a retrospective study. *Archives of Iranian Medicine, 14*, 58-60.

Mason, S., Wardrope, J., Turpin, G., & Rowlands, A. (2002). Outcomes after injury: A comparison of workplace and nonworkplace injury. *The Journal of Trauma: Injury, Infection, and Critical Care, 53,* 98-103.

Meyers, J.E., Volbrecht, M., Axelrod, B.N., & Reinsch-Boothby, L. (2011). Symptom validity tests and overall neuropsychological test performance. *Archives Of Clinical Neuropsychology, 26*, 8-15.

Miller, H. (1961a). Accident neurosis: Lecture I. *British Medical Journal, 1*(5230), 919-925.

Miller, H. (1961b). Accident neurosis: Lecture II. *British Medical Journal, 1*(5231), 992-998.

Miller, H., & Cartlidge, N. (1972). Simulation and malingering after injuries to the brain and spinal cord. *The Lancet, 299*, 580-585.

Mooney, G., Speed, J., & Sheppard, S. (2005). Factors related to recovery after mild traumatic brain injury. *Brain Injury, 19,* 975–987.

Moore, B. A., & Donders, J. (2004). Predictors of invalid neuropsychological test performance after traumatic brain injury. *Brain Injury*, *18*, 975–984.

Nelson, N. W., Sweet, J. J., Berry, D. T., Bryant, F. B., & Granacher, R. P. (2007). Response validity in forensic neuropsychology: Exploratory factor analytic evidence of distinct cognitive and psychological constructs. *Journal of the International Neuropsychological Society*, 13, 440–449.

O'Leary, A. (1985). Self-efficacy and health. *Behaviour Research and Therapy*, *23*, 437-451.

Ord, J. S., Boettcher, A. C., Greve, K. W., & Bianchini, K. J. (2010). Detection of malingering in mild traumatic brain injury with the Conners' Continuous Performance Test–II. *Journal of Clinical and Experimental Neuropsychology, 32*, 380–387.

Paniak, C., Reynolds, S., Toller-Lobe, G., Melnyk, A., Nagy, J., & Schmidt, D. (2002). A longitudinal study of the relationship between financial compensation and symptoms after treated mild traumatic brain injury. *Journal of Clinical and Experimental Neuropsychology, 24,* 187–193.

Rees, L. M., Tombaugh, T. N., & Boulay, L. (2001). Depression and the Test of Memory Malingering. *Archives of Clinical Neuropsychology, 16,* 501-506.

Rey, A. (1964). *L'examen Clinique en Psychologie.* Paris: Presses Universitairs de France.

Rohling, M. L., Green, P., Allen, L. M., & Iverson, G. L. (2002). Depressive symptoms and neurocognitive test scores in patients passing symptom validity tests. *Archives of Clinical Neuropsychology, 17,* 205-222.

Ross, S. R., Putnam, S. H., Millis, S. R., Adams, K. M., & Krukowski, R. A. (2006). Detecting insufficient effort using the Seashore Rhythm and Speech-Sounds Perception Tests in head injury. *The Clinical Neuropsychologist, 20,* 798–815.

Ruff, R. M., Iverson, G. L., Barth, J. T., Bush, S. S., Broshek, D. K., & the NAN Policy and Planning Committee. (2009). Recommendations for diagnosing a mild traumatic brain injury: A National Academy of Neuropsychology education paper. *Archives of Clinical Neuropsychology, 24,* 3–10.

Ruocco, A.C., Swirsky-Sacchetti, T., Daines, D., Chute, D.I., Platek, S.M., Mandel, S., & Zillmer, E.A. (2005). Dissociation between neuropsychological and psychiatric forms of malingering [abstract]. *Archives of Clinical Neuropsychology, 20*, 883-883.

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 (Ed.), *Assessment of feigned cognitive impairment* (pp. 405-427). New York: The Guilford Press.

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*, 437–453.

Seligman, M. E. P., & Csikszentmihalyi, M. (2000). Positive psychology: An introduction. *American Psychologist, 55,* 5–14.

Shores, E. A., Marosszeky, J. E., Sandanam, J., & Batchelor, J. (1986) Preliminary validation of a clinical scale for measuring the duration of post-traumatic amnesia. *Medical Journal of Australia*, *144*, 569-572.

Slick, D. J., Sherman, E. M. S., & Iverson, G. L. (1999). Diagnostic criteria for malingered neurocognitive dysfunction: Proposed standards for clinical practice and research. *Clinical Neuropsychologist*, *13*, 545–561.

Statacorp (2011). *Stata Release 12*. Statistical Software. College Station, TX: Statacorp LP.

Stulemeijer, M., Andriessen, T. M. J. C., & Brauer, J. M. P., Vos, P. E., Van der Werf, S. (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, 309-318.

Suhr, J. A. & Barrash, J. (2007). Performance on standard attention, memory, and psychomotor speed tasks as indicators of malingering. In G. J. Larrabee (Ed.), *Assessment of malingered neuropsychological deficits* (pp. 334-371). Oxford: Oxford University Press.

Suhr, J., Tranel, D., Wefel, J., & Barrash, J. (1997). Memory performance after head injury: Contributions of malingering, litigation status, psychological factors, and medication use. *Journal of Clinical and Experimental Neuropsychology, 19*, 500–514.

Sullivan, M. J. L., Bishop, S. R., & Pivik, J. (1995). The Pain Catastrophizing Scale: Development and validation. *Psychological Assessment, 7*, 524-532.

Sumanti, M., Boone, K. B., Savodnik, I., & Gorsuch, R. (2006). Noncredible psychiatric and cognitive symptoms in a worker's compensation "stress" claim sample. *The Clinical Neuropsychologist, 20*, 754-765.

Teasdale, G. M. (1995). Head Injury. *Journal of Neurology, Neurosurgery and Psychiatry, 58*, 526-539.

Teasdale, G. M., & Jennett, B. (1974). Assessment of coma and impaired consciousness: A practical scale. *The Lancet*, *304,* 81-84.

Tombaugh, T. N. (1996). *Test of Memory Malingering (TOMM).* North Tonawanda, NY: Multi-Health Systems.

Tombaugh, T. N. (1997). The Test of Memory Malingering (TOMM): Normative data from cognitively intact and cognitively impaired individuals. *Psychological Assessment, 9,* 260-268.

Torres, V., Collado, A., Arias, A., Peri, J. M., Bailles, E., Salamero, M., & Valdes, M. (2009). Pain locus of control predicts return to work among Spanish fibromyalgia patients after completion of a multidisciplinary pain program. *General Hospital Psychiatry, 31*, 137–145.

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*, 297-313.

Waddell, G. (1993). A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low-back pain and disability. *Pain, 52,* 157-168.

Wenzel, M. (2004). An analysis of norm processes in tax compliance. *Journal of Economic Psychology, 25*, 213–228.

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*, 114-125.

White, B., Driver, S., & Warren, A. (2008). Considering resilience in the rehabilitation of people with traumatic disabilities. *Rehabilitation Psychology, 53*, 9-17.

Whiteside, D. M., Dunbar-Mayer, P., & Waters, D. P. (2009). Relationship between TOMM performance and PAI validity scales in a mixed clinical sample. *The Clinical Neuropsychologist*, *23*, 523-533.

Williams, D. H., Levin, H.S., & Eisenberg, H. M. (1990). Mild head injury classification. *Neurosurgery*, *27*, 422-428.

CHAPTER 6: PSYCHOLOGICAL PREDICTORS OF EFFORT TEST FAILURE

Paper prepared for submission to the journal *The Clinical Neuropsychologist*

Abstract

Psychological and personal history predictors of effort test failure (ETF) were examined in an archival sample of 540 traumatically brain injured (TBI) adults. Logistic regression analyses were undertaken to examine potential predictors of ETF including: current diagnosis with major depressive disorder, current diagnosis with an anxiety disorder, having a current pain disorder, and having a current substance use disorder. Historical predictors examined included: having a premorbid psychiatric history, having a childhood history of sexual or physical abuse, and having a history of substance abuse. In each case, the predictive power of variables was adjusted for variables previously established to be of importance (Webb et al., 2012) namely: years of education, being foreign born, severity of injury, being compensation-seeking, and having sustained a workplace accident (Webb et al. 2012). Results showed that only a current diagnosis of depressive disorder was predictive of ETF (OR: 4.55, 95% CI [2.56, 8.08]) once adjustment was made for the effects of the variables listed above.

A subset of participants completed mood measures (Beck Depression Inventory-2; BDI-2; n = 207; State Trait Anxiety Inventory; STAI; n = 90). An analysis of group differences showed that scores on the BDI-2 significantly discriminated those participants giving valid effort from those displaying ETF (t (171) = 5.90, $p < .001$, d = .91) as did scores on the STAI-State subscale (U = 798, p = .01, r = .36) and scores on the STAI-Trait subscale (U = 826, p = .001, r = .40). Effects of the STAI were smaller than that seen with the BDI-2. Further regression analyses suggested that the STAI likely gained its discriminative power from compensation-seeking status.

Results indicated that depressive symptom reporting is significantly and strongly associated with ETF and vigilance to low effort is particularly indicated when confronted by displays of depressive symptomatology during neuropsychological assessment. Limitations of the study and suggestions for further research are identified.

Around three million cases of TBI occur each year in the United States and rates of TBI in most developed industrialised nations are similar, being estimated at 175-200 per 100,000 (Granacher, 2008). Most of those injuries represent mild traumatic brain injury (mTBI) (Ruff, 2011) from which the majority of people fully recover within weeks to months of the injury (e.g., Belanger, Curtiss, Demery, Lebowitz, & Vanderploeg, 2005; Carroll et al., 2004; Schretlen & Shapiro, 2003). However, slow and incomplete recovery is noted in a minority of mTBI patients. Poor outcome from TBI has been attributed to severity of injury and associated neuropathology (Bigler, 2008; De Guise, Le Blanc, Feyz, & Lamoureux, 2006; LeBlanc, de Guise, Gosselin, & Feyz, 2006), however, non-injury variables including litigation/compensation-seeking status have also been reliably found to predict poor recovery from TBI (Belanger et al., 2005; Carroll et al., 2004).

Iverson (2005) reviewed the literature on variables that are associated with poor recovery from mTBI and suggested that a number of pre-existing psychological variables, including psychiatric conditions and substance abuse problems, could be associated with poor outcomes. He also reported that co-morbid conditions, such as chronic pain, affective disturbance, or substance abuse are associated with poor outcome from mTBI. Whether those factors are also associated with effort test failure (ETF) or malingering has received little attention.

Pain has been examined as a predictor of ETF. Although experimental laboratory studies have not supported the notion that pain is a predictor of low effort (Etherton, Bianchini, Ciota, & Greve, 2005; Etherton, Bianchini, Greve, & Ciota, 2005), studies employing clinical samples of chronic pain patients have found that cognitive test-taking effort is reduced in this population (Greiffenstein, Gervais, Baker, Artiola & Smith, 2013; Greve, Etherton, Ord, Bianchini, & Curtis, 2009; Greve, Ord, Curtis, Bianchini, & Brennan, 2008; Johnson-Greene, Brooks, & Ference, 2013). Mooney, Speed, and Sheppard (2005) also found that pain was a predictor of poor outcome from TBI. To our knowledge there have been no previous studies that have examined whether having a co-morbid pain disorder is a specific risk factor for low effort in a TBI sample.

Depression and anxiety have been examined as predictive of poor effort in the TBI population. Findings have varied with some reporting a relationship between affective disturbance and effort (Gorissen, Sanz & Schmand, 2005) and others finding that no such relationship exists (Goldberg, Back-Madruga & Boone, 2007). To date, six studies have examined the relationship between affective disturbance and effort in a cohort restricted to TBI patients (Armistead-Jehle, 2010; Lange, Iverson, Brooks, & Rennison, 2010; Locke, Smigielski, Powell, & Stevens, 2008; Stulemeijer, Andriessen, Brauer, Vos, & Van der Werf, 2007; Suhr, Tranel, Wefel, & Barrash, 1997; Thomas & Youngjohn, 2009). Four groups of authors reported a relationship between depressed mood and effort (Armistead-Jehle, 2010; Lange et al., 2010; Stulemeijer et al., 2007; Thomas & Youngjohn, 2009) with the remaining two demonstrating no such effect (Locke et al., 2008; Suhr et al., 1997).

To date, only a limited analysis of the relationship between anxiety and effort has been undertaken in the TBI population. Lange et al., (2010) examined endorsement of anxiety symptoms on the Post-Concussion Scale (Lovell et al., 2006) and found that those failing the Test of Memory Malingering (TOMM; Tombaugh, 1996) were more likely to endorse nervousness (d = 1.12) than those passing the TOMM. In contrast, Locke et al., (2008) found that their anxious cohort were no more likely to fail Trial 2 of the TOMM than the non-anxious cohort. Stulemeijer et al., (2007) found that the mean score on a measure of posttraumatic stress symptoms was higher in those who had displayed inadequate effort on the Amsterdam Short Term Memory Test (Schagen, Schmand, Sterke, & Lindeboom, 1997) than those who had displayed adequate effort ($p < .05$) but found no difference between the groups on another measure of anxiety. Inconsistencies between the results of previous studies suggest that further investigation of the relationship between the variables of depressive symptomatology and anxiety and effort is indicated in the TBI population.

TBI and substance abuse are recognised as having a high rate of co-morbidity (Sacks et al., 2009). If substance abuse is associated with poor recovery from TBI (Iverson, 2005; Kirsch et al., 2010), it is possible that substance abuse is associated with an increased risk of affording a suboptimal effort during neuropsychological testing. Iverson, Slick, & Franzen (2000) examined a non-litigating cohort of people receiving inpatient treatment for substance abuse and found no significant evidence of poor effort on the Wechsler Memory Scale – Revised Malingering Index, a finding that was essentially replicated by Miller, Ryan, Carruthers, and Cluff (2004), using embedded Wechsler effort indices. Having a substance abuse history was also examined by Moore and Donders (2004) in their investigation of effort in a TBI sample and was not found to contribute significantly to a model predictive of invalid performance on the California Verbal Learning Test-II or the TOMM. Locke et al. (2008) found no evidence that having a history of substance abuse was associated with poor TOMM effort in a cohort of TBI rehabilitation seekers. Although a positive history of substance abuse has not been found to be predictive of low effort in TBI samples, whether having a co-morbid substance abuse problem is a factor that predicts cognitive effort has not been examined.

ETF has been found to be associated with other psychological variables. Moore and Donders (2004) examined the predictors of low effort in a sample of 132 TBI patients. They found that, after controlling for compensation-seeking, having a psychiatric history was predictive of low effort, a finding supported by Stulemeijer and colleagues (2007). Donders and Boonstra (2007) found that having either a psychiatric history or a history of physical or sexual abuse was predictive of suboptimal effort in a sample of 87 TBI participants. Similarly, Williamson, Holsman, Chaytor, Miller, and Drane (2012) found that having a physical, emotional, or sexual abuse history was related to ETF in a sample of people displaying psychogenic nonepileptic seizures.

Webb et al. (2012) revealed that both compensation-seeking and a concurrent self-reported Diagnostic and Statistical Manual-IV-TR (DSM-IV-TR) diagnosed mood disorder were independently predictive of low test-taking effort. The aim of the current study was to undertake further analysis of the relationship between psychological factors and effort and specifically, to examine whether current self-reported depression or anxiety was associated with ETF, whether those displaying ETF scored more highly on self-report measures of depression and anxiety than those displaying valid effort (VE) , whether having a co-morbid substance use disorder was predictive of ETF, whether having a co-morbid pain disorder was predictive of ETF, and whether having a psychiatric history, a history of sexual or physical abuse or a history of substance abuse was predictive of ETF.

Method

Participants

The sample consisted of 540 traumatically brain injured adults aged 16 years and above. The injury characteristics and demographic breakdown of the sample has been described in detail in Webb et al., (2012) and full details will not be repeated here. The data set was collected over the period 2001 to 2007 inclusive. Participants were excluded on the basis of having a pre-existing history of mental retardation, pre-existing or post-injury psychotic illness, or dementia.

Demographic variables.

Males made up 72% of the sample. The mean age was 41 years (*SD* = 12.40, range = 16-76). Ethnicity was as follows: European/White; (*n* = 409, 75.7%), Maori/Pacific Island (*n* =

101, 18.7%), Indo/Asian (*n* = 30, 5.6%). The sample had, on average, 11.8 (*SD* = 2.4, range = 6-21) years of education.

Injury variables

Injury variables and the method employed for categorizing injury severity has been previously described in detail (Webb et al., 2012). In brief, severity of injury was based on the acute Glasgow Coma Scale scores in the Emergency Department, duration of loss of consciousness, and duration of posttraumatic amnesia. TBI severity was classified as shown in Table 1. The mTBI-complicated and Moderate TBI groups were combined as were the Very Severe TBI and Extremely Severe TBI groups to ensure that parameter estimates were based on adequate numbers of cases and that there were no empty cells.

Table 1

Injury Severity Criteria and Sample Characteristics

Descriptor	Criteria	*n* (%)
mTBI	LOC < 30 mins, PTA < 24 hours, GCS 13-15 at 30 mins	272 (50.4)
mTBI complicated	LOC <30 mins, PTA <24 hours, GCS 13-15 at 30 mins, visible (on CT brain imaging) intracranial abnormality not requiring surgery	38 (7.0)
Moderate	PTA 1-24 hours, GCS 9-12	49 (9.1)
Severe	PTA 1-7 days, GCS 3-8	73 (13.5)
Very severe	PTA 1-4 weeks, GCS 3-8	69 (12.8)
Extremely severe	PTA >4 weeks, GCS 3-8	39 (7.2)

Note. LOC = Loss of consciousness; mTBI = mild traumatic brain injury; PTA = posttraumatic amnesia; GCS = Glasgow Coma Scale.

Procedures

This study was approved by the Human Research Ethics Committees of the University of New England and Macquarie University.

Measures.

All participants completed three measures of effort including one forced-choice symptom validity test. All participants completed the Test of Memory Malingering (TOMM; Tombaugh, 1996), the Fifteen Item Test (FIT; Rey, 1964), and Reliable Digit Span (RDS; Greiffenstein, Baker, & Gola, 1994).

From 2005 increased funding availability provided for the ability to introduce self-report affect measures. From 2005 all participants (*n* = 207) completed the Beck Depression Inventory-II (Beck, Steer, & Brown, 1996). The Beck Depression Inventory-II (BDI-2) is the most widely used self-report measure of depressed affect (Farmer, 2012), assessing depressive symptomatology and severity. A subset of those participants (*n* = 90) who displayed clinically significant anxiety also completed the State Trait Anxiety Inventory (Spielberger, 1983). The State Trait Anxiety Inventory is the most thoroughly researched measure of anxiety (Dreger, 2012), and provides a measure of anxiety in the transient state (STAI-s) and anxiety as a more stable and enduring personality characteristic or trait (STAI-t).

Effort classification criteria.

Effort was classified dichotomously – Valid Effort (VE) or ETF. Cut-offs employed for each measure were set to ensure 90% specificity following Baker, Donders, and Thompson (2000) and Boone, Salazar, Lu, Warner-Chacon, and Razani (2002). Consequently the Greve, Bianchini and Doane (2006) cut-off of <47 was employed for TOMM2 and TOMM Retention; An RDS cutoff of <8 was employed according to guidelines of Suhr and Barrash's (2007) review; An FIT cutoff of <9 was used in accordance with the findings of Boone et al.

(2002). As in Webb et al. (2012), ETF was operationalized as failure on any two measures in accord with the findings of Victor, Boone, Serpa, Buehler, and Ziegler (2009) or below chance performance on either trial 2 of TOMM or the retention trial (<18/50 at the 95% confidence interval using the binomial distribution). VE was classified using the requirement that participants pass all three effort measures at the cutoffs described above.

Psychological dimensions.

Psychological data were available from the archive records for each participant. All participants had undergone a semi-structured clinical interview of approximately 60 minutes duration with a licensed psychologist trained in clinical psychology (JW). Participants reporting affective disturbance and who met the DSM-IV-TR (American Psychiatric Association, 2000) for a depressive disorder or anxiety disorder were coded as having a self-reported depressive disorder or self-reported anxiety disorder. Individuals reporting specific phobias were not coded positive for anxiety disorder unless the phobia was a focus of treatment (eg, driving phobia following motor vehicle accident).

Self-reported pain disorders and substance use disorders were also diagnosed according to the DSM-IV-TR diagnostic criteria. Patients were coded positive for a current and/or historical substance use disorder if they reported sufficient current and/or historical symptoms for a DSM diagnosis with substance dependence and/or substance abuse.

All participants were screened via interview for a past history of psychiatric disorder and a history of childhood physical and/or sexual abuse. Psychiatric history was categorized as positive in any participant reporting a pre-injury history of having sought any individual psychiatric or psychological treatment, including any history of family physician prescribed psychotropic medication for the purposes of psychological management.

Physical abuse history was categorized as positive in any individual case upon endorsement of the question: “Were you ever physically abused as a child?” When necessary this question was clarified by further questioning: whether they were hit with a belt, stick or other object and whether the patient considered that the punishment was fair and reasonable discipline or excessive and harsh. For the purposes of this screening assessment physical abuse was dichotomized as endorsed/not endorsed and severity and frequency of abuse were not entered as variables for analysis.

Sexual abuse exposure was screened similarly: “Were you ever sexually abused as a child?” Where necessary this question was clarified by questioning whether a patient was involved in non-consensual sexual activity before the legal age of consent (16 years). For the purposes of this screening, sexual abuse was dichotomised endorsed/not endorsed and other variables such as age of abuse, frequency of abuse, familial/non-familial abuse were not entered for analysis.

Retrospective adult reports of childhood abuse and neglect have been shown to have good test-retest reliability and to be robust to affective states (Dube, Williams, Thompson, Felitti, & Anda, 2004; Yancura & Aldwin, 2009) and as such, self-reporting in this screening manner was likely to have resulted in reliable reporting.

Results

Of the entire cohort 25.7% (*n* = 139) were diagnosed with a depressive disorder, 13.1% (*n* = 71) with an anxiety disorder, 5.9% (*n* = 32) were diagnosed with a pain disorder and 5.7% (*n* = 31) with a substance abuse disorder. A total of 22.8% (*n* = 123) reported having a pre-injury psychiatric history, 17.4% (*n* = 94) reported a history of childhood physical abuse, 10.6% (*n* = 57) reported a history of childhood sexual abuse, and 19.8% reported a history of substance abuse.

The psychological variables of interest were found to be significantly intercorrelated (Table 2).

Table 2

Correlations of psychological variables, ETF and Compensation-Seeking

Variable	Anxiety	Pain dis.	Subst. use disorder	BDI-2	STAI-s	STAI-t	Psych hx	Sexual abuse hx	Physical abuse hx	Subst. use dis. hx	Comp. seeking	ETF
Depression	.47***	.06	-.30	.65***	.44***	.47***	.45***	.50***	.22	.21	.10	.68***
Anxiety		-.20	.01	.14*	.50***	.46***	.20	.50**	.27	-.06	-.06	.45**
Pain dis			.44	-.01	.10	.13	-.24	-.07	-.07	-.28	.29	.01
Subst. use disorder				.12	.12	-.04	.09	.12	.53*	.91***	.12	-.25
BDI-2					.62***	.68***	.12	.13	.13	.12	.08	.41***
STAI-s						.87***	.17	.22*	.20	-.04	.29**	.35**
STAI-t							.19	.24*	.22*	.07	.34**	.32**
Psych hx								.65***	.55***	.29*	-.03	-.04
Sexual abuse hx									.65***	.30	.15	.38*
Physical abuse hx										.44**	-.12	.07
Subst. use dis. hx											.03	-.21
Comp. seeking												.65***

Note. Relationships among dichotomous variables are expressed as Gamma coefficients. Relationships between continuous/ordinal variables and dichotomous variables are expressed as Pearson product-moment correlation coefficients/point-biserial correlation coefficients. * $p < .05$. ** $p < .01$. *** $p < .001$. Depression = depressive disorder. Anxiety = anxiety disorder. Pain dis. = pain disorder. Subst. use disorder = substance use disorder. BDI-2 = Beck Depression Inventory-2. STAI-s = State Trait Anxiety Inventory state scale. STAI-t = State Trait Anxiety Inventory trait scale. hx = history. Subst. use dis. hx = Substance use disorder history. Comp. seeking = compensation-seeking. ETF = effort test failure.

Diagnosis with a depressive disorder was significantly correlated with being diagnosed with an anxiety disorder, with BDI-2, STAI-s, STAI-t, with having a sexual abuse history, and with ETF. Diagnosis with an anxiety disorder was significantly correlated with ETF, having a sexual abuse history and with scores on both measures of the STAI but not with the BDI-2. Having a current substance use disorder significantly correlated with having a substance abuse history and with having a physical abuse history. Having a physical abuse history correlated significantly with STAI-t, with having a sexual abuse history, having a substance use disorder, and having a psychiatric history. Having a psychiatric history correlated with a diagnosis of depressive disorder and with having physical abuse, sexual abuse and substance use disorder histories.

The BDI-2 correlated significantly with a diagnosis of a depressive disorder and with a diagnosis of an anxiety disorder, with both measures of anxiety, with ETF, but not with any of the personal history variables. Both state and trait measures of anxiety were highly correlated. Both correlated with having a sexual abuse history, a substance abuse history, with being compensation-seeking, and with ETF. ETF was significantly correlated with a diagnosis of major depressive disorder, with a diagnosis of anxiety disorder, with both the BDI-2 and STAI measures, with sexual abuse history, and with compensation-seeking status.

Group differences

Table 3 shows the mean scores on the three psychological measures by level of effort (VE versus ETF). For the purposes of statistical analysis, alpha was adjusted employing a Bonferonni correction to .02. BDI-2 data was sufficiently normally distributed (Skew = .59), and a Levene's test showed essentially equal variances ($F = .03$, $p = .86$), to allow a parametric analysis of the BDI-2 differences between the means of the VE and ETF groups.

A significant difference between the VE and ETF groups was evident on the BDI-2 (t (171) = 5.90, p < .001, d = .91) and the effect size was large (Cohen, 1988).

Table 3

Mean (SD) psychological measure results for the adequate effort and inadequate effort groups

	VE	ETF	*p*	*Effect size* (*d*, *r*)
	(*n* = 104)	(*n* = 69)		
BDI-2	15.45 (11.68)	26.38 (12.29)	<.001	.91
	(*n* = 43)	(*n* = 26)		
STAI-s	62.86 (32.84)	85.08 (23.05)	.003	.36
STAI-t	71.53 (27.62)	89.12 (20.40)	.001	.40

Note. Effect size for t-test = *Cohen's d;* Effect size for Mann-Whitney U test = *r*.

STAI-s and STAI-t data was significantly skewed (Skew = -1.16 and -1.50 respectively) and a Levene's tests of equality of variances showed the variances of the groups were unequal (F = 10.18, p = .002, and F = 9.38, p = .003, respectively). As such the Mann-Whitney U test was employed for both anxiety measures. The results showed that for the STAI-s, state anxiety was significantly higher in the low effort group (U = 798, p = .01, r = .36). Reported trait anxiety was also significantly higher in the ETF group than the VE group (U = 826, p = .001, r = .40).

Predictive analyses

The power of the psychological variables to predict ETF was examined using logistic regression. Given the evident collinearities among psychological and personal history predictors and the large number of possible predictors to cases, it was not possible to evaluate a predictive model employing all psychological variables.

The power of variables to predict ETF was examined unadjusted and adjusted for known predictors of low effort (Webb et al., 2012) (with the exception of having a diagnosed (self-reported) mood disorder and self-reported psychotic symptomatology), namely: years of education, being foreign born, displaying behavioural floridity (defined as extreme displays of symptom-related behavior), TBI injury severity, having had a workplace accident, and being compensation-seeking. In light of multiple comparisons being undertaken and the increased risk of a Type I error, significance levels were subjected to a Bonferroni adjustment where the alpha level was set at .005

Considering the unadjusted odds, having a diagnosed depressive disorder and an anxiety disorder were significant predictors of low effort, such that a diagnosis of major depressive disorder was associated with about five times the odds of ETF relative to those not receiving such a diagnosis. Odds of ETF were around 2.5 times higher in those diagnosed with an anxiety disorder relative to those not receiving such a diagnosis. Having a co-morbid pain disorder or substance use disorder was not predictive of ETF, nor was reporting a psychiatric history, a physical abuse history, or a history of substance use disorder. Reporting a sexual abuse history trended toward an association with ETF (p = .02) but that relationship was not significant after correcting for multiple comparisons.

After adjusting for the known predictors of ETF, only having a diagnosis of depressive disorder significantly predicted ETF, such that the odds of ETF increased around 4.5 times for those diagnosed with depression relative to those without the diagnosis. Again, having a sexual abuse history trended toward being a significant predictor, with the odds of ETF approximately doubling for those reporting a sexual abuse history compared to those without that reported history, but the finding was not significant once the alpha level was corrected for multiple comparisons.

Table 4

Logistic Regression Analysis of ETF as a Function of Psychological Variables

Variable	*n* (%) with characteristic	Adjusted OR[1] (95% CI)	*p* value	Unadjusted OR (95% CI)	*p* value
Diagnosis[1]					
Depressive disorder	139 (25.7)	4.55 [2.56, 8.08]	<.001	5.17 [3.21, 8.32]	<.001
Anxiety disorder	71 (13.1)	1.66 [.78, 3.55]	.19	2.66 [1.50, 4.72]	.001
Pain disorder	32 (5.9)	.67 [.23, 1.95]	.59	1.02 [0.40, 2.62]	.96
Substance use disorder	31 (5.7)	.72 [2.0, 2.60]	.33	.60 [0.20, 1.79]	.36
Personal History[1]					
Psychiatric hx	123 (22.8)	.997 [.52, 1.90]	.74	.92 [0.54, 1.57]	.75
Sexual abuse hx	57 (10.6)	2.38 [1.08, 5.25]	.03	2.22 [1.16, 4.23]	.02
Physical abuse hx	94 (17.4)	1.27 [.64, 2.55]	.50	1.15 [0.64, 2.07]	.64
Substance abuse history	107 (19.8)	.82 [.41, 1.67]	.82	.66 [0.36, 1.21]	.18
Psychometric data[2]					
BDI-2	173	1.08 [1.05, 1.11]	<.001	1.08 [1.05, 1.11]	<.001
STAI-s	69	1.03 [1.005,1.05]	.02	1.03 [1.01, 1.05]	.008
STAI-t	69	1.03 [1.001,1.06]	.04	1.04 [1.001, 1.06]	.02

Note: 1. All variables were adjusted for years of education, being foreign born, compensation-seeking status, BFlor, TBI severity, and having a workplace accident. OR = Odds Ratio. hx = history. 2. Variables adjusted for compensation-seeking status only.

The power of the psychometric data to predict ETF was examined in unadjusted and adjusted analyses. The smaller number of cases and the large number of controlling variables made it inappropriate to control for all known predictors of ETF. Compensation-seeking status was entered alone as a controlling variable because the literature has established that there is a clear relationship between compensation-seeking and ETF (e.g., Bianchini, Curtis,

& Greve, 2006, Henry et al., 2011; Paniak et al., 2002) and also because preliminary bivariate correlations showed that moderate correlations existed between both STAI measures and compensation-seeking status.

When considering the unadjusted results, BDI-2 was significantly predictive of ETF such that each increase of 1 point on the BDI-2 was associated with an 8% increase in the odds of ETF. Controlling for compensation-seeking status did not modify the predictive relationship.

By contrast, STAI-s but not STAI-t was significantly predictive of ETF when considering the unadjusted odds. When the odds were controlled for compensation-seeking status, the predictive power of both STAI-s and STAI-t trended toward statistical significance, although neither variable was significantly predictive after controlling for multiple comparisons.

Discussion

The current study was conducted in order to examine whether psychological states and personal history factors were statistically predictive of ETF, holding other known predictors of low effort constant. Psychological predictors included having a DSM diagnosed self-reported depressive disorder, anxiety disorder, substance use disorder, and pain disorder; personal history predictors included having a history of sexual abuse, physical abuse, having a psychiatric history and having a substance use disorder.

When the unadjusted relationship between ETF and the psychological variables was considered, depression was found to be a significant predictor of ETF as was diagnosis with an anxiety disorder, while having a substance use disorder and having a pain disorder were not. Of the personal history variables, having a history of sexual abuse trended toward being a significant predictor of ETF but none of the personal history variables were predictive of

ETF once adequate control for multiple comparisons was undertaken. Of the psychological variables, only depression remained predictive of ETF once adequate control for known predictors of ETF and multiple comparisons was undertaken.

Results from the psychometric measures of negative affect (BDI-2, STAI) mirrored the findings based on clinical diagnosis. The results showed that those displaying ETF scored more highly than those displaying VE on the BDI-2, and the effect size was large. Results from the STAI also showed that STAI-s and STAI-t scores were significantly higher in the ETF group than the VE group. Effect sizes for the STAI analyses were medium (STAI-s: $r = .36$; STAI-t: $r = .40$). Unadjusted regression analyses showed that the BDI-2 was significantly predictive of ETF, while only STAI-s was predictive of ETF. When regression analyses were undertaken that controlled for the effect of compensation-seeking status and multiple comparisons of the psychometric measures, only the BDI-2 remained significantly predictive of ETF. The findings indicated that there was a significant relationship between self-reports of depressed affect and ETF that could not be accounted for by other known predictors including compensation-seeking. The finding provided some confirmation of the results of previous investigations (Armistead-Jehle, 2010; Lange et al., 2010; Stulemeijer et al., 2007; Thomas & Youngjohn, 2009).

Chronic pain and substance abuse are both variables that have been reported to be associated with a slow recovery from TBI (Iverson, 2005; Mooney, Speed, & Sheppard, 2005). The present study found that having a co-morbid pain disorder was not predictive of ETF nor was having a current or historical substance use disorder. The findings in respect to substance use and ETF were contrary to the hypothesis of Horton and Roberts (2005) who proposed that substance abusers might be prone to applying low effort. However, the findings were consistent with those of Moore and Donders (2004), Locke et al., (2008), and Miller and Donders (2001) who reported no relationship between a substance abuse *history* and effort and extend those findings to indicate that a current substance use disorder is not predictive of ETF. If having chronic pain or a substance use problem is predictive of poor outcome from TBI, the results of the current study are interpreted to suggest that those relationships do not appear to be significantly mediated by effort.

In contrast to findings by Moore and Donders (2004) and Donders and Boonstra (2007), the results revealed that having a psychiatric history did not predict ETF. The reason for the discrepancy between these findings and those of Donders and colleagues is not clear, but it is notable that the Moore and Donders (2004) and Donders and Boonstra (2007) participants were overwhelmingly (90% and 91% respectively) Caucasian in ethnicity, whereas this sample is only 75% Caucasian. In this study non-Caucasian ethnicities were significantly less likely (χ^2 (2, N = 555) = 19.00, $p < .001$) to report having a psychiatric history than participants of Caucasian ethnicity (it is unclear whether this difference in reporting is due to real differences in the groups, the relative influence of stigmatisation, reduced access to psychiatric services in the non-Caucasian participants or some other variable). An ethnic reporting bias may account for the failure to replicate previous findings.

In the current study an attempt was made to clarify whether a history of sexual abuse or physical abuse were predictive of ETF. A trend toward sexual abuse being predictive of ETF was evident, although that relationship was not statistically significant after employing a Bonferroni alpha adjustment and adjusting for other known predictors. Having a history of physical abuse was not significantly predictive of ETF. Donders and Boonstra (2007) previously found that having an abuse history was predictive of ETF. Of note is the fact that those authors undertook no statistical correction for multiple comparisons. Working from their published findings, had they undertaken such an adjustment their Personal Abuse variable would have shown a statistically non-significant relationship with ETF. However, Williamson et al. (2012) found that abuse was related to EFT in their sample, suggesting that the relationship may be somewhat sample dependent (their sample had a very high incidence of abuse). It is possible that if abuse, including sexual abuse, is significantly predictive of ETF it may be gaining its predictive power from a shared relationship with a potentially more powerful and proximal predictor – self-reported depressive symptoms – with which the variable sexual abuse strongly correlated in the current results (Gamma = 0.5, $p < .001$).

It should be noted that the findings were not sufficient to indicate that high levels of depression are associated with ETF. The BDI-2 is a self-report measure that assesses subjective symptoms and makes make no effort to correct for symptom over-reporting and although the diagnosis of major depressive disorder was made on the basis of strict DSM criteria, that diagnosis was also made on the basis of self-report data. Arguably, the seeking of secondary gains could equally be driving both the fabrication/over-reporting of negative emotional symptoms and the display of ETF. That explanation is conceivable but appears relatively unlikely given that other independent groups of researchers have consistently indicated that psychological symptom over-reporting and poor cognitive effort load on independent factors (e.g., Nelson, Sweet, Berry, Bryant, & Granacher, 2007; Ruocco et al., 2005; Whiteside, Dunbar-Mayer, & Waters, 2009; Williamson et al., 2012). In keeping with that finding, although a point biserial correlation between ETF and BDI-2 ($r = .41$, $p < .001$) was statistically and clinically significant, only around 17% of the variance in depressive symptom reporting was shared with ETF. Additionally, no significant correlation between compensation-seeking status and BDI-2 scores or between compensation-seeking status and diagnosis with a depressive disorder was found. Personality factors including Type-D personality have been previously found to be related to low test-taking effort (Stulemeijer et al., 2007) and notably, the current study replicated the findings of Stulemeijer et al. who also showed that higher levels of depression reporting on the BDI were associated with low test-taking effort, and the effect size in that study was also large ($d = .80$).

The present study had a number of weaknesses. Dichotomizing personal history variables such as personal abuse history increases the risk of Type II error by reducing statistical power. Although the large sample size may have mitigated that risk, future studies should consider collecting continuous data in relation to historical information (e.g., Kubany & Haynes, 2004). Second, the absence of any psychological symptom validity measure prevented clarification of whether psychological symptom *over-reporting* or fabrication accounted for the relationship between the BDI-2 and ETF. Future studies should consider employing the MMPI-2 Fp scale or other emerging scales of affective dissimulation (eg, Structured Inventory of Malingered Symptomatology; Smith & Burger, 1997), to allow for analysis of psychological symptom over-reporting. The administration of self-report measures to subgroups of the sample introduced the risk of sampling bias. That is, it may the case that those participants administered the BDI-2 and STAI differ from the remaining sample in important and undetermined characteristics. This risk is relatively low in respect of the BDI-2 findings because the BDI-2 was administered consecutively from 2005. However, in that period the STAI was administered to those who were clinically showing signs of significant anxiety. The STAI results may only be valid in a more anxious population and it remains unclear whether findings would generalise to a less anxious population. Lastly, insufficient cases relative to variables precluded incorporating BDI-2 and STAI data into a more sophisticated regression analysis, controlling for known covariates of ETF. Although some limited statistical analysis was undertaken to control for the influence of compensation-seeking on the relationship between ETF and BDI-2, the possibility remains that the relationship between ETF and BDI-2 is accounted for by another, as yet undetermined variable.

In summary, the results of the current study demonstrated that self-reported depressed affect is significantly statistically associated with low test taker effort in a TBI sample. It is concluded that high levels of self-reported negative affect, particularly high scores on the BDI-2, should cause the clinician to pay close attention to test taker effort during neuropsychological assessment.

References

Armistead-Jehle, P. (2010). Symptom validity test performance in U.S. Veterans referred for evaluation of Mild TBI. *Applied Neuropsychology, 17*, 52-59.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.

Baker, R., Donders, J., & Thompson, E. (2000). Assessment of incomplete effort with the California Verbal Learning Test. *Applied Neuropsychology, 7*, 111–114.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for Beck Depression Inventory-II*. San Antonio, TX: Psychological Corporation.

Belanger, H. G., Curtiss, J. A., 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,* 215–227.

Bianchini K. J., Curtis, K. L., & Greve, K. W. (2006.) Compensation and malingering in traumatic brain injury: A dose-response relationship? *The Clinical Neuropsychologist, 20*, 831-847.

Bigler, E. D. (2008). Neuropsychology and clinical neuroscience of persistent postconcussive syndrome. *Journal of the International Neuropsychological Society, 14*, 1–22.

Boone, K. B., Salazar, X., Lu, P., Warner-Chacon, K., & Razani, J. (2002). The Rey 15-Item Recognition Trial: A technique to enhance sensitivity of the Rey 15-Item Memorization Test. *Journal of Clinical and Experimental Neuropsychology, 24*, 561–573.

Carroll, L., Cassidy, J., Peloso, P.M., Borg, J., von Holst, H., & Holm, L. (2004). Prognosis for mild traumatic brain injury: results of the WHO Collaborating Centre Task Force on Mild Traumatic Brain Injury, *Journal of Rehabilitation Medicine, 43,* 84–105.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (second ed.). New York: Lawrence Erlbaum Associates.

De Guise, E., Le Blanc, J., Feyz, M., & Lamoureux, J. (2006). Prediction of outcome at discharge from acute care following traumatic brain injury. *Journal of Head Trauma Rehabilitation, 21,* 527-536.

Donders, J., & Boonstra, T. (2007). Correlates of invalid neuropsychological test performance after traumatic brain injury. *Brain Injury, 21,* 319–326.

Dreger, R. (2012). Test review of the State Trait Anxiety Inventory. In J. F. Carlson, K. F. Geisinger, & R. A. Spies (Eds.), *The nineteenth mental measurements yearbook*. Retrieved from the Buros Institute's *Test Reviews Online* website: http://www.unl.edu/buros

Dube, S. R., Williamson, D. F., Thompson, T., Felitti, V. J., Anda, R. F. (2004). Assessing the reliability of retrospective reports of adverse childhood experiences among adult HMO members attending a primary care clinic. *Child Abuse and Neglect, 28,* 729–737.

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,* 101-106.

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,* 375–384.

Farmer, R. F. (2012). Test review of the Beck Depression Inventory-II. In J. F. Carlson, K.

F. Geisinger, & R. A. Spies (Eds.), *The nineteenth mental measurements yearbook*. Retrieved from the Buros Institute's *Test Reviews Online* website: http://www.unl.edu/buros

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 (Ed.), *Assessment of feigned cognitive impairment* (pp. 281-309). New York: The Guilford Press.

Gorissen, M., Sanz, J. C., & Schmand, B. (2005). Effort and cognition in schizophrenia patients. *Schizophrenia Research, 78*, 199-208.

Granacher, R. P. (2008). *Traumatic brain injury: Methods for clinical and forensic neuropsychiatric assessment*. New York: CRC Press.

Greiffenstein, M. F., Baker, W. J., & Gola, T. (1994). Validation of malingered amnesia measures with a large clinical sample. *Psychological Assessment*, *6,* 218-224.

Greiffenstein, M., Gervais, R., Baker, W. J., Artiola, L., & Smith, H. (2013). Symptom Validity Testing in Medically Unexplained Pain: A Chronic Regional Pain Syndrome Type 1 Case Series. *The Clinical Neuropsychologist*, *27*, 138-147.

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*, 1176–1190.

Greve, K.W., Etherton, J.L., Ord, J., Bianchini, K.J., & Curtis, K.L. (2009). Detecting malingered pain-related disability: classification accuracy of the Test of Memory Malingering. *The Clinical Neuropsychologist, 23*, 1250–1271.

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,* 896-918.

Henry, G.K., Heilbronner, R.L., Mittenberg, W., Enders, C., Stevens, A., & Dux, M. (2011). Noncredible performance individuals with external incentives: Empirical derivation and cross-validation of the Psychosocial Distress Scale (PDS). *Applied Neuropsychology, 18,* 47-53.

Horton, A. M., Roberts, C. (2005). Derived Trail Making Test cutoffs and malingering among substance users. *International Journal of Neuroscience, 115*, 1083-1096.

Iverson, G. L. (2005). Outcome from mild traumatic brain injury. *Current Opinion in Psychiatry, 18*, 301–317.

Iverson, G. L., Slick, D. J., & Franzen, M. D. (2000). Evaluation of a WMS-R Malingering Index in a non-litigating clinical sample. *Journal of Clinical and Experimental Neuropsychology, 22,* 191–197.

Johnson-Greene, D., Brooks, L., & Ference, T. (2013). Relationship Between Performance Validity Testing, Disability Status, and Somatic Complaints in Patients With Fibromyalgia. *The Clinical Neuropsychologist*, *27*, 148-158.

Kirsch, N. L., de Leon, M. B., Maio, R. F., Millis, S. R., Tan-Schriner, C. U., & Frederiksen, S. (2010). Characteristics of a mild head injury subgroup with extreme, persisting distress on the Rivermead Postconcussion Symptoms questionnaire. *Archives of Physical Medicine and Rehabilitation*, *91*, 35-42.

Kubany, E. S., & Haynes, S. N. (2004). Traumatic life events questionnaire. Los Angeles, CA: Western Psychological Services.

Lange, R.T., Iverson, G.L., Brooks, B.L., & Rennison, V.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,* 961–972.

Le Blanc, J., De Guise, E., Gosselin, N., & Feyz, M. (2006). Comparison of functional outcome following acute care in young, middle-aged and elderly patients with traumatic brain injury. *Brain Injury, 20,* 779-790.

Locke, D. E. C., Smigielski, J. S., Powell, M. R., & Stevens, S. R. (2008). Effort issues in post acute outpatient acquired brain injury rehabilitation seekers. *Neurorehabilitation, 23,* 273–281.

Lovell, M. R., Iverson, G. L., Collins, M. W., Podell, K., Johnston, K. M., Pardini, D., et al. (2006). Measurement of symptoms following sports-related concussion: Reliability and normative data for the post-concussion scale. *Applied Neuropsychology, 13*(3), 166–174.

Miller, L. J., & Donders, J. (2001). Subjective symptomatology after traumatic head injury. *Brain Injury, 15,* 297-304.

Miller, L. J., Ryan, J. J., Carruthers, C. A., Cluff, R. B. (2004). Brief Screening Indexes for Malingering: A Confirmation of Vocabulary Minus Digit Span from the WAIS-III and the Rarely Missed Index from the WMS-III. *The Clinical Neuropsychologist, 18,* 327-333.

Mooney, G., Speed, J., & Sheppard, S. (2005). Factors related to recovery after mild traumatic brain injury. Brain Injury, 19(12), 975–987.

Moore, B. A., & Donders, J. (2004). Predictors of invalid neuropsychological test performance after traumatic brain injury. *Brain Injury, 18,* 975–984.

Nelson, N. W., Sweet, J. J., Berry, D. T., Bryant, F. B., & Granacher, R. P. (2007). Response validity in forensic neuropsychology: Exploratory factor analytic evidence of distinct cognitive and psychological constructs. *Journal of the International Neuropsychological Society*, 13, 440–449.

Paniak, C., Reynolds, S., Toller-Lobe, G., Melnyk, A., Nagy, J., & Schmidt, D. (2002). A longitudinal study of the relationship between financial compensation and symptoms after treated mild traumatic brain injury. *Journal of Clinical and Experimental Neuropsychology, 24,* 187–193.

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

Ruff , R. M. (2011). Mild traumatic brain injury and neural recovery: Rethinking the debate. *Neurorehabilitation, 28,* 167-180.

Ruocco, A. C., Swirsky-Sacchetti, T., Daines, D., Chute, D. I., Platek, S. M., & Mandel, S. (2005). Dissociation between neuropsychological and psychiatric forms of malingering. *Archives of Clinical Neuropsychology*, *20*, 883-883.

Sacks, A. L., Fenske, c. L., Gordon, W. A., Hibbard, M. R., Perez, K., Brandau, S., Cantor, J., Ashman, T., & Spielman, L. A. (2009). Co-Morbidity of Substance Abuse and Traumatic Brain Injury. *Journal of Dual Diagnosis, 5,* 404–417.

Schagen, S., Schmand, B., Sterke, S. D., & 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*, 43-51.

Schretlen, D. J., & Shapiro, A. M. (2003). A quantitative review of the effects of traumatic brain injury on cognitive functioning. *International Review of Psychiatry*, *15*, 341-349.

Smith, G.P., & Burger, G.K. (1997). Detection of malingering: A validation of the Structured Inventory of Malingered Symptomatology (SIMS). *Journal of the American Academy of Psychiatry and the Law*, *25*, 183–189.

Spielberger, C. (1983). *State-Trait Anxiety Inventory (Form Y)*. Palo Alto, California: Mind Garden.

Stulemeijer, M., Andriessen, T. M. J. C., & Brauer, J. M. P., Vos, P. E., Van der Werf, S. (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*, 309-318.

Suhr, J. A. & Barrash, J. (2007). Performance on standard attention, memory, and psychomotor speed tasks as indicators of malingering. In G. J. Larrabee (Ed.), *Assessment of malingered neuropsychological deficits* (pp. 334-371). Oxford: Oxford University Press.

Suhr, J., Tranel, D., Wefel, J., & Barrash, J. (1997). Memory Performance after Head Injury: Contributions of Malingering, Litigation Status, Psychological Factors, and Medication Use. *Journal of Clinical and Experimental Neuropsychology, 19*, 500-514.

Thomas, M. L., & Youngjohn, J. R. (2009). Let's not get hysterical: Comparing the MMPI- 2 validity, clinical, and RC scales in TBI litigants tested for effort. *The Clinical Neuropsychologist*, *23,* 1067–1084.

Tombaugh, T. (1996). *Test of Memory Malingering (TOMM).* North Tonawanda, NY: Multi-Health Systems.

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,* 297-313.

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.

Whiteside, D. M., Dunbar-Mayer, P., & Waters, D. P. (2009). Relationship between TOMM performance and PAI validity scales in a mixed clinical sample. *The Clinical Neuropsychologist*, *23*(3), 523-533.

Williamson, D. J., Holsman, M., Chaytor, N., Miller, J. W., & Drane, D. L. (2012). Abuse, not financial incentive, predicts non-credible cognitive performance in patients with psychogenic non-epileptic seizures. *The Clinical Neuropsychologist*, *26*(4), 588-598.

Yancura, L. A., & Aldwin, C. M. (2009). Stability and change in retrospective reports of childhood experiences over a 5-year period: findings from the Davis Longitudinal Study. *Psychology and Aging*, *24*, 715–721.

CHAPTER 7: ACCULTURATION AS A PREDICTOR OF EFFORT TEST FAILURE

Paper prepared for submission to the journal *The Clinical Neuropsychologist*

Abstract

English as a second language and acculturation variables were examined as predictors of effort test failure (EFT) in an archival sample of 555 traumatically brain injured (TBI) adults.

Bivariate and logistic regression models were used to examine whether acculturation variables (being foreign born, English as a second language, age at which English was learned, years educated in New Zealand, and number of years resident in New Zealand) were associated with effort test failure. Bivariate analyses showed that ETF was significantly associated with all acculturation variables. Logistic regression analyses revealed that being foreign born (OR = 4.21, 95% CI [2.16, 8.20]) and age at which English was learned (OR = 1.22, 95% CI [1.21, 1.31) were significantly predictive of ETF when the variables' predictive relationships with ETF were unadjusted for other variables. When included in a predictive model that incorporated compensation-seeking, diagnosis with a self-reported mood disorder, exaggerated displays of symptom behavior, psychotic illness, having sustained a workplace accident, and mild TBI severity relative to very severe TBI severity, only age at which English was learned made a significant independent contribution to the predictive model (OR = 1.28, 95% CI [1.13, 1.44]). It was concluded that certain acculturation variables were significantly associated with ETF. Directions for further research were discussed.

Numerous studies of healthy populations have shown that ethnicity and acculturation impact on cognitive test performance. Studies typically show a bias favoring those of Caucasian and English-speaking background such that those of minority cultures or those who have English as a second language (ESL) perform relatively poorly, with effect sizes across a wide range of cognitive and neuropsychological measures typically being medium to large (Walker, Batchelor, & Shores, 2009). Those findings have been widely demonstrated and include results of studies undertaken in the USA with African American vs Caucasian samples (Patton et al., 2003), Hispanic vs Caucasian samples (Coffey, Marmol, Schock, & Adams, 2005), Asian and Middle Eastern vs Caucasian samples (Razani, Burciaga, Mador, & Wong, 2006), as well as studies undertaken in Australia with a culturally and linguistically diverse (CALD) vs English-speaking sample (Carstairs, Myors, Shores, & Fogarty, 2006), in New Zealand (NZ) with Maori vs Caucasian samples (Ogden, Cooper, & Dudley, 2003), in South Africa with White vs Black and Indian samples (Owen & Lynn, 1993), and in the UK with African Caribbean vs White samples (Stewart, Richards, Brayne, & Mann, 2001).

Those findings have been replicated in studies that have been undertaken to examine the impact of ethnicity and acculturation variables in clinical samples. Kennepohl, Shore, Nabors, & Hanks (2004) found that lower levels of acculturation were associated with significantly poorer test performance on a number of neuropsychological measures in a traumatic brain injury (TBI) sample. Boone, Victor, Wen, Razani, & Pontón (2007) found that having English as a second language (ESL) and having low levels of acculturation (conceptualized as years educated in the United States, age at which English was learned, and number of years in the United States) were associated with poorer neuropsychological test performance in a diverse clinical sample. Walker, Batchelor, Shores, & Jones (2010) found that in a moderate-to-severe TBI sample, neuropsychological performances were significantly related to cultural background such that those with a CALD background performed more poorly than those with English speaking, and English-educated backgrounds.

Very little research has been undertaken to examine whether symptom validity test performance is influenced by ethnicity and acculturation variables, however, there is a growing call to investigate the role of acculturation in effort testing research (Faust, 2012). Vilar-Lopez et al. (2007; 2008a, b), has conducted one of the few cross-cultural neuropsychological studies on effort to date. Vilar-Lopez and colleagues examined performance of three effort measures (Victoria Symptom Validity Test (VSVT), Test of Memory Malingering (TOMM), and the b Test) in a Spanish population resident in Spain. The results of that study supported the contention that the use of those tests in the Spanish population was appropriate in that the tests were able to discriminate between groups of brain-injured individuals who were not suspected of malingering and those litigating with a high probability of malingering.

Salazar, Lu, Wen, & Boone (2007) employed the research methodology used by Boone et al. (2007) to investigate whether there was any relationship between acculturation variables and effort measures in a sample of 167 participants with varied neuropsychiatric conditions. Salazar et al. examined whether ethnicity, ESL, and acculturation variables including the age at which English was learned, number of years in the United States, and number of years educated in the United States were associated with ETF. Covarying for age and education significant differences between ethnic groups were found on Digit Span effort measures (Babikian, Boone, Lu, & Arnold, 2006), Rey Auditory Verbal Learning Test effort measures (Boone, Lu, & Wen, 2005; Suhr, Tranel, Wefel, & Barrash, 1997) and Rey-Osterrieth effort measures (Lu, Boone, Cozolino, & Mitchell, 2003; Sherman, Boone, Lu, & Razani, 2002) such that Caucasians scored significantly higher than ethnic minorities on all of those measures. Native English speakers outperformed the ESL participants on the

Reliable Digit Span. Digit Span measures were found to be significantly related to the age at which English was learned but no effort scores were related to number of years resident in the USA or number of years educated in the USA.

Burton, Vilar-Lopez, and Puente (2012) progressed the earlier investigations of Vilar-Lopez (2007; 2008a, b) to examine the use of the Dot Counting Test (DCT), the Rey Fifteen Item Test (FIT) and the TOMM in Spanish speaking residents of the USA involved in assessments in a forensic context. They determined that although the FIT and the TOMM successfully discriminated Spanish-speaking groups with different motivations to give adequate effort, the DCT did not.

Strutt and colleagues (2012) investigated the use of the TOMM with a small (N = 20) US foreign-born Spanish-speaking TBI cohort (compensation-seeking status not documented). They found that scores on TOMM were generally lower than expected. The mean TOMM Retention score for participants that were determined to be affording valid effort (on the basis of a collection of embedded measures and clinical judgment; n = 16) was close to the traditional low effort cut-score (M = 46.2, SD = 5.70) and the mean TOMM Retention score for those deemed to be affording a suboptimal effort (n = 4) was extremely low (M = 26.5, SD = 4.66).

Schroeder, Twumasi-Ankrah, Baade, and Marshall (2012), examined a diverse clinical pool of patients (N = 807) and showed that those with English as a second language had lower performance on the Reliable Digit Span than those with native English, a finding that was broadly supported by Yang and colleagues (2012) investigating the use of Digit Span-based effort indicators in a sample of 132 Taiwan Chinese participants.

Webb, Batchelor, Meares, Taylor, and Marsh (2012) examined the role of ethnicity as a predictor of effort test failure (ETF) in the context of a multivariable logistic regression model that included age, years of education, being of foreign born immigrant status, self-reported mood disorder, psychotic symptomatology, exaggerated behavioral displays of symptoms, injury severity, workplace accident, and compensation seeking. When ethnicity was considered as a predictor of ETF alone (unadjusted), those of Pacific Island ethnicity or of Indo/Asian ethnicity were significantly more likely to display ETF than those of European/White ethnicity. When incorporated into a model that included being of foreign born origin, ethnicity was no longer predictive of ETF but foreign born status made a significant, unique contribution to the predictive model such that the odds of ETF were five times greater among the foreign born sample than the NZ born sample.

Although specific findings vary from study to study, there is a clear trend in the neuropsychological and cognitive testing literature showing that those of CALD backgrounds perform more poorly than those of the majority culture (White/Caucasian) on cognitive measures. In addition, there is some emerging evidence that those of CALD backgrounds may perform more poorly on effort measures than those of majority cultures and those for whom English is their first language.

The purpose of the current study was to more closely examine the role of acculturation variables in predicting ETF and specifically, to examine those variables identified as potentially important by Boone et al. (2007) and by Salazar and colleagues (2007). Those study findings have not been examined or replicated in another sample or in a culturally diverse context outside of the USA and as such it is not clear to what extent the findings generalize. Additionally, the Webb et al. (2012) findings suggested that being foreign born may account for the acculturation effect on effort tests but that study did not examine the effect of being foreign born in the context of other acculturation variables. Therefore, the current research was conducted in order to determine whether, holding constant the Webb et al. (2012) known predictors of ETF, acculturation variables were predictive of ETF. The acculturation variables of interest included being foreign born, years educated in NZ, age at which English was learned, number of years resident in NZ, and English as a second language.

Method

Participants

An archival sample of 555 consecutive referrals to a private clinical neuropsychology practice in Auckland, NZ was examined. Participants had sustained a TBI and were over the age of 16 years at the time of assessment. The data set was collected over the period 2001 to 2007 inclusive. Participants were excluded on the basis of having a pre-existing history of mental retardation or dementia.

Demographic variables.

Males made up 72.8% of the sample. The mean age was 41 years (*SD* = 12.31, range = 16-76). Ethnicity was as follows: European/White (*n* = 418, 75.3%), Maori/Pacific Island (*n* = 105, 18.9%), Indo/Asian (*n* = 32, 5.8%). The foreign-born subgroup consisted of any non-NZ born individuals and comprised 18% (*n* = 99) of the sample. The sample had, on average, 11.8 (*SD* = 2.5, range = 2-21) years of education and the foreign born sample were, on average, more educated than the NZ born sample (t (553) = 2.80, p = .006, d = .35).

Injury variables

Injury variables and the method employed for categorizing injury severity has been previously described in detail (Webb et al., 2012). In brief, severity of injury was based on Glasgow Coma Scale scores at the Emergency Department, duration of loss of consciousness, and duration of posttraumatic amnesia. TBI severity was classified as shown in Table 1. The mTBI-complicated and Moderate TBI groups were combined as were the Very Severe TBI and Extremely Severe TBI groups to ensure that parameter estimates were based on adequate numbers of cases and that there were no empty cells. There was no significant association between severity of injury and NZ born/Foreign born status (χ^2 (3, N = 555) = 1.26, p = .74).

Table 1

Injury Severity Criteria and Sample Characteristics

Descriptor	Criteria	NZ born *n* (%)	Foreign born *n* (%)
mTBI	LOC < 30 mins, PTA < 24 hours, GCS 13-15 at 30 mins	222 (48.7)	53 (53.5)
mTBI complicated	LOC <30 mins, PTA <24 hours, GCS 13-15 at 30 mins, visible (on CT brain imaging) intracranial abnormality not requiring surgery	26 (5.7)	12 (12.1)
Moderate	PTA 1-24 hours, GCS 9-12	49 (10.7)	5 (5.1)
Severe	PTA 1-7 days, GCS 3-8	66 (14.5)	13 (13.1)
Very severe	PTA 1-4 weeks, GCS 3-8	61 (13.4)	8 (8.1)
Extremely severe	PTA >4 weeks, GCS 3-8	32 (7.0)	8 (8.1)

Note. LOC = Loss of consciousness; mTBI = mild traumatic brain injury; PTA = psttraumatic amnesia; GCS = Glasgow Coma Scale. NZ = New Zealand.

Measures.

All participants completed three measures of effort including one forced-choice symptom validity test. All participants completed the Test of Memory Malingering (TOMM; Tombaugh, 1996), the Fifteen Item Test (FIT; Rey, 1964), and Reliable Digit Span (RDS; Greiffenstein, Baker, & Gola, 1994).

Effort classification criteria.

Effort was classified dichotomously – Valid Effort (VE) or ETF. Cut-offs employed for each measure were set to ensure 90% specificity following Baker, Donders, and Thompson (2000) and Boone, Salazar, Lu, Warner-Chacon, and Razani (2002). Consequently the Greve, Bianchini, and Doane (2006) cut-off of <47 was employed for TOMM2 and TOMM

Retention; An RDS cutoff of <8 was employed according to guidelines of Suhr and Barrash's (2007) review; An FIT cutoff of <9 was used in accordance with the findings of Boone et al. (2002). In line with Webb et al. (2012), ETF was operationalized as failure on any two measures in accord with the findings of Victor, Boone, Serpa, Buehler, and Ziegler (2009) or below chance performance on either trial 2 of TOMM or the retention trial (<18/50 at the 95% confidence interval using the binomial distribution). VE was classified using the requirement that participants pass all three effort measures at the cutoffs described above.

Compensation-seeking.

Most of the sample (n = 470; 84.7%) were seeking compensation continuance or seeking entitlement to compensation. Compensation was defined as worker's compensation income replacement (insurance) payments or disability social security benefits. None of the sample were engaged in litigation (litigation for damages is specifically precluded under the NZ no-fault accident compensation legislation). A sizeable minority (n = 84; 15.3%) of the total sample were ineligible for compensation.

Procedure

This study was approved by the Human Research Ethics Committee of Macquarie University.

Results

Of the 555 participants, 111 (20.0%) were classified with ETF. Of these, seven participants (1%) scored below chance on either Trial 2 or the Retention Trial of the TOMM. A total of 352 participants were classified with VE.

Preliminary Analyses

Table 2 shows the results of exploratory bivariate statistical analyses employing the acculturation variables of interest. Because the bivariate analysis was exploratory, no adjustment for multiple comparisons was undertaken.

Table 2

Acculturation Characteristics of the Participants Grouped According to Effort

Variable	VE *n* = 352		ETF *n* = 111		*p*	Effect Size (*d*/*ϕ*)
	M	*SD*	*M*	*SD*		
Years educated NZ	10.88	3.91	7.20	5.17	<.001	.80
Age English learned	.55	2.32	4.79	9.43	<.001	.62
Years resident NZ	37.53	14.92	33.73	16.34	.03	.24
	n	%	*n*	%		
Immigration status					<.001	.27
NZ born	309	81.5	70	18.5		
Foreign born	43	51.2	41	48.8		
Language					<.001	.28
English	329	80.4	80	19.6		
ESL	23	42.6	31	57.4		

Note. *p*-values are from independent *t*-tests or Chi-Square tests. Effect sizes for *t*-tests are Cohen's *d* and Cramer's Phi (*ϕ*) for Chi-Square tests. ESL = English as a second language. VE = valid effort. ETF = effort test failure. NZ = New Zealand.

All acculturation variables were effective in discriminating between the VE and ETF groups, with the largest effect sizes apparent for years of education in NZ (d = .80) and ESL status (ϕ = .28). Correlations between the variables were calculated and are displayed in Table 3.

Table 3

Bivariate Relationships Among Acculturation Variables and ETF

	ESL	Years of education in NZ	Age English learned	Foreign born	Years resident in NZ
ETF	.70***	-.35***	.34***	.62***	-.10*
ESL		-.63***	.81***	.69***	-.44***
Years Edn. in NZ			-.57***	-.78***	.46***
Age English learned				.60***	-.42***
Foreign born					-.54***

Note. Relationships between continuous/ordinal variables and dichotomous variables are expressed as Pearson product-moment correlation coefficients/point-biserial correlation coefficients. Relationships between dichotomous variables are expressed as Gamma coefficients. ETF = effort test failure. ESL = English as a second language. Edn = education. NZ = New Zealand.
* p < .05. ** p < .01. *** p <.001

The acculturation variables were all significantly intercorrelated, with measures of association ranging from r = -.42 between age at which English was learned and years living in NZ, to r = .81 between ESL and age at which English was learned. All acculturation variables were significantly correlated with ETF, although the relationship between ETF and years living in NZ was weak (r = -.10, p < .05).

Predictive Models

The power of the acculturation variables to predict ETF was examined utilising logistic regression.

Variables in the models.

To minimize problems of multicollinearity and dependency of variables and because ESL is implied in the variable age at which English was learned, ESL was not included for further analysis. As years of education in NZ was strongly correlated with being foreign born and because years of education in the United States was not found to be a significant predictor of ETF by Salazar and colleagues (2007), years of education in NZ was not included for further analysis. Three acculturation variables remained for further analysis: foreign born status, age at which English was learned, and years living in NZ.

The power of variables to predict ETF was first examined unadjusted and then adjusted for the predictors of low effort as revealed in the analyses conducted by Webb et al., (2012) namely: years of education, displaying behavioural floridity (exaggerated displays of symptom-related behaviour), having a self-reported mood disorder (as defined by Diagnostic and Statistical Manual-IV-TR criteria), TBI severity, having sustained a workplace accident, and being compensation-seeking.

Dichotomous variables included foreign born status, presence of a self-reported mood disorder, displays of behavioral floridity, psychosis, having a workplace accident, and compensation-seeking. Injury severity was analyzed as categorical data. Years living in NZ and age at which English was learned were analysed as continuous data.

Logistic regression analyses.

In light of multiple comparisons being employed significance levels were subjected to a Bonferroni adjustment where the alpha level was set at .02.

The fit of the model was examined via regression diagnostics. Examination of residuals revealed 5 cases with very large standardised residuals (Z > 5.00). Those cases were eliminated from the predictive analyses.

Considering the odds of ETF (unadjusted), years resident in NZ was not significantly predictive of ETF, but both foreign born status and age at which English was learned were significantly predictive of ETF. Being foreign born increased the odds of ETF by four times, and each one year increase in age at which English was learned increased the odds of ETF by 22%.

When the predictors of years of education, diagnosis with a self-reported mood disorder, florid displays of exaggerated symptom-related behavior, psychosis, workplace injury, compensation-seeking were included in a predictive model along with the three acculturation variables (being foreign born, age at which English was learned, and years resident in NZ), only age at which English was learned contributed unique variance to the model; foreign born status was no longer significantly predictive of ETF. For every yearly increase in age at which English was learned, odds of ETF increased by about 28%.

The significant correlation (r_{pbs} = .60) between foreign born status and age at which English was learned was likely to have accounted for the failure of foreign born status to make a significant contribution in the multivariable model.

Table 4

Logistic Regression Analysis of ETF as a Function of Acculturation Variables

Acculturation Variable	Adjusted OR[1] (95% CI)	*p* value	OR[2] (95% CI)	*p* value
Foreign Born	1.46 [.45, 4.74]	.53	4.21 [2.16, 8.20]	<.001
Age English learned	1.28 [1.13, 1.44]	<.001	1.22 [1.21, 1.31]	<.001
Years resident in NZ	1.02 [.99, 1.04]	.37	.98 [.97, 1.001]	.07

Note: 1. All three predictive variables were included in a logistic regression model that adjusted for years of education, diagnosis with a self-reported mood disorder, displays of behavioral floridity (exaggerated symptom-related behaviour), psychosis, TBI severity, and having a workplace accident (Webb et al., 2012). OR = Odds Ratio. 2. Variables were analysed unadjusted.

Goodness of fit.

Goodness of fit of the final logistic regression model including all predictors was assessed via the Hosmer and Lemeshow test (χ^2 (8, N = 458) = 10.41, p = .24). The area under the receiver operating characteristic curve for the model was .91 (95% CI [.88, .94]), indicating excellent discrimination in predicting those with ETF and those without (Hosmer & Lemeshow, 2000) and Nagelkerke R^2 = .58.

Discussion

The aim of the current study was to examine the relationship between acculturation variables and ETF. Acculturation variables that have previously been highlighted as potentially predictive of ETF (Salazar, Lu, Wen, & Boone, 2007) were examined in an ethnically diverse sample of participants, not previously included in cross-cultural effort studies. Webb et al. (2012) showed that ethnicity was a significant predictor of ETF when considered independently but when included in a predictive model that included foreign born status, ethnicity was no longer significantly predictive of ETF. Whether other acculturation variables including years resident in NZ, years educated in NZ, ESL, and age at which English was learned were predictive of ETF was examined in the current study.

When considered in bivariate analyses, all acculturation variables were significantly associated with ETF. Selected acculturation variables were further examined in a predictive model that included established statistical predictors of ETF, namely: years of education, diagnosed with a self-reported mood disorder, displaying behavioural floridity, psychosis, injury severity, having sustained a workplace accident, and being compensation-seeking (Webb et al., 2012).

Findings indicated that all acculturation variables assessed were predictive of ETF but years living in NZ had only a weak association with effort, trending toward statistical significance when considered independent of the other known predictors but clearly non-significant when considered holding other known predictors of ETF constant. The findings were consistent with those of Salazar and colleagues (2007) who reported that number of years living in the USA was not significantly related to ETF in their sample.

Foreign born status and age at which English is learned have previously been shown to be significantly related to ETF (Salazar, Lu, Wen, & Boone, 2007; Webb et al., 2012). In the present study, when considered as independent predictors of ETF in logistic regression analyses both variables were significantly predictive of ETF. However, when considered holding known predictors constant, foreign born status failed to contribute independently to the predictive model, while age at which English was learned made a statistically significant independent contribution. The correlations among the variables showed that age at which English was learned was significantly correlated with both being foreign born (r_{pbs} = .60) and with ESL (r_{pbs} = .81) both of which were significantly associated with ETF. Those results suggested that age at which English was learned (a strong predictor of ETF) is a useful proxy measure of both ESL and foreign born acculturation variables in English-speaking contexts. The findings supported those of Salazar and colleagues and suggest that among the acculturation variables, age at which English is learned has the strongest association with ETF. Evidence from the social psychology literature suggests that as social distance increases and group identification decreases self-interested behavior increases and ethical/fair behavior becomes less likely (e.g., Hoffman, McCabe, Shachat, & Smith, 1996; Wenzel, 2004). These findings of reduced test-taking effort in those who learn the dominant culture's language late in life may be a reflection of greater social distance and lower group identification (with the dominant culture) by those people.

A limitation of the current study related to difficulties obtaining independent and uncorrelated acculturation variables. That limitation is intrinsic to the nature of acculturation variables and makes it difficult to isolate the most important aspect of acculturation that is contributing to ETF. Findings point to ESL and foreign born variables being most strongly contributory, results which require further examination in independent samples. Other acculturation variables might be considered in future research. Hofstede's five dimensions of cultures (Hofstede, 2001) have been found to be associated with illness behaviors (Deschepper et al., 2008) and have been posited to be associated with psychological illness variables (Draguns & Tanaka-Matsumi, 2003). These cultural dimensions could be examined as predictors of ETF in another large and culturally diverse sample.

The present findings indicate that in the current sample English learned later in life was associated with an increased risk of ETF during cognitive testing. In the present sample there were no strong associations between being compensation-seeking and the acculturation variables of foreign born status (Gamma = .27, p = .07), English as a second language (Gamma = .27, p = .14) or age at which English was learned (r_{pbs} = .06, p = .14), which suggested that ETF was not clearly or simply motivated by access to compensation in the

CALD sample. Further research is needed to determine the reasons for increased ETF in those foreign born subjects who learn English at a later age.

The results of the current study revealed that a wide range of acculturation variables are predictive of ETF. Webb et al. (2012) showed that ethnicity was predictive of ETF but that it gained its predictive power from foreign born status such that when ethnicity was included in a model with foreign born status, ethnicity failed to contribute uniquely to the predictive model. The present study showed that ESL, age at which English was learned, years resident in NZ, and years educated in NZare associated with ETF but only age at which English was learned was predictive of ETF once all other known predictors of ETF were held constant. The clinical implication of the latter finding is that vigilance to low test taker effort should be maintained when assessing those of foreign born status, particularly when English is a second language and when those language skills are learned later in life.

References

Babikian, T., Boone, K.B., Lu, P., & Arnold, G. (2006). Sensitivity and specificity of various Digit Span scores in the detection of suspect effort. *The Clinical Neuropsychologist, 20,* 145-159.

Baker, R., Donders, J., & Thompson, E. (2000). Assessment of incomplete effort with the California Verbal Learning Test. *Applied Neuropsychology, 7,* 111–114.

Boone, K. B., Lu, P., & Wen, J. (2005). Comparison of various RAVLT scores in the detection of noncredible memory performance. *Archives of Clinical Neuropsychology, 20*(3), 301-319.

Boone, K. B., Victor, T. L., Wen, J., Razani, J., Pontón, M. (2007). The association between neuropsychological scores and ethnicity, language, and acculturation variables in a large patient population. *Archives of Clinical Neuropsychology, 22,* 355–365.

Boone, K. B., Salazar, X., Lu, P., Warner-Chacon, K., & Razani, J. (2002). The Rey 15-Item Recognition Trial: A technique to enhance sensitivity of the Rey 15-Item Memorization Test. *Journal of Clinical and Experimental Neuropsychology, 24,* 561–573.

Burton, V., Vilar-López, R., & Puente, A. E. (2012). Measuring Effort in Neuropsychological Evaluations of Forensic Cases of Spanish Speakers. *Archives of Clinical Neuropsychology, 27,* 262-267.

Carstairs, J.R., Myors, B., Shores, E.A., & Fogarty, G. (2006). Influence of language background on tests of cognitive abilities: Australian data. *Australian Psychologist, 41,* 48-54.

Coffey, D. M., Marmol, L., Schock, L., & Adams, W. (2005). The influence of acculturation on the Wisconsin Card Sorting Test by Mexican Americans. *Archives of Clinical*

Neuropsychology, *20*, 795–803.

Deschepper, R., Grigoryan, L., Stålsby-Lundborg, C., Hofstede, G., Cohen, J., Van Der Kelen, G., Deliens, L., & Haaijer-Ruskamp, F. M. (2008). Are cultural dimensions relevant for explaining cross-national differences in antibiotic use in Europe? *BMC Health Services Research*, *8*.

Draguns, J. G., & Tanaka-Matsumi, J. (2003). Assessment of psychopathology across and within cultures: issues and findings. *Behaviour Research and Therapy*, *41*, 755-776.

Faust, D., Ahern, D. C., Bridges, A. J., & Yonce, L. J. (2012). Assessment of Malingering and Falsification: Conceptual Foundations and Sources of Error. In C. R. Reynolds & A. M. MacNeill Horton (Eds.). *Detection of malingering during head injury litigation* (2nd ed., pp. 1-67). New York: Springer.

Greiffenstein, M. F., Baker, W. J., & Gola, T. (1994). Validation of malingered amnesia measures with a large clinical sample. *Psychological Assessment*, *6*, 218-224.

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*, 1176–1190.

Hofstede, G. (2001). *Culture's consequences: comparing values, behaviors, institutions and organizations across nations*. London: Sage Publications.

Hosmer, D. W., and Lemeshow, S. (2000). *Applied logistic regression* (2nd ed.). John Wiley & Sons, Inc: New York.

Kennepohl, S., Shore, D., Nabors, N., & Hanks, R. (2004). African American acculturation and neuropsychological test performance following traumatic brain injury. *Journal of the International Neuropsychological Society*, *10*, 566–577.

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.

Ogden, J. A., Cooper, E., & Dudley, M. (2003). Adapting neuropsychological assessments for minority groups: A study comparing white and Maori New Zealanders. *Brain Impairment*, *4*, 122–134.

Owen, K., & Lynn, R. (1993). Sex differences in primary cognitive abilities among Blacks, Indians and Whites in South Africa. *Journal of Biosocial Science, 25*, 557-560.

Patton, D., Duff, K., Schoenberg, M., Mold, J., Scott, J., & Adams, R. (2003). Performance of cognitively normal African Americans on the RBANS in community dwelling older adults. *The Clinical Neuropsychologist*, *17*, 515–530.

Razani, J., Burciaga, J., Madore, M., & Wong, J. (2007). Effects of acculturation on tests of attention and information processing in an ethnically diverse group. *Archives of Clinical Neuropsychology, 22*, 331–341.

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

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 (Ed.), *Assessment of feigned cognitive impairment* (pp. 405-427). New York: The Guilford Press.

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*, 21-30.

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.

Stewart, R., Richards, M., Brayne, C., & Mann, A. (2001). Cognitive function in UK community-dwelling African Caribbean elders: normative data for a test battery. *International Journal of Geriatric Psychiatry, 16,* 518-527.

Strutt, A. M., Scott, B. M., Lozano, V. J., Tieu, P. G., & Peery, S. (2012). Assessing sub-optimal performance with the Test of Memory Malingering in Spanish speaking patients with TBI. *Brain Injury*, *26*(6), 853-863.

Suhr, J. A. & Barrash, J. (2007). Performance on standard attention, memory, and psychomotor speed tasks as indicators of malingering. In G. J. Larrabee (Ed.), *Assessment of malingered neuropsychological deficits* (pp. 334-371). Oxford: Oxford University Press.

Suhr, J., Tranel, D., Wefel, J., & Barrash, J. (1997). Memory performance after head injury: Contributions of malingering, litigation status, psychological factors, and medication use. *Journal of Clinical and Experimental Neuropsychology*, *19*(4), 500-514.

Tombaugh, T. (1996). *Test of Memory Malingering (TOMM).* North Tonawanda, NY: Multi-Health Systems.

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,* 297-313.

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

Vilar-Lopez, R., Gomez-Rio, M., Caracuel-Romero, A., Llamas-Elvira, J.M., & Perez-Garcia, M. (2008a). Use of specific malingering measures in a Spanish sample. *Journal of Clinical and Experimental Neuropsychology, 30*, 710-722.

Vilar-Lopez, R., Gomez-Rio, M., Santiago-Ramajo, S., Rodrıguez-Fernandez, A., Puente, A.E., Perez-Garcia, M. (2008b). Malingering detection in a Spanish population with a known-groups design. *Archives of Clinical Neuropsychology 23*, 365–377.

Walker, A. J., Batchelor, J., & Shores, A. (2009). Effects of education and cultural background on performance on WAIS-III, WMS-III, WAIS-R and WMS-R measures: systematic review. *Australian Psychologist, 44*, 216-223.

Walker, A. J., Batchelor, J., & Shores, E. A., & Jones, M. (2010). Effects of cultural background on WAIS-III and WMS-III performances after moderate-severe traumatic brain injury. *Australian Psychologist, 45*, 112-122.

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.

Wenzel, M. (2004). An analysis of norm processes in tax compliance. *Journal of Economic Psychology, 25*, 213–228.

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 on Suboptimal Effort Detection: An Example of the Digit Span Subtest of the WAIS-III in Taiwan. *Archives of Clinical Neuropsychology*, *27*, 869-878.

CHAPTER 8: SELF-REPORTED DEPRESSIVE SYMPTOMATOLOGY AND EFFORT TEST FAILURE: A SYSTEMATIC REVIEW

This paper has been submitted to the journal *Brain Injury* for publication.

Abstract

The current review was conducted in order to examine evidence pertaining to the relationship between self-reported depressive symptomatology (SRDS) and effort test performance in compensation-seeking samples. Systematic searches were conducted in MEDLINE, PsycINFO, and Web of Knowledge between 1950 and 2012 (inclusive). A total of 9,501 articles were screened, of which 19 satisfied inclusion criteria. Studies were rated for their level of evidence and methodological quality according to a structured quality assessment tool. Studies were rated of a high quality overall but methodologies employed typically afforded a low level of evidence. Studies revealed a medium to large effect of SDRS on test taking effort (mean *d* = .73, *SD* = .43) with effects being greatest in those studies that employed the most sensitive measures of test taker effort, including the Word Memory Test and Test of Memory Malingering. Studies that examined psychological symptom-validity using measures including the Minnesota Multiphasic Personality Inventory and Personality Assessment Inventory, showed that psychological symptom-reporting is increased in this population but frank malingering of depressive symptomatology is not indicated. Limitations of the literature and recommendations for future research were discussed.

Studies dating from the 1970's clearly indicated that neuropsychologists were poorly equipped to detect simulation of cognitive impairment using traditional neuropsychological tools (Faust, Hart, Guilmette, & Arkes, 1988; Heaton, Smith, Lehman, & Vogt, 1978). Those findings prompted efforts to develop more sophisticated and specialised tools for the detection of low test taker effort (Bianchini, Mathias, & Greve, 2001); tools that have become known as symptom validity tests (SVTs; Pankratz, 1988).

With the advent of SVTs, researchers examined whether the standalone measures were robust to the influence of various factors including depressed mood. A number of early studies suggested that SVTs were uninfluenced by depression (e.g., Inman et al., 1998; Rees, Tombaugh, & Boulay, 2001) but other studies (e.g., Green, Rohling, Lees-Haley, & Allen, 2001) found that depression was significantly associated with effort test failure (ETF). That apparent bifurcation of findings has continued and two apparently contradictory streams are now evident in the effort literature – those suggesting that it is possible that the cognitive impairments that are seen in depression are accounted for by ETF and are not due to any form of neuropathology (Benitez, Horner, & Bachman, 2011; Green, 2009) and those that find no significant evidence of suboptimal effort in those with affective disturbance (Schroeder & Marshall, 2011).

Goldberg, Back-Madruga, and Boone (2007) conducted a narrative review of the impact of psychiatric disorders on cognitive SVTs and concluded that depression and non-psychotic psychiatric illness had no appreciable impact on effort test performance. However, to date there has been no systematic review of the literature on ETF and self-reported depressive symptomatology (SRDS). The aim of the current review was to systematically summarise the available body of evidence pertaining to whether SRDS as assessed by self-report measures or personality testing was associated with ETF. Because the extant literature has been dominated by research in the context of compensation-seeking and litigation and as there are grounds for assuming that the compensation-seeking and non-compensation-seeking populations might present differently, the focus of the systematic review was solely on compensation-seeking samples.

Method

Systematic literature search

Studies were identified by searching electronic databases: MEDLINE, PsycINFO and Web of Knowledge between 1950 and 2012 (inclusive). The search was undertaken using combinations of the terms: (a) depression, OR (b) mood, OR (c) personality, OR psychiatr* AND (d) effort, OR (e) malinger*, OR (f) symptom validity test*, OR (g) response bias. A MEDLINE MeSH qualifier was included: Psychology. Web of Knowledge Research Area limits included: Psychology, Behavioral Sciences, Psychiatry.

Studies were limited to English-language peer-reviewed journals; dissertations and books/book chapters were not included. Studies were included where an analysis of the relationship between SRDS and effort was undertaken. SRDS was conceptualised as the state of depression or the depressive trait. Effort was conceptualised as test-taking effort as determined by performance on standalone neuropsychological SVTs or embedded cognitive validity measures. Case studies were excluded as were healthy simulator studies.

Levels of evidence

For the purpose of the current review, the levels of evidence suggested by the Australian National Health and Medical Research Council (NHMRC, 2009) were employed to determine the level of evidence of each study. The NHMRC levels of evidence consist of four different levels, with Level I indicating the strongest methodological design and Level IV representing the weakest. Table 1 describes the levels of evidence employed.

Table 1

NHMRC Levels of Evidence

Level	Description
I	A systematic review of Level II studies
II	Prospective cohort study
III-1	All or none[1]
III-2	A retrospective cohort study
III-3	A case-control study
IV	A cross-sectional study or case series

Note. [1] All or none = All or none of the people with the risk factor(s) experience the outcome; and the data arises from an unselected or representative case series which provides an unbiased representation of the prognostic effect

Assessment of study quality and analysis

The methodological quality of each study was assessed in detail using a Critical Appraisal Tool (CAT) based on the Heacock, Koehoorn, &Tan (1997) checklist for the critical appraisal of study results. The CAT has been shown to have good levels of inter-rater reliability (Heacock, Koehoorn, & Tan, 1997). Study quality was assessed across 11 dimensions represented by 14 individual items. A CAT total score was obtained by summing the item scores. Higher scores represent methodologically stronger studies, with a maximum possible score of 12 and a minimum score of 0. Therefore each study received a Level of Evidence rating based on the study type (rating I-IV) and a CAT score based on the quality of methodology (CAT range: 0-12). Table 2 describes the results of employing the CAT for each study.

Table 2

Critical Appraisal Tool Scores for each Study

First Author, Year	CAT Item Number														CAT Total
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	
Boone, 1995	1	1	1	1	.5	.5	0	0	1	0	0	2	.5	.5	9
Suhr, 1997	1	1	1	1	.5	.5	1	0	1	0	0	2	.5	0	9.5
Boone, 1999	1	1	1	0	0	0	0	0	0	0	0	0	.5	0	3.5
Rohling, 2002	1	1	1	1	.5	.5	0	1	.5	.5	0	2	0	0	9
Larrabee, 2003	1	1	0	0	0	0	1	0	1	.5	0	2	.5	.5	7.5
Temple, 2003	1	1	1	1	0	0	0	0	1	.5	.5	2	.5	.5	9
Sumanti, 2006	1	1	1	1	.5	.5	0	0	.5	0	0	2	.5	.5	8.5
Yanez, 2006	1	1	1	1	0	.5	1	0	.5	.5	0	2	.5	.5	9.5
Stulemeijer, 2007	1	1	1	1	.5	.5	0	0	1	.5	.5	2	.5	.5	10
Stevens, 2008	0	1	1	1	0	0	0	1	1	.5	0	2	.5	.5	8.5
Tsanadis, 2008	1	1	1	.5	.5	.5	0	0	1	.5	0	2	.5	.5	9
Henry, 2009	1	1	1	0	.5	.5	0	0	1	.5	0	2	.5	0	8
Thomas, 2009	1	1	1	1	0	.5	0	1	1	.5	.5	2	.5	0	10
Amistead-Jehle, 2010	1	1	1	.5	0	0	0	0	0	0	0	2	0	.5	6
Lange, 2010	1	1	1	1	0	.5	0	0	1	.5	.5	2	.5	0	9
Brooks-Johnson, 2012	1	1	1	.5	0	.5	0	0	1	.5	0	2	.5	.5	8.5
Jones, 2012	1	1	1	0	.5	.5	0	0	1	.5	.5	2	.5	.5	9
Lange, 2012	1	1	1	1	0	.5	0	0	1	.5	0	2	.5	0	8.5
Williamson, 2012	1	1	0	1	0	.5	0	1	0	.5	0	2	.5	.5	8

Note. Studies are ordered chronologically. CAT = Critical Appraisal Tool.

Where possible and when not reported, effect sizes (e.g., Cohen's *d*) were calculated from published data. Reporting of effect sizes (ES) has been strongly recommended by the APA as a means of avoiding exclusive reliance on null hypothesis statistical significance testing (APA, 2010) and Liberati et al. (2009) have recommended reporting ES in systematic review and meta-analytic studies to allow comparisons of findings across-studies. ES calculations were based on the means and standard deviations where these were available. If not available, ES estimations were based on the *t*- or *F*-statistics. Meta-analysis was precluded by the wide variability of SVTs and SRDS measures employed.

Results

After excluding duplicates a total of 9,501 studies were initially screened for inclusion. After initial screening on the basis of article title and abstract and then more detailed manuscript review applying inclusion criteria, a total of 19 studies met the inclusion criteria (see flow diagram, Figure 1). Of the 19 studies, 68% (n = 13) reported finding significant effects of SRDS on effort.

Levels of Evidence

Studies were generally rated with low levels of evidence; only one study (Stulemeijer, Andriessen, Brauer, Vos, & Van der Werf, 2007) rated above level III-3 (case control study) and those researchers employed a retrospective cohort design, examining consecutive emergency department admissions (Level of Evidence: III-2). Most studies (n = 13) rated at level IV and comprised non-consecutive cross-sectional designs. The remaining (n = 5) studies were rated level III-3 based on a case control design methodology.

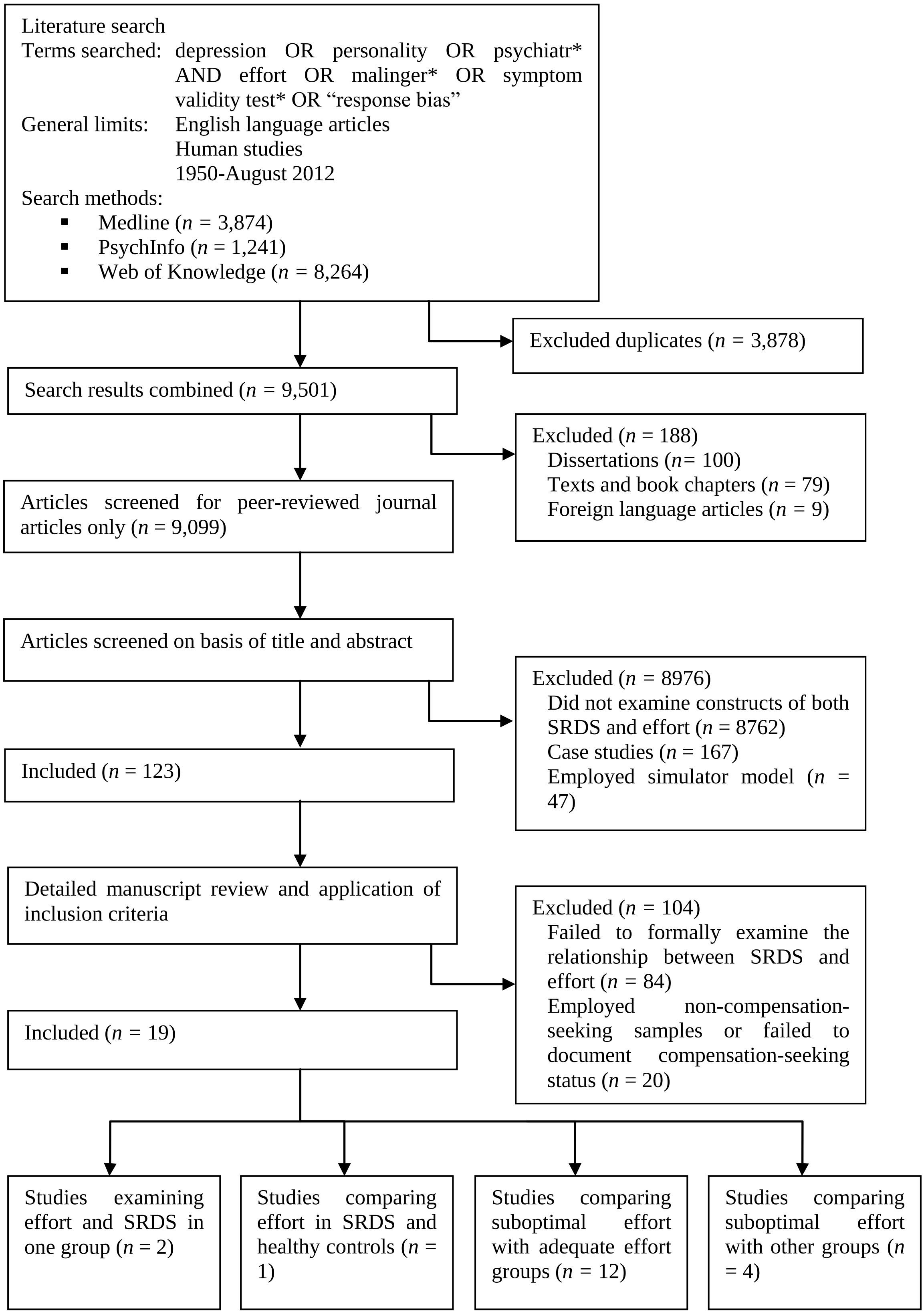

Figure 1. Flow diagram of study selection.

Methodological Quality

Studies were generally of high methodological quality (mean CAT score = 8.5, *SD* = 1.5). The quality of the studies ranged from a CAT score low of 3.5 (Boone & Lu, 1999) to a high of 10.0 (Stulemeijer et al., 2007; Thomas & Youngjohn, 2009). The distribution of CAT scores was negatively skewed (skewness = -2.22) with only three studies receiving a CAT score <8.0.

Methodological considerations fell into three main domains:

i) sample sizes,

ii) reporting of statistical significance and ES,

iii) statistical treatment.

i) Sample sizes.

Sample sizes were low in some of the studies reviewed. Power analysis utilising G*Power (Buchner, 1997) revealed that when undertaking a Wilcoxon Mann-Whitney test of mean differences, a total sample of *N* = 824 is required at a power of .80 and an alpha of .05 to detect a small effect (*d* = .20). Considering the whole sample of studies analysed, of those undertaking two-group mean comparisons, none were found to have adequate sample sizes to detect a small main effect.

Power analysis indicated that to detect a moderate effect size (*d* = .50) of SRDS on effort at a power of .80 and an alpha of .05, the sample size required is 134. Of the studies examined, eight had an insufficient sample size to detect a moderate effect (Armistead-Jehle, 2010; Boone and Lu, 1999; Brooks, Johnson-Greene, Lattie, & Ference, 2012; Lange, Iverson, Brooks, & Rennison, 2010; Stulemeijer et al., 2007; Thomas & Youngjohn, 2009; Williamson, Holsman, Chaytor, Miller, & Drane, 2012; Yanez, Fremouw, Tennant, Strunk, & Coker, 2006).

ii) Reporting of statistical significance and ES.

Across the sample of studies only ten reported ES data and of those, two reported ESs in the course of reporting correlation coefficients. For the sample of 19 studies there was a relationship evident between reporting of ESs and reporting an effect of SRDS on effort. Those studies that reported ESs were more likely than those only reporting *p*-value statistics to report an effect of SRDS on effort (χ^2 (1, N = 19) = 5.63, p = .02, ϕ = .54).

For six of the nine studies that did not report ESs, sufficient published data was available to hand calculate ESs. That data is included in Table 3. For the entire sample of studies for which ESs were published or calculable the mean ES of SRDS on effort was d = .73 (SD = .43), a moderate to large ES.

iii) Statistical treatment.

Statistical tests are typically bound by assumptions about the independence of groups and the distribution of observations. Effort test data tends to be skewed (Schoenberg & Scott, 2011), therefore assumptions required for parametric analyses need careful testing. Studies failed to identify whether the data sets met statistical assumptions for parametric analyses. Across the entire pool of studies, skew was formally considered in only one study (Stulemeijer et al., 2007).

Multivariable data analytic procedures were seldom employed. Logistic regression was employed in only one paper (Henry, Heilbronner, Mittenberg, Enders, & Domboski, 2009), ANCOVA was employed in only one paper (Henry et al., 2009) and MANOVA was employed in two papers (Stulemeijer et al., 2007; Thomas & Youngjohn, 2009). The number of cases was insufficient for any formal analysis of the relationship between use of multivariable analytic techniques and finding an effect of SRDS on effort but it is noteworthy that the studies cited above which employed more sophisticated analytic methods reported finding an effect of SRDS on effort and ESs reported were medium to large (*d's* range: .51 – 2.19).

Choice of effort measure

Across the 19 studies reviewed, 23 separate measures of effort were applied. The WMT was the most widely employed measure, used in 9 of the 19 studies, with the TOMM being the next most frequently employed (*n* = 6), followed by the Rey Fifteen Item Test (FIT) (*n* = 4) and Dot Counting (*n* = 4), then the Portland Digit Recognition Test (PDRT) (*n* = 3). The remaining measures were included in only 1 or 2 studies.

Studies that employed the WMT were most likely to report a main effect of SRDS with seven of the nine studies reporting finding an effect of SRDS. Eight of the nine studies reported or had calculable ESs across 21 comparisons. ESs ranged from small (*d* = .16 (Stevens, Friedel, Mehren, & Merten, 2008)) to large (*d* = 2.19 (Henry et al., 2009)), and the mean ES was large (*M* = .83, *SD* = .10).

Four of the six studies that employed the TOMM reported finding an effect of SRDS on effort. Analysis of ESs, both reported and independently calculated, revealed that for each of the six studies, an effect of SRDS on effort was evident with a large mean ES (*M* = .80, *SD* = .46, range: .29 – 2.19) across 14 comparisons.

The FIT and Dot Counting appeared less sensitive to the effects of SRDS with only three (Boone, et al., 1995; Sumanti, Boone, Savodnik, & Gorsuch, 2006; Thomas & Youngjohn, 2009) of the five studies that employed those measures reporting an effect of SRDS on effort. ESs were not available and could not be calculated for two of the five studies but for those where ESs were available or calculable they were typically smaller than those seen in studies employing TOMM and WMT (ES for FIT studies: *M* = .27 , *SD* = .03; ES for Dot Counting studies: *M* = .43, *SD* = .17). Insufficient data was available in relation to the other effort measures, including embedded measures, to draw any useful comparisons.

Effort test cut-scores to indicate ETF were not applied consistently across studies. Larrabee (2003) employed the FIT using a cut-off of <10, while Sumanti and colleagues (2006) employed a cut-off of <9, while each of the four studies that employed the Dot Counting test employed different cut-scores, as did each of the three studies that employed the PDRT. Consistent (test manual-published) cut-scores were employed in studies using TOMM, WMT, and the Medical Symptom Validity Test (MSVT).

Measures of SRDS

SRDS measures employed also varied widely across studies with a total of 9 separate measures of depression used across the 19 studies. The MMPI was the most frequently employed measure, used in 9 (47%) of the articles, either with or without other formal psychometrics. The BDI was the next most widely employed, used in four studies; the PAI was employed in three studies. Across all studies employing the MMPI scale 2 (Depression) for which ESs were reported or were calculable, the mean ES of SRDS on effort was large (d = .81, SD = .52, range: .29 – 2.19).

Similar results were evident in the four studies which employed the BDI. One of those studies (Suhr, Tranel, Wefel, & Barrash, 1997) reported no effect of SRDS on effort but failed to report specific statistics, while the mean ES (Cohen's *d*) for the remaining three studies, across five comparisons, was medium (M = .65, SD = .16).

Of the three studies employing the PAI, all found a significant effect of SRDS on effort with a medium-to-large mean ES across four comparisons (d = .72, SD = .64).

Methodologies employed

Studies employed one of four research methodologies to examine the role of SRDS in effort test performance: single group studies employing correlation analyses (n = 2); studies comparing ETF in SRDS and healthy groups (n = 1); studies comparing SRDS in those affording adequate effort with those displaying ETF (n = 12; typically comparing group means using performance on SRDS measures as the dependent variable); studies comparing ETF groups with other contrast samples (*n* = 4; comparing group means using performance on SRDS measures as the dependent variable).

Studies examining effort and SRDS in one group.

Two studies (Temple, McBride, Horner & Taylor, 2003; Sumanti et al., 2006; see Table 3) examined the relationship between effort and SRDS in a single group and conducted correlation analyses. Both employed samples of participants with mixed clinical diagnoses and each included different measures of effort. Both studies reported finding a significant effect of SRDS on effort, reporting small to moderate negative correlations between SRDS and effort in each case. The mean absolute correlation across the three comparisons equalled .28 (*SD* = .06), which represented a small-to-medium ES (Cohen, 1988).

Studies comparing effort in SRDS and healthy control groups.

Yanez et al. (2006) examined the impact of SRDS on effort by comparing effort test performances of SRDS (*n* = 20) and unmatched healthy controls (see Table 3). Participants were paid US $5.00 for participation and testing was conducted *following* disability evaluation. The authors found no statistically significant differences between the SRDS and control groups on TOMM Trial 1 (TOMM1), TOMM Trial 2 (TOMM2) or TOMM Retention trials (*p*'s > .05), however, ESs calculated from published data revealed clinically significant ESs for each comparison including TOMM1 (*d* = .64), TOMM2 (*d* = .55) and TOMM Retention (*d* = .40).

Studies comparing ETF with adequate effort groups.

Twelve studies employed a methodology of examining SRDS in groups defined by participants' level of effort – adequate effort and ETF (Armistead-Jehle, 2010; Boone et al., 1995; Boone & Lu, 1999; Brooks et al., 2012; Henry et al., 2009; Jones, Ingram, & Ben-Porath, 2012; Lange et al., 2010; Lange, Pancholi, Bhagwat, Anderson-Barnes, & French, 2012; Stevens et al., 2008; Stulemeijer et al., 2007; Thomas & Youngjohn, 2009; Williamson et al., 2012). Those studies are described in Table 3.

Across the 12 studies, 18 SRDS comparisons were made across groups categorised by level of effort. The range of ESs was *d* = .16 (Stevens et al., 2008) to *d* = 2.19 (Henry et al., 2009) with a large mean ES: *d* = .86, (*SD* = .49).

Three studies reported no relationship between SRDS and ETF (Boone & Lu, 1999; Stevens et al., 2008; Williamson et al., 2012). Boone and Lu (1999), in the methodologically weakest study in the sample (CAT score: 3.5), examined MMPI-2 scale performances in a small (*N* = 19) mixed sample of patients. They reported that no MMPI-2 clinical scale differences were evident between those displaying ETF (*n* = 13) and adequate effort (*n* = 6), however, no statistics or relevant specific results were reported. Stevens and colleagues found that in a mixed sample of compensation-seeking adults (*N* = 233) those displaying ETF (MSVT or WMT) were no more likely to be diagnosed with depression than those displaying adequate effort (χ^2 (1, 64) = .40, $p > .01$, $\phi = .04$, $d = .16$). Finally, Williamson et al. (2012) examined a sample of participants with psychogenic non-epileptic seizures (*N* = 91) and reported that those displaying ETF did not report greater levels of psychopathology on the MMPI-2 than those displaying adequate effort, however, no relevant specific statistics were reported.

Studies comparing ETF group with other contrast groups.

Four studies employed a methodology of comparing an ETF group with another clinical group, in two instances employing depressed comparison samples (Larrabee, 2003; Suhr et al., 1997) and in two instances employing moderate-to-severe TBI comparison samples (Rohling, Green, Allen, & Iverson, 2002; Tsanadis et al., 2008).

Each of the four studies reported a significant relationship between SRDS and effort. Three relevant ESs were reported or calculable across the four studies, with the range of ESs being *d* = .29 (Larrabee, 2003) to *d* = .85 (Rohling, Green, Allen, & Iverson, 2002) and the mean ES: *d* = .59, typically considered to be a medium ES. Suhr et al. (1997) compared the MMPI-2 Scale 2 performance of a DSM diagnosed depressed sample (*n* = 30) with a probable mild traumatic brain injury (mTBI) malingerer group (*n* = 31) and found, consistent with the other studies, that there was no significant differences between the probable malingerer group and the depression group on the MMPI-2 Scale 2 or on the BDI (*p's* > .05). Insufficient detail was reported to calculate relevant ESs.

Table 3

Studies Examining SRDS and Effort

Studies Examining Relationships between SRDS and Effort in One Group					
Study	**Sample**	**Effort measures**	**SRDS measures**	**Results**	**Methodological issues**
					Adjustment for multiple comparisons. One measure of effort employed. Heterogeneous sample. No attempt to control for covariates but confounds were examined. Pearson correlations on non-normal data. Considered psychological symptom validity.
Temple, McBride, Horner, & Taylor (2003)	Mixed-diagnosis veterans referred for neuropsychological evaluation (*N* = 50)	PDRT "easy" and "hard" scores	MMPI-2	MMPI-2 Scale 2 was significantly correlated with PDRT easy trials ($r = -.33$, $p < .05$)*, $d = .35$**	Level of Evidence[1]: IV cross-sectional study CAT score: 9.0
Sumanti, Boone, Savodnik, & Gorsuch (2006)	Sequential psychiatric referrals for workers compensation (*N* = 233)	FIT and Dot Counting	PAI DEP	PAI DEP was significantly correlated with: Rey FIT ($r = -.24$, $p < .001$)*, $d = .25$** and;	Relatively insensitive effort measures. Adjustment for education without explanation. No adjustment for multiple

				Dot Counting (r = .22, p <.01)*, d = .23**	comparisons. Large *N*. Level of Evidence[1]: IV cross-sectional study CAT score: 8.5

Studies Comparing Effort in SRDS and Healthy Controls

Study	**Sample**	**Effort measures**	**SRDS measures**	**Results**	**Methodological issues**
Yanez, Fremouw, Tennant, Strunk, & Coker (2006)	Depressed sample undertaking assessment following social security disability evaluation (n = 20), normal controls (n = 20)	TOMM	BDI-2	On TOMM1 no significant differences between the depressed and controls (p < .06*, d = .64**). On TOMM2 no significant difference between groups (p > .05*, d = .55**). On TOMM Retention no significant difference between groups (p > .05*, d = .40**).	Participants were paid $5 for participating. The study was undertaken *following* the disability evaluation and participants were informed the study was unrelated to the evaluation. Small sample. No consideration of skew, kurtosis and unequal variances (p < .0001) of TOMM data. Did not meet the assumptions for ANOVA, namely equal variance of groups and normal distribution. One measure of effort employed. ESs not calculated. Level of Evidence[1]: III-3 case-control

					study CAT score: 9.5

Studies Comparing ETF With Adequate Effort Groups

Study	**Sample**	**Effort measures**	**SRDS measures**	**Results**	**Methodological issues**
Armistead-Jehle (2010)	Veterans with TBI (*N* = 45)	MSVT	PAI	A higher number of individuals classified as depressed failed MSVT than passed MSVT (χ^2 (1, 45) = 6.3, $p < .05$)*. d = .81**	Small *N*. Potential sampling bias. All injury severity data based on self-report. One measure of effort employed. PAI only administered to a portion (not specified). No ESs reported. Level of Evidence[1]: IV cross-sectional study CAT score: 6.0
Boone, Savodnik, Ghaffarian, Lee, Freeman, & Berman (1995)	Psychiatric claimants of worker's compensation insurance (*N* = 154)	FIT, Dot Counting	MCMI	12% of participants failed one or both effort measures. Those that failed one or both effort measures scored significantly higher on Psychotic Depression subscale (χ^2 = 4.39, p = .04)	Small number failed effort tests (*n* = 17). No reporting of ESs and insufficient data reported to calculate ESs. No controlling for multiple comparisons. Relatively insensitive measures of effort.

					Level of Evidence[1]: IV cross-sectional study CAT score: 9.0
Boone & Lu (1999)	Mixed clinical sample with non-credible cognitive symptoms (*N* = 19)	FIT, Dot Counting, Rey Word Recognition, *b* Test, WRMT, RAVLT 30-minute recognition trial	MMPI-2 clinical scales including Scale 2	No significant differences between those that failed one SVT (*n* = 13) and those that failed zero SVTs (*n* = 6) on MMPI/MMPI-2 clinical scales. No relevant statistics are reported.	Small *N*. No reporting of any statistics. Insufficient data to calculate ESs. Level of Evidence[1]: IV cross-sectional study CAT score: 3.5
Brooks, Johnson-Greene, Lattie, & Ference (2012)	Participants diagnosed with fibromyalgia (*N* = 73)	WMT and Spanish version of WMT	MCMI-III	SVT fail (*n* = 11) scored significantly higher than SVT pass (*n* = 46) on: Depressive Clinical Personality scale ($p \leq .001$, $\eta^2 = .17$)*, $d = .91$** Negativistic Clinical Personality scale ($p \leq .001$, $\eta^2 = .15$)*, $d = .84$** Dysthymia Clinical Syndrome scale ($p \leq .001$, $\eta^2 = .31$)*, $d = 1.34$** Major Depression Severe Clinical Syndrome scale ($p \leq .001$, $\eta^2 = .24$)*, $d = 1.12$** Depression (Axis I) correlated with WMT-IR (*r* =	Unclear method of participant selection. Data suggests that combining Spanish-speaking (22%) and English-speaking (78%) participants potentially biased results. MCMI-III not validated in Spanish language. One measure of effort employed. Level of Evidence[1]: IV cross-sectional study CAT score: 8.5

				-.49*, *d* = .56**) and with WMT DR (*r* = -.46*, *d* = .52**) Depression (Axis II) correlated with WMT-IR (*r* = -.29*, *d* = .30**) and with WMT-DR (*r* = -.25*, *d* = .26**)	
Henry, Heilbronner, Mittenberg, Enders, & Domboski (2009)	Non-litigant head injured controls (‘Non-malingerer’; *n* = 77), ‘Probable Malingerers’ (*n* = 84)	TOMM, CARB, VSVT or WMT	MMPI-2	‘Probable Malingerers’ scored more highly than ‘non-malingerer’ groups on: MMPI-2 RC2 ($p < .001$, *d* = .85)* MMPI-2 Scale 2 ($p < .001$, *d* = 2.19)*	Use of multiple effort measures but relied on failure on only one measure for classification of ETF. Limited description of selection criteria. Well defined group of malingerers. Use of sophisticated statistical procedures. Controlling for covariates. Reported ESs. Omitted to partial out effect of MMDS from the regression of Scale 2 on Effort. Level of Evidence[1]: IV cross-sectional study CAT score: 8.0
Jones, Ingram, & Ben-Porath (2012)	Active duty military members primarily with mTBI (*N* = 501)	TOMM, VSVT, WMT, RBANS EI	MMPI-2-RF	Those failing 3 SVTs (*n* = 60) scored higher than those passing all SVTs (*n* = 220) on: RCd (*d* = .90)*	Appropriate use of MMPI and SVT data. Large *N*. Reporting of ES.

				RC2 (d = 1.14)* RC7 (d = .80)* SUI (d = .68)* HLP (d = .63)*	Employed multiple effort measures. No adjustment for multiple comparisons. Level of evidence[1]: IV cross-sectional study CAT score: 9.0
Lange, Iverson, Brooks, & Rennison (2010)	mTBI (N = 63)	TOMM	DS items of the PCS	‘TOMM fail’ scored significantly higher than ‘TOMM pass’ on items: Sadness (p = .004, d = .93)* Feel more emotional (p = .049, d = .56)*	Unequal group sizes. One measure of effort employed. Appropriate consideration of covariates. Appropriate use of nonparametric analyses. Reported ESs. Level of Evidence[1]: IV cross-sectional study CAT score: 9.0
Lange, Pancholi, Bhagwat, Anderson-Barnes, & French (2012)	Military personnel suffering mTBI (N = 143)	WMT, embedded measures from: TMT, CPT-II, CVLT-II, DSY	PAI DEP scale	mTBI SVT fail group (n = 21) scored higher than severe TBI SVT pass group (n = 35) on PAI DEP (p < .001, d = 1.60)*	No adjustment for multiple comparisons. Appropriate use of multiple SVTs. Reported ESs. Consideration of psychological symptom validity.

					Level of Evidence[1]: IV cross-sectional study CAT score: 8.5
Stevens, Friedel, Mehren, & Merten (2008)	*N* = 233 adults with mixed diagnoses (Depression: *n* = 62)	MSVT or WMT	DSM-IV-TR	Those passing effort measures were no more likely to be diagnosed with depression than those failing effort measures (χ^2 (1, 64) = .40, *p* > .01*, (ϕ = .04, *d* = .16)**	No healthy comparison group. One measure of effort employed. No reporting of ESs. Unclear treatment of multiple comparisons. Large *N*. Unclear how participants were selected. Level of Evidence[1]: IV cross-sectional study CAT score: 8.5
Stulemeijer, Andriessen, Brauer, Vos, & Van der Werf (2007)	Consecutive ED admits with mTBI (*N* = 110)	ASTM	BDI-PC, DS-14	BDI-PC scores were significantly higher in the 'Poor Effort' group than the 'Adequate Effort' group (*p* = .004, *d* = .80)* 'Poor Effort' group were more likely to be classified with Type-D personality (*p* < .001)*, *d* = .66**	One measure of effort employed. Appropriate treatment of skew. Considered multiple comparisons. Appropriate treatment of confound of education. Limited reporting of ESs. Appropriate description of subject pool. Level of Evidence[1]: III-2 retrospective cohort study

					CAT score: 10.0
Thomas & Youngjohn (2009)	TBI patients (n = 83)	PDRT, WMT, Dot Counting	MMPI-2/MMPI-2-RF	Variance in SVT status (Pass/Fail) accounted for by MMPI-2 Scale 2 = 7% (η^2_p = .07)*, d = .55** Variance in SVT status (Pass/Fail) accounted for by MMPI-2-RF RC2 = 6% (η^2_p = .06)*, d = .51**	Use of multiple effort measures. Defined ETF as failure on any one SVT. Multiple comparisons addressed. Appropriate use of multivariable analytic techniques Unclear whether statistical assumptions for DFA were met. Level of Evidence[1]: IV cross-sectional study CAT score: 10.0
Williamson, Holsman, Chaytor, Miller, & Drane (2012)	Patients with PNES (N = 91)	WMT	MMPI-2	Those failing WMT did not report greater levels of psychopathology on MMPI-2 (specific statistics are not reported)	No relevant statistics are reported, nor are ESs. One measure of effort employed. Very selected sample of participants. Unclear generalizability. Level of Evidence[1]: IV cross-sectional study CAT score: 8.0

Studies Comparing ETF Group With Other Contrast Groups

Study	**Sample**	**Effort measures**	**SRDS measures**	**Results**	**Methodological issues**
Larrabee (2003)	Personal injury litigants (*n* = 33), non-litigating TBI (*n* = 47), spinal cord injury (*n* = 42), multiple sclerosis (*n* = 66), chronic pain (*n* = 502), depression (*n* = 30)	PDRT, TOMM, FIT, RMT, RDS<8, Mittenberg WAIS-R DF, and WCST	MMPI-2	'Probable Malingerer' cohort scored *higher* than 'Depressed' cohort on MMPI-2 Scale 2 but the difference was not statistically different (p = .26, d = .29)*	Unclear sampling selection procedure. Heterogeneous sample. Employed multiple SVTs. Controlled for multiple comparisons. ESs reported. Level of Evidence[1]: III-3 case control study CAT score: 7.5
Rohling, Green, Allen, & Iverson (2002)	Compensation-related evaluations (*N* = 719)	CARB, WMT	BDI/MMPI-2/SCL-90	41.6% failed effort measures, 'Low Effort' group had significantly higher BDI scores than a moderate-to-severe TBI group (p < .0001)*, d = .85** 'Low Effort' cohort had significantly poorer performance on effort measures than the brain injury reference group (p < .001)*	Employed multiple SVTs. Limited reporting of data. No control for potential covariates. Large *N*. Level of Evidence[1]: III-3 case control study CAT score: 9.0
Suhr, Tranel,	Probable mTBI	Hiscock forced	BDI,		Appropriate treatment of potential

Wefel, & Barrash (1997)	malingerers (*n* = 31), mTBI comp-seek (*n* = 30), mild/mod TBI not comp-seek (*n* = 20), severe TBI not comp-seeking (*n* = 15), somatizing (*n* = 29), depression (*n* = 30)	choice measure	MMPI-2	'Probable Malingering' and 'Depression' groups were not significantly different on the BDI or MMPI-2 Scale 2 (*p*'s > .05)*	covariates. One measure of effort employed. Attention to multiple comparisons but omitted to make explicit what the significance level was or how it was applied. Small *n*s and Bonferonni control may have increased risk of Type II error. No healthy comparison group. Level of Evidence[1]: III-3 case control study CAT score: 9.5
Tsanadis et al. (2008)	Consecutive neuropsychology evaluations (*N* = 158), 'Moderate-to-Severe TBI' (*n* = 133), versus 'Poor Effort' (*n* = 25)	RMT, TOMM, WMT	Depression Item from the PCSQ	'Poor Effort' sample endorsed more depression symptoms on the PCSQ than the 'Moderate-to-Severe TBI' sample (*p* = .007, *d* = .63)*	Employed multiple SVTs. ETF defined by failing >1 SVT. Effort tests are sensitive. Lacked attention to potential covariates. Lacked control for multiple comparisons. Reported ESs. Sample thoroughly described. Level of Evidence[1]: III-3 case control study CAT score: 9.0

Note. [1] NHMRC = National Health and Medical Research Council (2009) criteria. SRDS = depressive symptomatology. CAT score = Methodological Strength Score. Comp-seek = compensation-seeking/litigating. TBI = traumatic brain injury. mTBI = mild traumatic brain injury. SVT = symptom validity test. PNES = psychogenic non-epileptic seizures. ES = effect size. ETF = effort test failure. DFA = Discriminant function analysis.

Effort Measures: ASTM = Amsterdam Short Term Memory Test. CARB = Computerized Assessment of Response Bias. CPT-II = Conners' Continuous Performance Test - 2nd Edition. CVLT-II = California Verbal Learning Test – II. DSY = Digit Symbol Coding. FIT = Rey 15-Item Test. MSVT = Medical Symptom Validity Test. PDRT = Portland Digit Recognition Test. RBANS EI = Repeatable Battery for the Assessment of Neuropsychological Status Effort Index. RDS = Reliable Digit Span. RAVLT = Rey Auditory Verbal Learning Test. RMT = Recognition Memory Test. TMT = Trail Making Test. TOMM = Test of Memory Malingering. VSVT = Victoria Symptom Validity Test. WAIS-R DF = Wechsler Adult Intelligence Scale (Revised) Discriminant Function. WCST = Wisconsin Card Sorting Test. WMT = Word Memory Test. WMT-DR = Word Memory Test delayed recall. WMT-IR = Word Memory Test immediate recall. WRMT = Warrington Recognition Memory Test.

SRDS Measures: BDI = Beck Depression Inventory. BDI-2 = Beck Depression Inventory (2nd Edition). BDI-PC = Beck Depression Inventory for Primary Care. SRDS-14 = Type D Scale-14. DSM-IV-TR = Diagnostic and Statistical Manual (fourth edition, text revision). HLP = Minnesota Multiphasic Personality Inventory (2nd Edition) Restructured Format helplessness/hopelessness scale. MMPI = Minnesota Multiphasic Personality Inventory. MCMI = Millon Clinical Multiaxial Inventory. MCMI-III = Millon Clinical Multiaxial Inventory-III. MMPI-2 = Minnesota Multiphasic Personality Inventory (2nd Edition). MMPI-2-RF = Minnesota Multiphasic Personality Inventory (2nd Edition) Restructured Format. PAI = Personality Assessment Inventory. PAI DEP = Personality Assessment Inventory Depression Scale. PCS = Postconcussion Scale. PCSQ = Postconcussive Symptom Questionnaire. RC = Minnesota Multiphasic Personality Inventory (2nd Edition) Restructured Format restructured Clinical Scale. SCL-90 = Symptom Checklist-90. SUI = Minnesota Multiphasic Personality Inventory (2nd Edition) Restructured Format suicidal/death ideation scale.

Results: * = Published statistics. ** = Statistics and/or ES calculated from published data. DV = Dependent Variable. IV = Independent Variable.

Consideration of SRDS Validity

Assessment of potential exaggeration of SRDS was undertaken by 11 (58%) of the studies in the sample (Armistead-Jehle, 2010 ; Boone & Lu, 1999; Brooks et al., 2012; Henry et al., 2009; Jones et al., 2012; Larrabee, 2003; Lange et al., 2012; Suhr et al., 1997; Sumanti et al., 2006; Temple et al., 2003; Thomas & Youngjohn, 2009). The results of investigations into the relationship between SRDS validity measures and effort are presented in Table 4.

Some 21 separate scales assessing psychological symptom validity were included in the 11 studies but the majority were represented by the MMPI and its various editions, employed in 7 of 11 studies; the PAI in 3 of 11 studies, and the MCMI-III in 1 of 11 studies.

MMPI studies have examined whether participants have displayed signs of psychological symptom exaggeration on validity scales including the F-scales. Group mean T-scores in the range of T-90 to T-100 are traditionally considered suggestive of marked exaggeration or malingering of psychological symptomatology (Graham, 2000; Nichols, 2001). Two studies omitted to report specific statistics for the validity scales but depicted findings graphically (Boone & Lu, 1999; Suhr et al., 1997). Both concluded that the ETF groups did not display significant SRDS over-reporting. In the sole study employing correlation analyses, validity scales correlated weakly (Mean r = .13, SD = .09) with effort scores (Temple et al., 2003). For the remaining studies that examined ETF groups' performances on psychological validity scales, across 12 observations, mean T-scores for the ETF groups did not reach traditional malingering cut-scores (M = 59.25, SD = 13.36).

Three studies employed the PAI (Armistead-Jehle, 2010; Lange et al., 2012; Sumanti et al., 2006). Sumanti et al. (2006) reported similar findings to those seen with the MMPI, showing weak to moderate correlations (r's = .02 to -.30) between effort measures and psychological validity scales. In the two remaining studies, seven observations involving six

PAI validity scales showed that mean T-scores did not meet traditional cut scores to indicate exaggerated or malingered SRDS (*M* = 61.34, *SD* = 18.4).

Only Brooks and colleagues (2012), who employed the MCMI-III (Disclosure, Desirability, Debasement scales) found evidence of significant psychological symptom exaggeration in the ETF group. About half the ETF group failed psychological validity cut-offs (T-score ≥ 85) on Disclosure and Debasement scales.

Table 4

Studies Employing SRDS Validity Measures

Study	Sample	SRDS validity measure	Findings
Armistead-Jehle (2010)	Veterans with TBI (*N* = 45)	PAI NIM, PAI MAL, Roger's Discriminant Function	No differences evident between the 'Pass MSVT' and 'Fail MSVT' groups evident on PAI measures of exaggeration (*p*'s > .05)* Range of *d:* .1–.45** Mean T-scores on measures of exaggeration below traditional cutoffs (*M* = 61.34, *SD* = 14.43).
Boone & Lu (1999)	Mixed clinical sample with non-credible cognitive symptoms (*N* = 19)	MMPI F, L, K scales	No clinically significant elevation of the validity scales seen in the 'noncredible' group (mean F = 53.1 (*SD* = 85).
Brooks, Johnson-Greene, Lattie, & Ference (2012)	Participants diagnosed with fibromyalgia (*N* = 73)	MCMI-III DIC, DES, DEB	SVT fail group (*n* = 11) scored significantly higher than SVT pass group (*n* = 46) on: DIS scale ($p \leq .001$, $\eta^2 = .23$)* DEB scale ($p \leq .001$, $\eta^2 = .24$)* SVT fail group scored significantly *lower* than SVT pass group on: DES scale ($p \leq .001$, $\eta^2 = .22$)* About half the ETF group scored above traditional protocol invalidity cutoffs
Henry, Heilbronner, Mittenberg, Enders, & Domboski (2009)	Non-litigant head injured controls ('Non-malingerer'; *n* = 77), 'Probable	MMPI-2 MMDS	'Probable Malingerers' scored more highly on MMPI-2 MMDS (*p* < .001, *d* = 1.65)*, than 'Non-malingerer' group.

	Malingerers' (*n* = 84)		
Jones, Ingram, Ben-Porath (2012)	Active duty military members primarily with mTBI (*N* = 501)	MMPI-2-RF scales: VRIN-r, TRIN-r, F-r, Fp-r, Fs, FBS, RBS, L-r, K-r	Those failing 3 SVTs scored higher on most validity scales including VRIN-r (*d* = .30), F-r (*d* = 1.10), Fp-r (*d* = .92), and K-r (*d* = -.65). Scores in the suboptimal effort group were generally not indicative of over-reporting however.
Larrabee (2003)	Personal injury litigants (*n* = 33), non-litigating TBI (*n* = 47), spinal cord injury (*n* = 42), multiple sclerosis (*n* = 66), chronic pain (*n* = 502), depression (*n* = 30)	MMPI-2 F, MMPI-2 L	Mean scores for 'Probable Malingerers' and 'Depression' groups were significantly below traditional malingering cutoffs for both MMPI-2 F (*M* = 66.5, *SD* = 16.7) and MMPI-2 L (*M* = 57.6, *SD* = 11.8)
Lange, Pancholi, Bhagwat, Andersion-Barnes, & French (2012)	Military personnel suffering mTBI (*N* = 143)	PAI INC, INF, NIM, PIM	mTBI SVT fail group (*n* = 21) scored higher than severe TBI SVT pass group (*n* = 35) on PAI NIM ($p < .001$, *d* = 1.13)* but no differences on INF, INC or PIM. Mean NIM score was not elevated into a clearly exaggerated range (*M* = 65.0, *SD* = 18.4)
Suhr, Tranel, Wefel, & Barrash (1997)	Probable mTBI malingerers (*n* = 31), mTBI compensation seeking (*n* = 30), mild/mod TBI not compensation-seeking (*n* = 20), severe TBI not compensation-seeking (*n* = 15), somatizing (*n* = 29), Depression (*n* = 30)	MMPI-2 F, MMPI-2 L	'Probable Malingerer' group showed no significant MMPI-2 F or L scale elevations relative to other groups. Mean F and L scales are not documented specifically but are depicted graphically at well below traditional malingering cutoffs (approximate T-65 and T-55 respectively).
Sumanti, Boone, Savodnik,	Sequential psychiatric	PAI NIM and MAL	Only 2-4% of participants failed both PAI

& Gorsuch (2006)	referrals for workers compensation (*n* = 233)		validity indicator and SVT.
Temple, McBride, Horner, & Taylor (2003)	Compensation seeking veterans referred for neuropsychological evaluation (*N* = 50)	MMPI-2 F, MMPI-2 L	Correlations between PDRT and MMPI-2 F were weak and statistically non-significant (PDRT-easy: *r* = -.14, *p* > .05*; PDRT-hard: *r* = .26, *p* > .05*). Correlations between MMPI-2 L and PRDT were weak and non-significant (PDRT-easy: *r* = -.05, *p* > .05*; PDRT-hard: *r* = .06, *p* > .05*).
Thomas & Youngjohn (2009)	Litigating TBI patients (*n* = 83)	MMPI-2 F, Fb, Fp	Mean F, Fb, & Fp scores not significantly elevated or suggestive of malingering. The variance in SVT status ('Fail' or 'Pass') that can be accounted for by MMPI-2 F, Fb or Fp is low ($\eta^2_p \leq .07$)*

Note. **Depressive symptomatology validity measures:** F-r = Infrequency Scale-Revised. Fp-r = Infrequency Psychopathology Scale-Revised. Fs = Infrequent Somatic Responses scale. L = Lie Scale. L-r = Lie Scale-Revised. K-r = Defensiveness-Revised. MCMI-DEB = Millon Clinical Multiaxial Inventory-III Debasement Index. MCMI-DES = Millon Clinical Multiaxial Inventory-III Desirability Index. MCMI-DIS = Millon Clinical Multiaxial Inventory-III Disclosure Index MMDS = Malingered Mood Disorder Scale. MMPI-2 F = Minnesota Multiphasic Personality Inventory (2nd Edition) Infrequency Scale. MMPI-2 Fb = Minnesota Multiphasic Personality Inventory (2nd Edition) Back F Scale. MMPI-2 Fp = Minnesota Multiphasic Personality Inventory (2nd Edition) Infrequency Psychopathology Scale. MMPI-2 FBS = Minnesota Multiphasic Personality Inventory (2nd Edition) Fake Bad Scale. MMPI-2-RF = Minnesota Multiphasic Personality Inventory (2nd Edition) Restructured Format. PAI INC = Personality Assessment Inventory Inconsistency index. PAI INF = Personality Assessment Inventory Infrequency index. PAI MAL = Personality Assessment Inventory Malingering index. PAI NIM = Personality Assessment Inventory Negative Impression Management index. PAI PIM = Personality Assessment Inventory Positive Impression Management index. P-SVT = Psychological-Symptom Validity Test. RBS = Response Bias Scale. SIMS = Structured Inventory of Malingered Symptomatology. TRIN-r = True Response inconsistency Scale-Revised. VRIN-r = Variable Response Inconsistency Scale-Revised.

Discussion

The purpose of conducting the current review was to systematically examine whether SRDS were associated with ETF in compensation-seeking samples. A consistent finding across the studies reviewed was that there was a relationship between SRDS and ETF. Reduced effort test performances in the SRDS samples did not always reach the level that would trigger effort test cut-offs or attain traditional *p*-value statistical significance (e.g., Yanez et al., 2006), nevertheless, across the entire pool of studies and in every case without exception, the trend or clearly identified statistical relationship was in the direction of SRDS being associated with lower effort. For the entire sample of papers for which ESs were published or calculable the mean ES of SRDS on effort is d =.73 (SD = .43), typically considered to be a moderate to large ES.

The majority of articles reviewed were of relatively high methodological quality but had low Levels of Evidence ratings. A number of studies were limited by potential sample bias and small sample sizes. Cohen (1962) pointed out that small sample size, causing a loss of power, was the single most concerning problem leading to failure to detect effects in psychological research and this problem continues to be evident in this field of research. Most investigations were single-centre studies with restricted samples and that may have limited the generalizability of results.

Studies have used a wide range of measures of test-taking effort. Because sensitivity and specificity of tests can vary widely, that lack of consistency may have contributed to the inconsistency of findings. Studies that employed the WMT and the TOMM most consistently reported an effect of SRDS on effort while those that utilized the FIT and Dot Counting did not. A number of studies have shown that the WMT and TOMM are among the most sensitive effort measures and that both Dot Counting and FIT lack sensitivity (Sollman & Berry, 2011; Vallabhajosula & van Gorp, 2001; Vickery, Berry, Inman, Harris, & Orey, 2001). This suggests that the ability to detect a relationship between SRDS and ETF may depend on the sensitivity of the effort measure.

Failure to document ESs has had a potentially biasing effect in the literature. The APA Taskforce on Statistical Inference stated that authors should "always provide some effect-size estimate when reported a *p* value," (Wilkinson, 1999, p. 599). In the sample of studies included in the current review only 42% specifically reported ESs and those reporting ESs were more likely than those not reporting ESs to detect an effect of SRDS on effort. Additionally, statistical techniques have not always been well tailored to the type of data seen in effort test research with potentially misleading parametric techniques being used on potentially highly skewed data and/or with groups of very unequal variance.

Some researchers have sought to consider the influence of SRDS over-reporting, however, none have undertaken any direct statistical analysis to partial out the effects of over-reporting on SRDS and its relationship with effort. Of those that have examined symptom over-reporting, the majority of researchers have shown that psychological symptom reporting is heightened in those displaying ETF but not to the extent to indicate significant over-reporting or malingering of psychological symptoms. That finding is consistent with repeated research findings that have shown that psychological symptom over-reporting and cognitive under-performance represent relatively (but not wholly) independent constructs (e.g., Haggerty, Frazier, Busch, & Naugle, 2007, Nelson, Sweet, Berry, Bryant, & Granacher, 2007).

Studies employing more sophisticated statistical techniques such as logistic regression techniques, partial correlation, or ANCOVA as a means of extracting the influence of symptom over-reporting from the relationship of SRDS and effort would be appropriate. Future studies might consider employing the F_p scale of the MMPI-2 rather than the more commonly adopted F scale because F scale (infrequency) elevation is seen routinely in cases of depression without symptom exaggeration and because the F_p scale is the most sensitive and specific measure of psychological symptom over-reporting in the MMPI-2 (Nichols, 2001). The Malingered Mood Disorder Scale (Henry, Heilbronner, Mittenberg, Enders, & Roberts, 2008) may prove to have utility in this research but that is a recently developed scale that has yet to be validated with a non-compensation-seeking depressed sample.

Limitations of the Current Review

A limitation of the current review relates to the use of only one rater for determining methodological quality and Levels of Evidence. An attempt to limit the impact of that potential bias was made by employing a structured Quality Assessment Tool with established psychometric properties (Heacock, Koehoorn, & Tan, 1997). It is conceivable that a bias was evident nevertheless. The use of multiple raters and establishing inter-rater reliability is recommended for any future systematic review in this field. Sampling bias may have been inadvertently introduced by inclusion of only research articles written in English as may have undiscovered and unpublished papers (the so-called *file drawer problem* (Rosenthal, 1979)).

The current review deliberately employed a broad conceptualization of the psychological construct of interest – depressive symptomatology – that allowed for the inclusion of studies that examined both the state of depression and the depressive personality trait. That approach was based on the understanding that measures of state and trait tend to be highly correlated, and that there are strong state effects of mood on personality measures (Barnett et al., 2010). Constraining the inclusion criteria to include only the depressed state would have conceivably resulted in a more homogenous sample of participants, however, that would demand a set of tools that better delineate state from trait affect and that may not be possible at this point in time.

In conclusion, the research detailed in this systematic review examined the relationship between SRDS and ETF in the compensation-seeking population. The evidence indicated that SRDS in that population was associated with ETF. Research is needed to systematically examine the role of SRDS in the non-compensation-seeking clinical population.

References

American Psychological Association. (2010). *Publication Manual of the American Psychological Association* (6th ed.). Washington, DC: Author.

Armistead-Jehle, P. (2010). Symptom validity test performance in U.S. Veterans referred for evaluation of Mild TBI. *Applied Neuropsychology, 17*, 52-59.

Barnett, J.H., Huang, J., Perlis, R. H., Young, M. M., Rosenbaum, J. F., Nierneberg, A. A., Saachs, G., Nimgaonkar, V. L., Miklowitz, D. J., & Smoller, J. W. (2010). Personality and bipolar disorder: Dissecting state and trait associations between mood and personality. *Psychological Medicine, 41*, 1593-1604.

Benitez, A., Horner, M. D., & Bachman, D. (2011). Intact cognition in depressed elderly veterans providing adequate effort. *Archives of Clinical Neuropsychology, 26*, 184-193.

Bianchini, K. J., Mathias, C. W., & Greve, K. W. (2001). Symptom validity testing: A critical review. *Clinical Neuropsychology, 15*, 19–45.

Boone, K. B., & Lu, P. H. (1999). Impact of somatoform symptomatology on credibility of cognitive performance. *The Clinical Neuropsychologist*, *13*, 414–419.

Boone, K. B., Savodnik, I., Ghaffarian, S., Lee, A., Freeman, D., & Berman, N. G. (1995). Rey 15-item memorization and dot counting scores in a ‘‘stress’’ claim worker’s compensation population: Relationship to personality (MCMI) scores. *Journal of Clinical Psychology, 51*, 457–463.

Brooks, L., Johnson-Green, D., Lattie, E., & Ference, T. (2012). The relationship between performances on neuropsychological symptom validity testing and the MCMI-III in patients with fibromyalgia. *The Clinical Neuropsychologist, 26,* 816-831.

Buchner, A., Erdfelder, E., & Faul, F. (1997). How to Use G*Power [WWW document]. URL http://www.psycho.uniduesseldorf.de/aap/projects/

gpower/how_to_use_gpower.html

Cohen, J. (1962). The statistical power of abnormal-social psychological research: A review. *Journal of Abnormal and Social Psychology*, *65*(3), 145-153.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (second ed.). New York: Lawrence Erlbaum Associates.

Faust, D., Hart, K., Guilmette, T.J., & Arkes, H. R. (1988). Neuropsychologists capacity to detect adolescent malingerers. *Professional Psychology: Research and Practice, 19*, 508-515.

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 (Ed.), *Assessment of feigned cognitive impairment* (pp. 281-309). New York: The Guilford Press.

Graham, J. R. (2000). *MMPI-2. Assessing personality and psychopathology* (3rd ed.). New York: Oxford University Press.

Green, P. (2009). Questioning common assumptions in depression. In J. Morgan, & J. Sweet (Eds.), *Neuropsychology of malingering casebook* (pp. 132-144). New York: Taylor & Francis.

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*, 1045–1060.

Haggerty, K. A., Frazier, T. W., Busch, R. M., & Naugle, R. I. (2007). Relationship among Victoria Symptom Validity Test indices and Personality Assessment Inventory validity scales in a large clinical sample. *The Clinical Neuropsychologist*, *21*, 917–928.

Heacock, H., Koehoorn, M., & Tan, J. (1997). Applying epidemiological principles to ergonomics: a checklist for incorporating sound design and interpretation of studies. *Applied Ergonomics, 28,* 165-172.

Heaton, R. K., Smith, H. H., Lehman, R. A. W., & Vogt, A. T. (1978). Prospects for faking believable deficits on neuropsychological testing. *Journal of Consulting and Clinical Psychology, 46,* 892-900.

Henry, G. K., Heilbronner, R. L., Mittenberg, W., Enders, C., & Domboski, K. (2009). Comparison of the MMPI-2 Restructured Demoralization scale, Depression scale, and Malingered Mood Disorder Scale in identifying non-credible symptom reporting in personal injury litigants and disability claimants. *The Clinical Neuropsychologist, 23,* 153–166.

Inman, T. H., Vickery, C. D., Berry, D. T. R., Lamb, D. G., Edwards, C. L., & Smith, G. T. (1998). Development and initial validation of a new procedure for evaluating adequacy of effort given during neuropsychological testing: The Letter Memory Test. *Psychological Assessment, 10,* 128–139.

Jones, A., Ingram, M. V., & Ben-Porath, Y. S. (2012). Scores on the MMPI-2-RF scales as a function of increasing levels of failure on cognitive symptom validity tests in a military sample. *The Clinical Neuropsychologist, 26,* 790–815.

Lange, R. T., Iverson, G. L., Brooks, B. L., & Rennison, V. 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,* 961–972.

Lange, R. T., Pancholi, S., Bhagwat, A., Anderson-Barnes, V., & French, L. M. (2012). Influence of poor effort on neuropsychological test performance in U.S. military personnel following mild traumatic brain injury. *Journal of Clinical and*

Experimental Neuropsychology, 34, 453-466.

Larrabee, G. J. (2003). Exaggerated MMPI-2 symptom report in personal injury litigants with malingered neurocognitive deficit. *Archives of Neuropsychology, 18*, 673-686.

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P., Clarke, M., Devereaux, P. J., Kleijen, J., & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Annals of Internal Medicine*, *151*, W-65-94.

National Health and Medical Research Council. (2009). *NHMRC levels of evidence and grades for recommendations for guideline developers.* Retrieved from https://www.nhmrc.gov.au/_files_nhmrc/file/guidelines/developers/nhmrc_levels_grades_evidence_120423.pdf

Nelson, N. W., Sweet, J. J., Berry, D. T., Bryant, F. B., & Granacher, R. P. (2007). Response validity in forensic neuropsychology: Exploratory factor analytic evidence of distinct cognitive and psychological constructs. *Journal of the International Neuropsychological Society, 13*, 440–449.

Nichols, D. S. (2001). *Essentials of MMPI-2 assessment*. New York: Wiley.

Pankratz, L. (1988). Malingering on intellectual and neuropsychological measures. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (pp. 169-192). New York: Guilford.

Rees, L. M., Tombaugh, T. N., & Boulay, L. (2001). Depression and the Test of Memory Malingering. *Archives of Clinical Neuropsychology, 16*, 501-506.

Rohling, M. L., Green, P., Allen, L. M., & Iverson, G. L. (2002). Depressive symptoms and neurocognitive test scores in patients passing symptom validity tests. *Archives of Clinical Neuropsychology, 17*, 205-222.

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, *8*, 638.

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*, 437–453.

Schoenberg, M. R., & Scott, J. G. (Eds.). (2011). *The little black book of neuropsychology*. New York: Springer.

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

Stevens, A., Friedel, E., Mehren, G., & Merten, T. (2008). Malingering and uncooperativeness in psychiatric and psychological assessment: Prevalence and effects in a German sample of claimants. *Psychiatry Research*, *157*, 191–200.

Stulemeijer, M., Andriessen, T. M. J. C., & Brauer, J. M. P., Vos, P. E., Van der Werf, S. (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*, 309-318.

Suhr, J., Tranel, D., Wefel, J., & Barrash, J. (1997). Memory performance after head Injury: Contributions of malingering, litigation status, psychological factors, and medication use. *Journal of Clinical and Experimental Neuropsychology*, *19*, 500-514.

Sumanti, M., Boone, K. B., Savodnik, I., & Gorsuch, R. (2006). Noncredible psychiatric and cognitive symptoms in a worker's compensation "stress" claim sample. *The Clinical Neuropsychologist*, *20*, 754-765.

Temple, R. O., McBride, A. M., Horner, M. D., & Taylor, R. M. (2003). Personality characteristics of patients showing suboptimal cognitive effort. *The Clinical Neuropsychologist*, *17*, 402–409.

Thomas, M. L., & Youngjohn, J. R. (2009). Let's not get hysterical: Comparing the MMPI- 2 validity, clinical, and RC scales in TBI litigants tested for effort. *The Clinical Neuropsychologist*, *23*, 1067–1084.

Tsanadis, J., Montoya, E., Hanks, R.A., Millis, S.R., Fichtenberg, N.L., & Axelrod, B.N. (2008). Brain injury severity, litigation status, and self-report of postconcussive symptoms. *The Clinical Neuropsychologist*, *22*, 1080–1092.

Vallabhajosula, B. & van Gorp, W. G. (2001). Post-Daubert admissibility of scientific evidence on malingering of cognitive deficits. *Journal of the American Academy of Psychiatry and the Law, 29*, 207–215.

Vickery, C. D., Berry, D. T., 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*, 45-73.

Wilkinson, L. (1999). Statistical methods in psychology journals: guidelines and explanations. *American Psychologist*, *54*(8), 594.

Williamson, D. J., Holsman, M., Chaytor, N., Miller, J. W., & Drane, D. L. (2012). Abuse, not financial incentive, predicts non-credible cognitive performance in patients with psychogenic non-epileptic seizures. *The Clinical Neuropsychologist*, *26*, 588-598.

Yanez, Y. T., Fremouw, W., Tennant, J., Strunk, J., Coker, K. (2006). Effects of severe depression on TOMM performance among disability-seeking outpatients. *Archives of Clinical Neuropsychology, 21*, 161–165.

CHAPTER 9: CONCLUSION

The current research investigated the statistical predictors of ETF in a clinical sample of TBI patients. The research was designed to meet the following objectives: (1) To clarify the relative importance of previously identified predictors of ETF (compensation-seeking, TBI severity, education) when those variables were incorporated into a multivariable predictive model; (2) to examine the relationship between other, previously unexplored variables (abnormal nonverbal behaviours, workplace accident) and ETF; (3) to examine the role that psychological and personal history variables (self-reported mood disorder, psychosis, pain disorder, substance use disorder) have in predicting ETF; (4) to examine the role of ethnicity and acculturation variables (being foreign born, English as a second language, age at which English was learned, years educated in NZ, and number of years resident in NZ) in predicting ETF and; (5) to systematically examine the extant literature that describes the relationship between negative affect and ETF in compensation-seeking samples.

To meet these objectives four studies were undertaken. The first study (Chapter 5) comprised a broad examination of the predictors of ETF, examining previously identified predictors (compensation-seeking, TBI severity, education) in a multivariable model alongside other variables that had not previously been examined (abnormal nonverbal behaviours, workplace accident) and also variables that had previously been examined but for which no clear or consistent relationship with ETF had been established (self-reported mood disorder, psychosis, age, education, ethnicity, and being foreign born). The second study (Chapter 6) extended findings from the first study, examining more closely the predictive relationship between psychological (depression, anxiety, pain disorder, substance use disorder, BDI-2, STAI) and personal history variables (psychiatric history, sexual abuse history, physical abuse history, substance use disorder history) and ETF. The third study (Chapter 7) further extended findings of Study 1 that showed that a demographic acculturation variable (being foreign born) was predictive of ETF. In Study 3 an attempt was made to reconcile Study 1 findings in respect to ethnicity and acculturation and ETF with the results of Salazar, Lu, Wen and Boone (2007) who found that certain ethnic and acculturation variables were associated with ETF. In Study 3 the variables identified by Salazar et al. (2007) were adapted to the NZ context (years educated in NZ, age at which English was learned, number of years resident in NZ, and English as a second language) and tested in a multivariable model and with a new and ethnically diverse population from that previously studied. In the fourth study, the somewhat divided and contentious topic of self-reported depressive symptomatology (SRDS) and its relationship with ETF was examined by undertaking a systematic review of the literature in that area.

Summary of Research Findings

Predictors of ETF

The results of the Study 1 confirmed that previously identified correlates of ETF, namely, compensation-seeking (Bianchini, Curtis, & Greve, 2006), low education (Babikian, Boone, Lu & Arnold, 2006) and diagnosis with psychotic illness (Goldberg, Back-Madruga, & Boone, 2007) have a statistically predictive relationship with ETF. Additionally, the previously and well-established inverse relationship between injury severity and ETF (Green, Iverson, & Allen, 1999) was supported by the results of the study. In contrast to the findings of Flach, Krol, & Groothoff (2008) the results indicated that in the current sample age was not a predictor of ETF. Importantly, the study identified new predictors of ETF. Florid behavioural displays of illness were reliably associated with ETF as was being diagnosed with a self-reported mood disorder, having had a workplace injury and being foreign born.

The findings indicated that although compensation-seeking was predictive of ETF, other psychological variables were more strongly associated with effort, such that the odds-ratios associated with being diagnosed with a self-reported mood disorder, displaying behavioural floridity and in particular, being diagnosed with psychotic illness, exceeded the odds-ratio associated with being compensation-seeking.

Psychological predictors of ETF

On the basis of the finding that psychological variables had a strongly predictive relationship with ETF, further investigations were undertaken to examine the relationship between current and historical psychological conditions, personal history variables and ETF. The psychological variables chosen were influenced by the data reported by Iverson (2005), who reviewed the literature examining variables that were associated with poor outcome from TBI. Variables examined included: a DSM diagnosis of pain disorder, depressive disorder, anxiety disorder, and substance use disorder. The predictive relationship between personal history variables and ETF was examined employing the following variables: a history of sexual abuse, a history of physical abuse, a psychiatric history, and a history of substance use disorder. Psychometric variables (patient responses on self-report measures including the BDI-2 and the STAI) were also considered as predictors of ETF.

The results demonstrated that once multivariable statistical analysis controlling for the variables established by Webb et al. (2012) was undertaken, of the psychological and personal history variables, only a diagnosis of depressive disorder was predictive of ETF. A trend toward findings of a predictive relationship between history of sexual abuse and ETF was evident but that relationship was not statistically significant following appropriate control for multiple comparisons. Responses on the BDI-2 were predictive of ETF after controlling for the effect of being compensation-seeking and multiple comparisons but responses on the STAI were not.

Acculturation as a predictor of effort test failure

In study three the predictive relationship between acculturation variables and ETF was examined. A range of acculturation variables including those posited by Salazar and colleagues (2007) were examined as predictors of ETF, namely: years educated in NZ, age at which English was learned, years resident in NZ, being foreign born, and having English as a second language. The predictive relationship between three acculturation variables (being foreign born, age at which English was learned and years resident in NZ) and ETF was examined employing logistic regression analysis adjusting for years of education, displaying behavioural floridity (exaggerated displays of symptom-related behaviour), having a self-reported mood disorder (as defined by Diagnostic and Statistical Manual-IV-TR criteria), TBI severity, having sustained a workplace accident, and being compensation-seeking. Of the acculturation variables, only age at which English was learned made a significant independent contribution to the predictive model.

Self-reported depressive symptomatology and ETF

The findings of previous research studies that have examined the relationship between depressive symptomatology and ETF have been conflicted and seemingly contradictory. While some have reported no significant relationship exists between SRDS and ETF (e.g., Inman et al., 1998), others have reported a relationship (e.g., Green, Rohling, Lees-Haley, & Allen, 2001). In an attempt to reconcile the disparate findings, a systematic review of the literature describing SRDS and ETF in compensation-seeking samples was undertaken. A total of 9,501 papers were screened, 19 of which fulfilled inclusion criteria. Studies were generally of high methodological quality but afforded a low level of evidence. Most studies consisted of non-consecutive cross-sectional designs or case control designs and none employed prospective samples. Across the sample of studies as a whole, a medium to large effect of SRDS on test taking effort was noted (mean d = .73, SD = .43) with effects varying with the measures employed. More specifically, effects were generally greatest in those studies that employed measures of effort that have been demonstrated to have relatively high sensitivity. Even in those studies that found no statistically significant effect of SRDS and effort, in every case and without exception, effort was lower in those with SRDS than in those without SRDS.

Methodological Strengths and Limitations

The literature pertaining to effort has often included heterogeneous samples, combining diverse groups of neurological and psychiatric patients, (e.g., Green, Rohling, Lees-Haley, & Allen, 2001) and mixed or ill-defined compensation-seeking and non-compensation-seeking samples (e.g., O'Bryant, Finlay, & O'Jile, 2007). While that has the potential to increase the generalizability of research findings, heterogeneous samples risk dilution of effects with small effects being rendered insignificant (Lynch, 1999). In the current studies, the participant sample was restricted to TBI patients as a means of minimising the confounding effects of sample heterogeneity and care was taken to identify the compensation-seeking status of participants. The sample employed was also sufficiently large that small effects would be less likely to be diluted statistically and the relatively large sample allowed the use of logistic regression – a powerful multivariable statistical procedure that has been seldom employed in this research field. The strength of this statistical analytic approach is in the ability to examine the predictive power of multiple predictors simultaneously, thereby gaining an impression of the relative importance of predictive variables. Logistic regression is relatively demanding of sample size. It has been recommended that logistic regression strategies be employed in large samples only ($N > 200$; Schutte & Axelrod, 2013), as a means of improving generalizability of findings and reducing risk of sample-dependent findings. The current study employed a sufficient sample to maximise the chances of stable findings and results that can be generalised to the greater population of TBI patients. Unfortunately, the number of variables to cases precluded including all variables of interest from studies 1 through 3 in a single multivariable model.

The use of the multivariable logistic regression technique allowed some appraisal of the relative predictive power of compensation-seeking as a predictor of ETF against the predictive power of psychological variables. The finding that the psychological variables are relatively stronger predictors than compensation seeking supports those of Williamson, Holsman, Chaytor, Miller and Drane (2012) and others (Donders & Boonstra, 2007; Stulemeijer, Andriessen, Brauer, Vos, & Van der Werf, 2007). The approach adopted also allowed an examination of various acculturation variables to determine which afforded the most unique variance to the predictive model of ETF. Of the acculturation variables considered, age at which English is learned proved to demonstrate the strongest predictive relationship with ETF – a result that supports the previous findings of Salazar and colleagues (2007).

Additionally, unlike most statistical procedures, logistic regression is a technique that has no assumptions about the distributions of predictor variables (Tabachnik & Fidell, 2007). The majority of research into effort conducted to date has employed parametric statistical analytic techniques, at times with no consideration of the skewed distribution that is typically seen in effort data (Brooks, Sherman, Iverson, Slick, & Strauss, 2011). The relatively robust logistic regression technique allows for improved confidence in the veracity of findings and reduced concern that findings might be spurious due to being obtained through contravention of the assumptions of (parametric) statistical techniques.

A further strength of the set of studies comprising the current dissertation is the use of multiple measures of effort to classify ETF rather than relying on any one measure alone to determine ETF. There has been a growing call to employ multiple measures to assess effort (Boone, 2007; Larrabee, 2003; Vickery et al., 2004; Victor, Boone, Serpa, Buehler, &

Ziegler, 2009) and the administration of three or more SVTs has been recommended during neuropsychological assessment (Victor et al., 2009). Employing two or more SVT failures as evidence of ETF is in line with the recommendations of Larrabee, Greiffenstein, Greve and Bianchini (2007) and serves to reduce the risk of failing to detect low effort due to the relatively low sensitivity of effort measures, or the risk of over-reporting low effort due to imperfect specificity of measures.

The studies comprising the current thesis do have a number of methodological limitations. First, the absence of litigation for damages following workplace (and other) injury in the NZ context means that it is not clear that the findings from these studies will directly map onto jurisdictions where litigation is required to seek damages following injury. It is not clear that compensation-seeking (insurance and disability payments) and litigation represent equal or different secondary gains that might differentially motivate ETF and malingering in the neuropsychological setting. Certainly, it is clear that compensation-seeking in the form of seeking disability payments and insurance entitlements is associated with a relatively high rate of ETF and estimates of malingering in litigating personal injury cases and those seeking disability have been found not to differ significantly (Mittenberg, Patton, Canyock, & Condit, 2002). Nevertheless, it remains to be established that these different forms of potential gain are indistinguishable in terms of their effect on effort. Arguably, suing for damages is a more combative and hostile process than seeking disability payments and that might impose a specific set of reinforcement schedules to support the emergence of malingering behaviours. Iverson (2003) noted that having a sense of justification, entitlement, frustration, manipulation, greed, and/or neediness might influence exaggeration on cognitive testing. There has been no systematic study of the role that those variables might play in ETF and it is not known whether those variables apply equally to the litigant and to the disability-seeker.

A second limitation of the current research related to the absence of any psychological symptom-validity measurement and the reliance on self-reporting for the determination of psychological symptoms. While the results did demonstrate that *self-reporting* psychological symptomatology, particularly depressive symptoms, is reliably associated with ETF, it remains unknown whether the reporting of psychological symptoms is valid and not another manifestation of a more general phenomenon of symptom over-reporting. As noted in Chapters 5 and 6, there is evidence that psychological symptom over-reporting and low cognitive effort may be independent constructs (Dandachi-Fitzgerald, Ponds, Peters, & Merckelbach, 2011; Demakis, Gervais & Rohling, 2007; Greiffenstein, Gola & Baker, 1995; Haggerty, Frazier, Busch, & Naugle, 2007; Nelson, Sweet, Berry, Bryant, & Granacher, 2007; Ruocco et al., 2005; Whiteside, Dunbar-Mayer & Waters, 2009). Additionally, findings by Jones, Ingram and Ben-Porath (2012) suggest that in those displaying ETF psychological symptom reporting is elevated but not to a level to suggest *over*-reporting. It is recommended that future students of the relationship between self-reported psychological symptoms and ETF include administration of a measure of validity of the former.

The inclusion of both compensation-seeking and non-compensation-seeking participants represents one of the strengths of the current research. That afforded the advantage of being able to examine the predictive role that compensation-seeking has alongside other variables to examine their relative power to predict ETF. However, the sample employed was mostly compensation-seeking (85%). Some caution should be adopted in extending the findings to the non-compensation-seeking population and the present findings in respect to that population should be considered preliminary. Note is again made that the systematic review of the relationship between SRDS and ETF (Chapter 8) examined only compensation-seeking samples and that the conclusions of that review are not necessarily expected to generalize beyond that population.

The use of an archival convenience sample from one private practice and one clinical psychologist for the purposes of the current research introduced the risk of sampling bias. Although similar convenience samples are most commonly seen in effort research, the risk is that unknown variables associated with the practice or the psychologist (e.g., the geographical catchment) bias findings. The large size of this sample somewhat mitigates against this risk as does the general consistency of findings when compared with the literature, however this remains untested. Further, the present sample was relatively heavily represented by those with more severe TBIs than is typically seen in effort research, which has primarily focused on mTBI patients. In the current research, the availability of participants with a wide range of injury severities allowed for appraisal of the relationship between severity and ETF while holding other variables constant; this afforded a useful confirmation of previous findings that have suggested that an inverse relationship exists between effort and injury severity (Green & Iverson, 2001). The relatively large proportion of those with severe TBI seen in this convenience sample is not representative of the epidemiology of TBI in the general population, where approximately 75% of all TBIs are mTBI (Langlois et al., 2003). As such, the predictive model may not reflect the natural history of ETF in prospective samples of consecutive admissions to hospital EDs and it is possible that some sample-dependent findings are evident. It remains to be determined whether the variables found to be predictive of ETF in the current sample, will replicate in another, more representative TBI sample.

As noted in Study 3, difficulty accessing uncorrelated acculturation variables meant that not all acculturation variables of interest could be tested using the logistic regression statistical strategy and entered into the full predictive model. That limitation is inherent to many acculturation variables where, for example, being foreign born correlates highly with years resident in NZ and with years educated in NZ. In the current research selection of which variables to test was guided by the findings of previous literature and rational decision-making processes but it is possible that other untested acculturation variables better explain the variance seen in ETF than those selected.

In instances of regression where the number of observations is low and the number of parameters is high, over-fitting of the data is a risk. In the studies detailed in the present dissertation the risk of over-fitting existed particularly in respect to two variables considered in Study 1 where the number of observations for each variable was small relative to the entire sample (specifically: psychosis (n = 15); BFlor (n = 36)). In each case diagnostic statistics (Menard, 2002) did not reveal any evidence of over-fitting and the findings matched *a priori* expectations on the basis of previous literature and hypotheses, however, it cannot be excluded that some over-fitting did occur across the studies. Independent replication of the findings with another sample is important to ensure the findings are reliable.

ETF has been conceptualised here as failure on effort tests each, of which simulate memory tests. It must be accepted that there is some risk that cases of insufficient effort on cognitive tasks other than those assessing memory were not detected by this methodology. There has been some move to develop effort measures in respect of cognitive domains other than memory (e.g., information processing speed (Tombaugh & Rees, 2000), finger tapping (Arnold et al., 2005)), however the majority of studies to-date have employed the paradigm of memory-styled effort measures and the measures employed here are the most frequently and widely employed effort measures. Futher development of effort measures examining effort across a range of cogntivie domains will allow for the predictive models introduced here to be further developed and refined.

Implications and Directions for Future Research

The current findings contributed to the cognitive effort literature by showing that there is a relationship between psychological variables and ETF that is not solely accounted for by compensation-seeking. Depressive symptom reporting and elevated scores on the BDI-2 are reliably associated with ETF in compensation-seeking patients and clinicians should be particularly vigilant to low effort in those patients suffering from the effects of TBI who are reporting depressive symptoms. Future studies examining the constructs of both depression and ETF might employ a formal measure of psychological symptom over-reporting as a means of extracting the effects of any psychological symptom over-reporting from the effects of ETF.

Although the preliminary statistical model of ETF that is reported here accounted for a significant proportion of the variance seen in ETF, it was clear that there was considerable variance in ETF not accounted for by the variables considered in those studies. There is a need to further broaden the examination of variables that are predictive of ETF to examine other occupational variables such as worker anger and resentment that Silver (2012) notes as problematic within insurance and litigation assessments. Levels of worker satisfaction, work monotony, work stress, and low levels of autonomy/control have been associated with prolonged absence from work following injury or illness (Dragano & Schneider, 2011; Krokstad, Johnsen & Westin, 2002; Kuoppala, Lamminpaa, Vaanen-Tomppo, & Hinkka, 2011) and might be examined as predictive of ETF. Workplace variables including effort-reward imbalance, low job security, and low social support at work have been associated with poor long-term mental health functioning (Wahrendorf et al., 2012) and may have some possible utility in predicting ETF.

Future studies of effort might examine psychological variables that have been identified in the pain literature as predictive of disability including catastrophization and symptom-focus (Sullivan, Bishop, & Pivik, 1995), self-efficacy (O'Leary, 1985), fear of pain and/or re-injury (Waddell, 1993) and external health locus of control (Torres et al., 2009). The abnormal illness behaviours identified here as Behavioural Floridity deserve further investigation in studies that lead to improved operational definition of the behaviours encompassed by that term and potentially, the development of reliable and valid behavioural observation measures of those behaviours. That would broaden and extend the range of tools available for the detection of over-reporting in the TBI population.

Neuroticism and resilience are known to be inversely related (Campbell-Sills, Cohan, & Stein, 2006) and resilience has been posited as a variable that might mediate outcomes in rehabilitation and disability following TBI (White, Driver, & Warren, 2008). The findings of the current studies and others that have examined the relationship between negative affect and ETF (e.g., Lange, Pancholi, Bhagwat, Anderson-Barnes, & French, 2012; Rohling, Green, Allen, & Iverson, 2002; Stulemeijer, 2007; Thomas & Youngjohn (2009); Tsanadis et al., 2008) raise the possibility that neuroticism and ETF might be related. Some examination of the roles of neuroticism and resilience in predicting ETF would be valuable additions to the developing model of ETF described in the current studies.

Acculturation variables might be further examined as predictors of ETF. Geert Hofstede identified five dimensions of cultures (Hofstede, 2001) that have been shown to impact on a wide range of behaviours including the expression of phobic fears (Arrindell et al., 2004), social anxiety (Heinrichs et al., 2006) and illness behaviors (Deschepper et al., 2008). Draguns & Tanaka-Matsumi (2003) also report evidence to suggest that variables related to culture are associated with psychopathology. The current studies have shown that ethnicity is not predictive of ETF once control for being foreign born is carried out but that age at which English was learned was strongly related to ETF. Unfortunately, the Hofstede cultural dimensions have not been adequately defined in South Pacific cultures and as such could not be examined in these studies. However, it is conceivable that other dimensions of culture (e.g., power-distance and collectivism/individualism) impact on effort during cognitive testing following TBI and that might be examined in further research in US or

European contexts.

The systematic review of SRDS and ETF described in Study 4 suggested that in compensation-seeking samples, SRDS is associated with reduced effort. Some examination of that relationship would be valuable in the non-compensation-seeking population to determine whether the findings seen generalize to that population. At this time there is only a relatively small collection of studies of ETF in clearly identified non-compensation-seeking samples and a systematic review of that literature is likely premature. Further original investigations that focus on negative affect in solely non-compensation-seeking samples or that use statistical procedures to partial out the effects of compensation-seeking on the relationship between ETF and SRDS are needed to further clarify that relationship.

The relationship between historical personal abuse and effort remains unclear in the TBI population. Williamson et al., (2012) convincingly showed that abuse not financial reward was associated with ETF in a sample of patients with non-epileptic seizures. That has not been convincingly demonstrated in TBI samples. The current study (Study 2) demonstrated only a relatively weak, statistically non-significant relationship between effort and abuse parameters and a previous study (Donders and Boonstra, 2007) reported that a relationship existed but that relationship was statistically weak and findings were somewhat inconclusive. The research to date, including Study 2 of the current thesis, may have been hampered by dichotomizing the abuse variables, thereby leading to a reduced ability to detect small effects. Further examination of the relationship between abuse and effort is indicated through the collection of data that allows analysis of abuse as a continuous variable (e.g., abuse frequency, abuse duration, specific type of abuse, age of onset of abuse).

Conclusion

The series of studies reported in the current thesis has contributed to the neuropsychological literature regarding effort during cognitive testing by developing a preliminary model of ETF, identifying a range of variables that are associated with ETF and broadening the model of ETF beyond the well-established malingering model of ETF. The findings demonstrate that a range of variables including demographic variables (education, being foreign born, age at which English is learned), psychological variables (self-reported depression, psychosis, behavioural floridity, responses on a self-report measure of depression), occupational variables (workplace accident), financial rewards (compensation-seeking) and injury-variables (mTBI versus severe TBI) are predictive of ETF. Assessors need to remain vigilant to suboptimal effort during cognitive testing. Because the range of variables that are associated with ETF is still only partially understood, effort testing should be undertaken with all patients, not only those for whom there is clear evidence of secondary gains.

References

Arnold, G., Boone, K., Dean, A., Wen, J., Nitch, S., Lu, P., et al. (2005). Sensitivity and specificity of finger tapping test scores for the detection of suspect effort. *The Clinical Neuropsychologist, 19,* 105–120.

Arrindell, W. A., Eisemann, M., Oei, T. P., Caballo, V. E., Sanavio, E., Sica, C., Bages, N., Feldman, L., Torres, B., Iwawaki, S., Hatzichristou, C., Castro, J., Canalda, G., Furnham, A., & van der Ende, J. (2004). Phobic anxiety in 11 nations: Part II. Hofstede's dimensions of national cultures predict national-level variations. *Personality and Individual Differences, 37,* 627-643.

Babikian, T., Boone, K.B., Lu, P., & Arnold, G. (2006). Sensitivity and specificity of various Digit Span scores in the detection of suspect effort. *The Clinical Neuropsychologist, 20,* 145-159.

Bianchini K. J., Curtis, K. L., & Greve, K. W. (2006.) Compensation and malingering in traumatic brain injury: A dose-response relationship? *The Clinical Neuropsychologist, 20,* 831-847.

Boone, K. B. (2007). A reconsideration of the Slick et al., (1999) criteria for malingered neurocognitive dysfunction. In K. B. Boone (Ed.), *Assessment of feigned cognitive impairment: A neuropsychological perspective*. New York: Guilford Publications, Inc.

Brooks, B. L., Sherman, E. M. S., Iverson, G. L., Slick, D. J., & Strauss, E. (2011). Psychometric foundations for the interpretation of neuropsychological test results. In M. R. Schoenberg & J. G. Scott (Eds.), *The little black book of neuropsychology* (pp. 893-922). New York: Springer.

Campbell-Sills, L., Cohan, S. L., & Stein, M. B. (2006). Relationship of resilience to personality, coping, and psychiatric symptoms in young adults. *Behaviour Research and Therapy*, *44*, 585-599.

Dandachi-Fitzgerald, B., Ponds, R. W., Peters, M. J., & Merckelbach, H. (2011). Cognitive underperformance and symptom over-reporting in a mixed psychiatric sample. *The Clinical Neuropsychologist*, *25*, 812-828.

Demakis, G. J., Gervais, R. O., & Rohling, M. L. (2008). The effect of failure on cognitive and psychological symptom validity tests in litigants with symptoms of post-traumatic stress disorder. *The Clinical Neuropsychologist*, *22*, 879-895.

Deschepper, R., Grigoryan, L., Lundborg, C. S., Hofstede, G., Cohen, J., Kelen, G. V., Deliens, L., & Haaijer-Ruskamp, F. M. (2008). Are cultural dimensions relevant for explaining cross-national differences in antibiotic use in Europe? *BMC Health Services Research*, *8*, 123-132.

Donders, J., & Boonstra, T. (2007). Correlates of invalid neuropsychological test performance after traumatic brain injury. *Brain Injury*, *21*, 319–326.

Dragano, N., & Schneider, L. (2011). Work related psychosocial factors and the risk of early disability pensioning: A contribution to assessing the need for rehabilitation. *Rehabilitation*, *50*, 28-36.

Draguns, J. G., & Tanaka-Matsumi, J. (2003). Assessment of psychopathology across and within cultures: issues and findings. *Behaviour Research and Therapy*, *41*, 755-776.

Flach, P. A., Krol, B., & Groothoff, J. W. (2008). Determinants of sick-leave duration: A tool for managers? *Scandinavian Journal of Public Health, 36*, 713–719.

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 (Ed.), *Assessment of feigned cognitive impairment* (pp. 281-309). New York: The Guilford Press.

Green, P., & Iverson, G. L. (2001). Validation of the Computerized Assessment of Response Bias in litigating patients with head injuries. *The Clinical Neuropsychologist*, *15*, 492-497.

Green, P., Iverson, G. L., & Allen, L. M. (1999). Detecting malingering in head injury litigation with the Word Memory Test. *Brain Injury*, *13*, 813-819.

Green, P., Rohling, M. L., Lees-Haley, P. R., & Allen, L. M. (2001). Effort has a greater effect on test scores than severe brain injury in compensation claimants. *Brain Injury*, *15*, 1045-1060.

Greiffenstein, M. F., Gola, T., & Baker, W. J. (1995). MMPI-2 validity scales versus domain specific measures in detection of factitious traumatic brain injury. *The Clinical Neuropsychologist*, *9*, 230–240.

Haggerty, K. A., Frazier, T. W., Busch, R. M., & Naugle, R. I. (2007). Relationships among Victoria Symptom Validity Test indices and Personality Assessment Inventory validity scales in a large clinical sample. *The Clinical Neuropsychologist*, *21*, 917–928.

Heinrichs, N., Rapee, R. M., Alden, L. A., Bögels, S., Hofmann, S. G., Ja Oh, K., & Sakano, Y. (2006). Cultural differences in perceived social norms and social anxiety. *Behaviour Research and Therapy*, *44*, 1187-1197.

Hofstede, G. (2001). *Culture's consequences: comparing values, behaviors, institutions and organizations across nations*. London: Sage Publications.

Inman, T. H., Vickery, C. D., Berry, D. T. R., Lamb, D. G., Edwards, C. L., & Smith, G. T. (1998). Development and initial validation of a new procedure for evaluating adequacy of effort given during neuropsychological testing: The Letter Memory Test. *Psychological Assessment, 10*, 128–139.

Iverson, G. L. (2003). Detecting malingering in civil forensic evaluations. In A. M. Horton & L. C. Hartlage (Eds.), *Handbook of forensic neuropsychology* (pp. 137-177). New York: Springer.

Jones, A., Ingram, M. V., & Ben-Porath, Y. S. (2012). Scores on the MMPI-2-RF scales as a function of increasing levels of failure on cognitive symptom validity tests in a military sample. *The Clinical Neuropsychologist*, *26*, 790-815.

Krokstad, S., Johnsen, R., & Westin, S. (2002). Social determinants of disability pension: a 10-year follow-up of 62 000 people in a Norwegian county population. *International Journal of Epidemiology, 31*, 1183–1191.

Kuoppala, J., Lamminpaa, A., Vaananen-Tomppo, I., & Hinkka, K. (2011). Employee well-being and sick leave, occupational accident, and disability pension a cohort study of civil servants. *Journal of Occupational and Environmental Medicine, 53*, 633-640.

Lange, R. T., Pancholi, S., Bhagwat, A., Anderson-Barnes, V., & French, L. M. (2012). Influence of poor effort on neuropsychological test performance in U.S. military personnel following mild traumatic brain injury. *Journal of Clinical and Experimental Neuropsychology, 34*, 453-466.

Langlois, J. A., Kegler, S. R., Butler, J. A., Gotsch, D. E., Johnson, R. L., Reichard, A. A., Webb, K. W., Coronado, V. G., Selassie, A. W., & Thurman, D. J. (2003). Traumatic brain injury-related hospital discharges: Results from a 14-state surveillance system, 1997. *Morbidity and Mortality Weekly Report*, *52*, 1–20.

Larrabee, G. J. (2003). Detection of malingering using atypical performance patterns of standard neuropsychological tests. *The Clinical Neuropsychologist, 17*, 410–425.

Larrabee, G. J. Greiffenstein, M.F., Greve, K.W., & Bianchini, K. (2007). Refining diagnostic criteria for malingering. In G. J Larrabee (Ed.), *Assessment of malingered neuropsychological deficits* (pp. 334-371). Oxford: Oxford University Press.

Lynch, J. G. (1999). Theory and external validity. *Journal of the Academy of Marketing Science*, *27*, 367-376.

Menard, S. (2002). *Applied logistic regression analysis* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Mittenberg, W., Patton, C., Canyock, E.M., & Condit, D.C. (2002). Base rates for malingering and symptom exaggeration. *Journal of Clinical and Experimental Neuropsychology, 24,* 1094-1102.

Moore, B. A., & Donders, J. (2004). Predictors of invalid neuropsychological test performance after traumatic brain injury. *Brain Injury*, *18*, 975–984.

Nelson, N. W., Sweet, J. J., Berry, D. T., Bryant, F. B., & Granacher, R. P. (2007). Response validity in forensic neuropsychology: Exploratory factor analytic evidence of distinct cognitive and psychological constructs. *Journal of the International Neuropsychological Society*, *13*, 440–449.

O'Bryant, S. E., Finlay, C. G., & O'Jile, J. R. (2007). TOMM performances and self-reported symptoms of depression and anxiety. *Journal of Psychopathology and Behavioral Assessment, 29,* 111–114.

O'Leary, A. (1985). Self-efficacy and health. *Behavior Research and Therapy*, *23,* 437-451.

Rohling, M. L., Green, P., Allen, L. M., & Iverson, G. L. (2002). Depressive symptoms and neurocognitive test scores in patients passing symptom validity tests. *Archives of Clinical Neuropsychology*, *17*, 205-222.

Ruocco, A.C., Swirsky-Sacchetti, T., Daines, D., Chute, D.I., Platek, S.M., Mandel, S., & Zillmer, E.A. (2005). Dissociation between neuropsychological and psychiatric forms of malingering [abstract]. *Archives and Clinical Neuropsychology, 20*, 883-883.

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 (Ed.), *Assessment of feigned cognitive impairment* (pp. 405-427). New York: The Guilford Press.

Schutte, C., & Axelrod, B. N. (2013). Use of embedded cognitive symptom validity measures in mild traumatic brain injury cases. In D. Carone & S. Bush (Eds.), *Mild traumatic brain injury: Symptom validity assessment and malingering* (pp. 159-181). New York: Springer Publishing.

Silver, J. M. (2012). Effort, exaggeration and malingering after concussion. *Journal of Neurology, Neurosurgery & Psychiatry, 83*, 836-841.

Stulemeijer, M., Andriessen, T. M. J. C., & Brauer, J. M. P., Vos, P. E., Van der Werf, S. (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*, 309-318.

Sullivan, M.J.L., Bishop, S.R., & Pivik, J. (1995). The Pain Catastrophizing Scale: Development and validation. *Psychological Assessment, 7,* 524-532.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics.* Boston, MA: Pearson Education Inc.

Thomas, M. L., & Youngjohn, J. R. (2009). Let's not get hysterical: Comparing the MMPI-2 validity, clinical, and RC scales in TBI litigants tested for effort. *The Clinical Neuropsychologist*, *23,* 1067-1084.

Tombaugh, T. N., & Rees, L. (2000). *Manual for the Computerized Tests of Information Processing (CTIP).* Ottawa, Ontario: Carleton University Press.

Torres, V., Collado, A., Arias, A., Peri, J. M., Bailles, E., Salamero, M., & Valdes, M. (2009). Pain locus of control predicts return to work among Spanish fibromyalgia patients after completion of a multidisciplinary pain program. *General Hospital Psychiatry, 31*,137–145.

Tsanadis, J., Montoya, E., Hanks, R.A., Millis, S.R., Fichtenberg, N.L., & Axelrod, B.N. (2008). Brain injury severity, litigation status, and self-report of postconcussive symptoms. *The Clinical Neuropsychologist, 22*, 1080–1092.

Vickery, C. D., Berry, D. T. R., Dearth, C. S., Vagnini, V. L., Baser, R. E., Cragar, D. E., & Orey, S. A. (2004). Head injury and the ability to feign neuropsychological deficits. *Archives of Clinical Neuropsychology*, *19*, 37–48.

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,* 297-313.

Waddell, G. (1993). A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear avoidance beliefs in chronic low-back pain and disability. *Pain, 52*, 157–168.

Wahrendorf, M., Sembajwe, G., Zins, M., Berkman, L., Goldberg, M., & Siegrist, J. (2012). Long-term Effects of Psychosocial Work Stress in Midlife on Health Functioning After Labor Market Exit—Results From the GAZEL Study. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *67*, 471-480.

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.

White, B., Driver, S., & Warren, A. (2008). Considering resilience in the rehabilitation of people with traumatic disabilities. *Rehabilitation Psychology, 53*, 9-17.

Whiteside, D. M., Dunbar-Mayer, P., & Waters, D. P. (2009). Relationship between TOMM performance and PAI validity scales in a mixed clinical sample. *The Clinical Neuropsychologist*, *23*, 523–533.

Williamson, D. J., Holsman, M., Chaytor, N., Miller, J. W., & Drane, D. L. (2012). Abuse, not financial incentive, predicts non-credible cognitive performance in patients with psychogenic non-epileptic seizures. *The Clinical Neuropsychologist*, *26*, 588-598.

APPENDIX A

Table 1

List of Acronyms Employed Throughout the Thesis

ACRONYM	DEFINITION
AACN	American Academy of Clinical Neuropsychology
ASTM	Amsterdam Short Memory Test
BDI-2	Beck Depression Inventory (2nd Edition)
BDI-PC	Beck Depression Inventory for Primary Care
BFlor	Behavioral Floridity
CALD	Culturally and linguistically diverse
CARB	Computerized Assessment of Response Bias
CAT	Critical Appraisal Tool
Comp-seek	Compensation-seeking/litigating
CPT-II	Conners' Continuous Performance Test – 2nd Edition
CTAM	Computerized Test of Attention and Memory
CVLT	California Verbal Learning Test
DCT	Dot Counting Test
DSM	Diagnostic and Statistical Manual
ETF	Effort test failure
ESL	English as a second language
Euro/W	Eusopean/White
FIT	Fifteen Item Test
GCS	Glasgow Coma Scale
HNRB-A	Halstead-Reitan Neuropsychological Battery for Adults
MCMI	Milton Clinical Multiaxial Inventory
MMDS	Malingered Mood Disorder Scale
MMPI	Minnesota Multiphasic Personality Inventory
MR	Mental retardation
MS-TBI	Moderate-to-Severe TBI
MSVT	Medical Symptom Validity Test
mTBI	Mild Traumatic Brain Injury
NAN	National Academy of Neuropsychology

NHMRC	National Health and Medical Research Council
NZ	New Zealand
PAI	Personality Assessment Inventory
PCS	Postconcussion Scale
PCSQ	Postconcussive Symptoms Questionnaire
PDRT	Portland Digit Recognition Test
PNES	Psychogenic Non-Epileptic Seizures
P-SVT	Psychological Symptom Validity Test.
PTA	Post-Traumatic Amnesia
RBANS	Repeatable Battery for the Assessment of Neuropsychological Status
RDS	Reliable Digit Span
RAVLT	Rey Auditory Verbal Learning Test
RMT	Recognition Memory test
SIMS	Structured Inventory of Malingered Symptomatology
SLC-90	Symptom Checklist-90
SRDS	Self-Reported Depressive Symptomatology
STAI	State Trait Anxiety Inventory
STAI-s	State Trait Anxiety Inventory – State Scale
STAI-t	State Trait Anxiety Inventory – Trait Scale
SVT	Symptom Validity Test
TMT	Trail Making Test
TOMM	Test of Memory Malingering
VDS	Vocabulary minus Digit Span
VE	Valid effort
VSVT	Victoria Symptom Validity Test
WAIS-R	Wechsler Adult Intelligence Scale
WCST	Wisconsin Card Sorting Test
WMS	Wechsler Memory Scale
WMS-R	Wechsler Memory Scale – Revised
WMT	Word Memory Test
WMT DR	Word Memory Test Delayed Recall
WMT-IR	Word Memory Test Immediate Recall
WRMT	Warrington Recognition Memory Test

APPENDIX B

Published paper:

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.

The Clinical Neuropsychologist, 2012, 26 (8), 1377–1396
http://www.psypress.com/tcn
ISSN: 1385-4046 print/1744-4144 online
http://dx.doi.org/10.1080/13854046.2012.728248

Psychology Press
Taylor & Francis Group

Effort Test Failure: Toward a Predictive Model

James W. Webb[1], Jennifer Batchelor[1], Susanne Meares[1], Alan Taylor[1], and Nigel V. Marsh[2]†
[1]Department of Psychology, Macquarie University, Sydney, NSW, Australia
[2]Department of Psychology, James Cook University, Singapore

Predictors of effort test failure were examined in an archival sample of 555 traumatically brain-injured (TBI) adults. Logistic regression models were used to examine whether compensation-seeking, injury-related, psychological, demographic, and cultural factors predicted effort test failure (ETF). ETF was significantly associated with compensation-seeking (OR = 3.51, 95% CI [1.25, 9.79]), low education (OR: .83 [.74, .94]), self-reported mood disorder (OR: 5.53 [3.10, 9.85]), exaggerated displays of behavior (OR: 5.84 [2.15, 15.84]), psychotic illness (OR: 12.86 [3.21, 51.44]), being foreign-born (OR: 5.10 [2.35, 11.06]), having sustained a workplace accident (OR: 4.60 [2.40, 8.81]), and mild traumatic brain injury severity compared with very severe traumatic brain injury severity (OR: 0.37 [0.13, 0.995]). ETF was associated with a broader range of statistical predictors than has previously been identified and the relative importance of psychological and behavioral predictors of ETF was evident in the logistic regression model. Variables that might potentially extend the model of ETF are identified for future research efforts.

Keywords: Symptom validity testing; Malingering; Effort; Depression; Psychosis; Illness behavior.

INTRODUCTION

An estimated 1.7 million traumatic brain injury (TBI) related emergency department visits occur each year in the United States, from which 275,000 individuals are hospitalized (Faul, Xu, Wald & Coronado, 2010). Following TBI many people undergo neuropsychological assessment, with those assessments frequently being conducted in order to determine whether or not cognitive impairments exist sufficient to entitle the individual to disability or insurance payments. It is recognized that a substantial proportion (up to 30 40%) of test examinees fail to put forth optimal effort during cognitive testing (Larrabee, 2007). Accordingly there has been a growing interest in objectively assessing effort, and identifying risk factors for, or predictors of, low test-taking effort (Chaftez & Prentkowski, 2011; Dandachi-Fitzgerald, Ponds, Peters, & Merckelbach, 2011; Donders & Boonstra, 2007; Moore & Donders, 2004).

The provision of financial gains via compensation-seeking/litigation for disability has been repeatedly shown to predict effort test failure (ETF) and malingering (Bianchini, Curtis, & Greve, 2006, Henry et al., 2011;

Address correspondence to: James Webb, Department of Psychology, Macquarie University, NSW 2109, Australia. E-mail: james.webb@xtra.co.nz
†Nigel V. Marsh is now at Sunway University, Malaysia. We thank Dr Mike Jones for statistical advice.
Accepted for publication: September 4, 2012. First published online: October 12, 2012.

Paniak et al., 2002). ETF describes suboptimal performance on specific symptom validity tests (SVTs) and embedded cognitive measures that have been validated as tests of effort. It is important to note that ETF does not equate to neurocognitive malingering. Malingering is behavior that is evident in a subset of the group of individuals displaying ETF and, although ETF is essential for the classification of neurocognitive malingering, ETF alone is not sufficient for that classification (Larrabee, Greiffenstein, Greve, & Bianchini, 2007; Slick, Sherman, & Iverson, 1999). Better knowledge of ETF predictors can help explain why some patients present with this response style besides secondary gain.

With the exception of gender, which has very seldom been found to be predictive of ETF, demographic variables of education, age, and ethnicity have been variably associated with ETF. Findings in respect to education have trended toward lower education being predictive of ETF (e.g., Babikian, Boone, Lu & Arnold, 2006; Greve, Etherton, Bianchini, & Curtis, 2009; Mahdavi, Mokari, & Amiri, 2011; Stulemeijer, Andriessen, Brauer, Vos, & Van Der Werf, 2007) and even in those studies that have reported no effect of education there has been a tendency for education to be lower in the group failing SVTs (e.g., Armistead-Jehle, 2010; Ord, Boettcher, Greve, & Bianchini, 2010).

Advanced age is a known predictor of prolonged disability from work following illness (Flach, Krol, & Groothoff, 2008) and as such there are grounds for reasoning that advanced age might be a predictor of ETF. In fact, findings in respect to age have been inconsistent and somewhat sample-dependent. For example, Grote et al., (2000) and Donders and Boonstra (2007) found an age-related relationship with older people more likely to display ETF, while Tombaugh (1997) and others (e.g., Lange, Iverson, Brooks, & Rennison, 2010) have reported no significant relationship.

Similarly inconsistent findings have been found in respect to ethnicity. For example Meyers, Volbrecht, Axelrod and Reinsch-Boothby (2011) found no evidence of ethnicity being a moderator of ETF, while Salazar, Lu, Wen, and Boone, (2007) found a significant impact of ethnicity on ETF. Caucasians scored significantly higher than Hispanics on the embedded measures of Digit Span age corrected scaled score and Reliable Digit Span, and higher than African Americans on Rey AVLT and Rey-Osterreith embedded measures. Salazar et al. found that in English as a second language groups effort scores were related to the age at which English was learned, but years living in the USA or years educated in the USA were not related to effort test scores. Foreign-born immigrant status, independent of ethnicity, has not been specifically investigated.

Psychological factors have been found to have a complex association with ETF. Although earlier researchers (e.g., Ashendorf, Constantinou, & McCaffrey, 2004; Rees, Tombaugh, & Boulay, 2001) found no relationship between depression and ETF in non-litigating psychiatric participants, other more recent studies including non-litigating psychiatric patients (Dandachi-Fitzgerald et al., 2011; Gorissen, Sanz, & Schmand, 2005), compensation-seeking depressed participants (Green, 2009), non-litigating, non-compensation seeking TBI participants (Bierley et al., 2001), and compensation-seeking neurological participants (Armistead-Jehle, 2010; Rohling, Green, Allen, & Iverson, 2002; Stulemeijer et al., 2007; Suhr, Tranel, Wefel, & Barrash, 1997), have detected a relationship, but Schroeder and Marshall (2011) did not. Schroeder and Marshall (2011) have suggested that reliance on single SVT failure to indicate ETF might account for some of the inconsistent findings in this field and this will be further examined in the present study. Psychosis is a factor that has been shown to be associated with ETF (Goldberg, Back-Madruga, & Boone, 2007), although again, Schroeder and Marshall (2011) have questioned that finding.

There have been very few studies of the relationship that might exist between interview behaviors and test-taking effort. Greiffenstein and Baker (2006) drew on Miller's (1961a, 1961b) landmark lectures on *accident neurosis*, by investigating a behavioral component of test-taker effort. They found that behavior, in the form of the *floridity* (frequency) of symptom reporting, was related to ETF such that higher symptom floridity predicted ETF. A similar relationship between over-reporting and ETF was reported by Dandachi-Fitzgerald and co-workers (2011).

Injury-related variables have been examined as predictors. Specifically, many researchers have noted a paradoxical finding of better effort test performance among those with more severe brain injuries (Carone, 2008; Green, Iverson, & Allen, 1999; Green, Rohling, Lees-Haley, & Allen, 2001; West, Curtis, Greve, & Bianchini, 2011). Having a workplace accident has not been specifically investigated as a correlate of poor effort independent of compensation-seeking/litigation. Workplace injuries are associated with an increased risk of prolonged disability, blaming others for the injury, post-traumatic stress, and litigation (Mason, Wardrope, Turpin, & Rowlands, 2002); this might afford a potential context for ETF during neuropsychological evaluation.

Overall the literature to date has identified a number of variables that increase the probability or likelihood of ETF but with few exceptions (Donders & Boonstra, 2007; Moore & Donders, 2004) these variables have seldom been examined concurrently using multivariable statistical techniques. The aim of the current study was to examine predictors of ETF in an archival sample of consecutive referrals to a private clinical neuropsychology practice. Based on the literature it was anticipated that compensation-seeking (Bianchini et al., 2006), demographic and psychological variables would predict ETF. Specifically, it was hypothesized that age (Flach et al., 2008), education (Babikian et al., 2006), ethnicity and immigration status (Salazar et al., 2007) would each be associated with ETF. In addition, a self-reported mood disorder (Green, 2009), displaying psychotic illness (Goldberg et al., 2007), and displaying exaggerated symptomatic or behavioral floridity (Greiffenstein & Baker, 2006) were anticipated to predict ETF. Finally, it was hypothesized that injury variables including having sustained a workplace accident (Mason et al., 2002), and that ETF would be more likely in those with mTBI than more severe brain injuries (Green et al., 1999).

METHOD

Participants

An archival sample of 555 consecutive referrals to a private clinical neuropsychology practice in Auckland, New Zealand was examined. Participants included had sustained a TBI and were over the age of 16 years at the time of assessment.

The data set was collected over the period 2001 to 2007 inclusive. Participants were excluded on the basis of having a pre-existing history of mental retardation or dementia.

Compensation-seeking

Most of the sample ($n = 470$; 84.7%) were seeking compensation continuance or seeking entitlement to compensation. Compensation was defined as worker's compensation income replacement (insurance) payments or disability social security benefits. None of the sample was engaged in litigation (litigation for damages is specifically precluded under New Zealand no-fault accident compensation legislation). The majority of the compensation-seekers ($n = 422$; 90%) were seeking continuance of worker's compensation payments while 46 (10%) were seeking continuance of social security benefits. A sizeable minority ($n = 85$; 15%) of the total sample were ineligible for compensation.

Demographic variables

Men made up 72.8% of the sample. The mean age was 41 ($SD = 12.31$, range = 16 76) years. Ethnicity was as follows: European/White; ($n = 418$, 75.3%), Maori/Pacific Island ($n = 105$, 18.9%), Indo/Asian ($n = 32$, 5.8%). The foreign-born subgroup consisted of any non-New Zealand born individuals and comprised 18% ($n = 99$) of the sample. The sample had, on average, 11.8 ($SD = 2.5$, range = 2 21) years of education.

Injury variables

Severity of injury was classified on the basis of Glasgow Coma Scale (GCS) scores at the Emergency Department, duration of post-traumatic amnesia (PTA) assessed using the Westmead Post-traumatic Amnesia (PTA) scale (Shores, Marosszeky, Sandanam, & Batchelor, 1986), and duration of loss of consciousness (LOC). When PTA data were unavailable, duration of PTA was assessed via clinical interview, which sought to establish the onset of continuous recall (Gronwall & Wrightson, 1980). PTA duration was estimated from clinical interview alone in 7.5% ($n = 42$) of cases where there had been no medical attention at the time of injury. In all cases severity was estimated as mTBI. Comparing the group with mTBI injury classified from interview alone with the mTBI group classified from interview and documented medical data (i.e., GCS, PTA, LOC) ($n = 225$) revealed no differences in respect of age, gender, years of education, or classification with ETF.

Duration of loss of consciousness (LOC) was assessed in accordance with the guidelines of Ruff et al. (2009), specifically: that the duration of LOC should result from impact, not from other medical causes, and that LOC was determined from collateral reports of witnesses present at the scene (e.g., paramedic) or from hospital medical records, not from self-report of the participant.

Severity of mild TBI (mTBI) was defined according to the WHO Collaborating Task Force mTBI diagnostic criteria (Carroll, Cassidy, Holm, Kraus, & Coronado, 2004). Complicated mTBI was differentiated from mTBI according to Williams, Levin and Eisenberg (1990). Moderate to severe TBIs were defined using the Teasdale and Jennett (1974) and Teasdale (1995) criteria.

TBI severity was classified as shown in Table 1. The mTBI-complicated and Moderate TBI groups were combined to ensure that parameter estimates were based on adequate numbers of cases and that there were no empty cells. This approach is supported by the findings of Kashluba, Hanks, Casey, and Millis (2008) who found that few differences in outcome are seen in cases of mTBI-complicated injuries and moderate TBI. Additionally, preliminary analyses that showed that the mTBI-complicated and Moderate TBI groups did not differ in respect of their predictive relationships with ETF. The Very Severe TBI and Extremely Severe TBI groups were similarly combined on the basis that preliminary analyses revealed that the two groups did not differ in respect of their predictive relationships with ETF.

Measures

All participants completed three measures of effort including one forced-choice symptom validity test. All participants completed the Test of Memory Malingering (TOMM; Tombaugh, 1996), the Fifteen Item Test (FIT; Rey, 1964), and Reliable Digit Span (RDS; Greiffenstein, Baker, & Gola, 1994).

Effort classification

Effort was classified dichotomously: valid effort (VE) or ETF. Cut-offs employed for each measure were set to ensure >90% specificity following Baker, Donders, and Thompson (2000) and Boone, Salazar, Lu, Warner-Chacon, and Razani (2002). Consequently the Greve, Bianchini, and Doane (2006) cut-off of <47 was employed for TOMM2 and TOMM Retention; An RDS cutoff of <8 was employed according to guidelines of Suhr and Barrash's (2007) review; An FIT cutoff of <9 was used in accordance with the findings of Boone et al. (2002). ETF was operationalized as failure on any two measures in accord with the findings of Victor, Boone, Serpa, Buehler, and Ziegler (2009) or below chance performance on either TOMM2 or TOMM Retention (<18/50 at the 95% confidence interval using

Table 1. Injury severity criteria and sample characteristics

Descriptor	Criteria	*n* (%)
mTBI	LOC < 30 mins, PTA < 24 hours, GCS 13–15 at 30 mins	275 (49.5)
mTBI complicated	LOC < 30 mins, PTA < 24 hours, GCS 13–15 at 30 mins, visible (on CT brain imaging) intracranial abnormality not requiring surgery	38 (6.8)
Moderate	PTA 1–24 hours, GCS 9-12	54 (9.7)
Severe	PTA 1–7 days, GCS 3-8	79 (14.2)
Very severe	PTA 1–4 weeks, GCS 3-8	69 (12.4)
Extremely severe	PTA >4 weeks, GCS 3-8	40 (7.2)

LOC = Loss of consciousness; mTBI = mild traumatic brain injury; PTA = post-traumatic amnesia; GCS = Glasgow Coma Scale.

the binomial distribution). VE was classified using the requirement that participants pass all three effort measures at the cutoffs described above.

Psychological dimensions

Psychological data were available from the archive records for each participant. All participants had undergone an approximately 60-minute semi-structured clinical interview with a licensed psychologist trained in clinical psychology (JW). Participants reporting a mood disorder and/or a psychotic disorder and who met the Diagnostic and Statistical Manual of Mental Disorders IV-TR diagnostic criteria (American Psychiatric Association, 2000) were coded as having a self-reported mood disorder or self-reported psychosis.

In a subsection of the sample exaggerated displays of behavior had been noted and were coded as Behavioral Floridity (BFlor). BFlor conceptually builds on the work of Miller (1961a, 1961b; Miller & Cartlidge, 1972) and Greiffenstein and Baker (2006). Miller's original lectures (1961a, 1961b) described other aspects of abnormal behavior seen in this population. He described florid behaviors: "gross dramatization of symptoms," (p. 922) including extreme behaviors such as "groaning," "slumping forward with head in hands," "quivering," and using what might now be described as catastrophizing language when symptom-reporting ("terrible," "agonizing"). Building on Miller's papers, BFlor is defined here as extreme displays of symptom-related behavior and no assessment of internal states including cognitive styles, beliefs, perceptions, etc., is implied.

For the purposes of this exploratory study, dichotomous coding of BFlor was made by clinical judgment of the principal author. A conservative approach was taken toward coding BFlor to ensure that only extreme and disproportionate displays of behavior comprised BFlor. Examples of BFlor included lying on the floor with complaints of extreme tiredness following interview, requiring each and every question to be repeated, dry-retching while reporting extreme fatigue, atypical levels of symptom endorsement, extreme slowness of all movements, dramatic facial displays of tiredness, pathos, or marked affective blunting in the absence of affective/psychotic disturbance, atypical displays of language use, e.g., answering yes/no in German language despite the participant having never been fluent in the German language; missing the first spoken phoneme from each word. Because only the most extreme forms of behavior were considered to be BFlor, relatively few participants were coded as such ($n = 36$, 6.5%).

The participants with diagnosed psychotic illness ($n = 15$) were all male, seven of whom were diagnosed with a psychotic disorder due to a general medical condition (TBI) with symptom onset ranging from emergence from PTA to 6 years post-injury. Three were diagnosed with schizophrenia with age of onset ranging from late teens to early twenties and pre-dating their injuries. Two were diagnosed with delusional disorder with onset in late teens and early twenties and before injury. Two were diagnosed with schizoaffective disorder, one developing following immigration 9 years before injury and one developing 4 months before TBI. One was differentially diagnosed with schizophrenia/substance-induced psychosis with symptom onset before injury. All participants but one were under psychiatric review, four were taking no psychotropic medications, the remainder taking atypical antipsychotics. Four participants were taking anticonvulsant medication for seizure disorders. Nine had been hospitalized for assessment and treatment, eight of which were involuntary committals. In seven participants injury was considered to be the precipitant for developing psychotic symptoms and in those, all injuries were classified as moderate to very severe. All but one of the participants were receiving compensation/compensation-seeking.

Ethical approval

This study was approved by the Human Research Ethics Committees of Macquarie University and the University of New England.

RESULTS

Of the 555 participants, 111 (20.0%) were classified with ETF. Of these, seven participants (1%) scored below chance on the TOMM trials. A total of 352 participants were classified with VE.

Preliminary analyses

Table 2 shows the results of exploratory bivariate statistical analyses. Participants were grouped according to ETF status and comparisons across the variables of interest were undertaken. Because the bivariate analysis was exploratory no adjustment for multiple comparisons was undertaken.

Variables that were noted to have small to moderate statistically significant positive relationships with ETF included age, being foreign-born, ethnicity, having a workplace accident, compensation-seeking, having a diagnosed (self-reported) mood disorder, having a diagnosed (self-reported) psychotic illness, and displaying BFlor. Variables with a small to moderate inverse relationship with ETF included years of education, and TBI severity. All relationships with variables in bivariate analyses were statistically significant and as such all were included for further analysis.

Bivariate relationships between the predictors that were identified by the preliminary analysis were calculated and are displayed in Table 3. Most relationships between the statistical predictors were relatively weak, ranging from .003 to −.21. Having a mood disorder was noted to correlate with BFlor (Gamma = .63, $p < .001$) and with having sustained a workplace accident (Gamma = .31, $p < .01$), and was negatively correlated with TBI severity (Gamma = −.36, $p < .001$). BFlor was also negatively correlated with TBI severity (Gamma = −.61, $p < .001$) and was positively correlated with having sustained a workplace accident (Gamma = .51, $p < .05$) and with being compensation-seeking (Gamma = .74, $p < .001$). Compensation-seeking was not significantly correlated with any other variable.

Predictive model

Logistic regression analyses with ETF as the dependent variable were undertaken. Years of education was included as a continuous variable.

Table 2. Characteristics of the participants grouped according to effort

Variable	Good effort *M*	Good effort *SD*	ETF *M*	ETF *SD*	*p*	Effect Size (*d*/φ)
Age	40.65	13.16	43.68	10.57	.01	.25
Education, years	12.09	2.46	11.19	2.52	.001	.36
	N	%	*n*	%		
Ethnicity					.004	.15
European/White	281	79.4	73	20.6		
Maori/Pacific Island	56	68.3	26	31.7		
Indo/Asian	15	55.6	12	44.4		
Immigration status					<.001	.27
NZ born	309	81.5	70	18.5		
Foreign-born	43	51.2	41	48.8		
Self-reported Mental Illness						
Affective disorder					<.001	.36
No	269	86.8	41	13.2		
Yes	83	54.2	70	45.8		
Psychosis					<.001	.17
No	347	77.3	102	22.7		
Yes	5	35.7	9	64.3		
Behavioral Floridity					<.001	.34
No	345	79.9	87	20.1		
Yes	7	22.6	24	77.4		
TBI Severity					<.001	.20
mTBI	164	68.9	74	31.1		
Moderate	61	78.2	17	21.8		
Severe	52	78.8	14	21.2		
Very Severe	75	92.6	6	7.4		
Place of accident					<.001	.28
Non-workplace	312	81.5	71	18.5		
Workplace	40	50.0	40	50.0		
Comp.seek/Continuance					<.001	.17
No	68	93.2	5	6.8		
Yes	284	72.8	106	27.2		

p-values are from independent *t*-tests or Chi-Square tests. Effect sizes for *t*-tests are Cohen's *d* and Cramer's Phi (φ) for Chi-Square tests. Indo/Asian = participants reporting themselves of Indian or Asian ethnicity; NZ = participants reporting themselves New Zealand born; mTBI = mild traumatic brain injury; BFlor = behavioural floridity; Comp. seeking/Continuance = compensation seeking or seeking continuance of compensation.

Dichotomous variables included immigrant status, presence of a self-reported mood disorder and self-reported psychotic illness, displays of BFlor, having a workplace accident, and compensation-seeking. Injury severity and ethnicity were analyzed as categorical data.

The fit of the model was examined in three ways. First, examination of residuals revealed 11 cases with large standardized residuals. Those cases were eliminated and the same results were obtained and for this reason all cases were

Table 3. Bivariate relationships between main predictors

	Age	Yrs ed	Foreign-born	Mood disorder	Psychotic illness	BFlor	TBI Severity	Wrkplc Acc	Comp-seek
Yrs ed	−.04								
Foreign-born	.04	.15**							
Mood disorder	.11*	.01	.33**						
Psychotic illness	−.03	−.12**	.26	−.34					
BFlor	−.01	−.03	.43*	.63***	−1.00**				
TBI severity	−.06	.04	−.05	−.36***	.25	−.61***			
Wkplc Acc	.15***	−.07	.09	.31**	−1.00***	.51*	−.48***		
Comp-seek	.08	−.06	.27	.01	.44	.74***	.10	.37	
Ethn[1]	.13	.19***	−.06	−.05	.34	.34	.03	−.06	.19

Relationships among dichotomous variables are expressed as Gamma coefficients. Relationships between continuous/ordinal variables and dichotomous variables are expressed as Pearson product-moment correlation coefficients/point-biserial correlation coefficients. [1]Relationships between Ethnicity (categorical data) and Age, Years education, and Severity are expressed as r (R^2) following one-way ANOVAs employing the continuous variables as dependent variables and Ethnicity as factor. Yrs ed = years of education; BFlor = behavioral floridity; TBI = traumatic brain injury; Wrkpl Acc = workplace accident; Comp-seek = compensation-seeking or seeking continuance of compensation; Ethn = ethnicity.
$^*p < .05$. $^{**}p < .01$. $^{***}p < .001$.

included in the analyses. Second, the fit of possible alternative models which might have been appropriate in the light of the relatively skewed binary dependent were examined. The fit of a model with a complementary log-log link was trivially worse than the original model with a log link (AIC = 360.3, BIC = 418.2 versus AIC = 359.3 and BIC = 417.2 for the original model). A model with a log-log link fitted markedly less well (AIC = 369.1, and BIC = 427.0). The original model was therefore retained. Third, multicollinearity was examined using a method described by Belsley, Kuh, and Welsch (1980). A program written for Stata (Statacorp, 2011) by Hardin (1995) provided condition numbers and variance decomposition proportions based on the singular value decomposition of the X matrix of independent variables. One of the condition numbers (19.5), while lower than the highest criterion of 30 suggested by Belsley et al. (1980), was higher than the next-lowest criterion, 15. Examination of the variance decomposition proportions showed that, as might be expected, age and years of education showed reasonably high collinearity. However, tests of reduced models showed (1) that age was never significant, adjusted or unadjusted for years of education, and years of education was significant whether or not it was adjusted for age, and (2) that the significance of other effects were very similar whether both or either of the variables were included in the model. The original model, which included both age and years of education, was therefore retained.

As presented in Table 4 a number of statistical predictors of ETF were positively identified as making unique contributions (i.e., with all other variables held constant) in the final logistic regression model. Unadjusted odds ratios are provided in Table 4 for comparison with the adjusted odds ratios.

Table 4. Logistic regression analysis of effort test failure as a function of demographic and injury predictors

Variable	Adjusted OR (95% CI)	*p* value	Unadjusted OR (95% CI)	*p* value
Age	1.01 [0.98, 1.03]	.36	1.01 [1.002, 1.03]	.02
Years education	.83 [0.74, 0.94]	.005	.85 [0.77, 0.94]	.001
Foreign-born	5.10 [2.35, 11.06]	<.001	3.90 [2.43, 6.26]	<.001
Ethnicity[1]		.19		.005
Pac.Isl vs Euro/W	1.85 [0.92, 3.69]	.08	1.78 [1.05, 3.04]	.03
IndoAs. vs Euro/W	.90 [0.26, 3.09]	.86	3.07 [1.38, 6.86]	.006
Mood disorder	5.53 [3.10, 9.85]	<.001	5.53 [3.50, 8.74]	<.001
Psychotic illness	12.86 [3.21, 51.44]	<.001	6.12 [2.00, 18.67]	.001
BFlor	5.84 [2.15, 15.84]	.001	13.59 [5.67, 32.58]	<.001
TBI Severity[2]		.25		.001
Moderate vs Mild	.87 [0.40, 1.86]	.71	.61 [0.33, 1.12]	.11
Severe vs Mild	.41 [0.30, 1.63]	.41	.59 [0.31, 1.14]	.12
V.Severe vs Mild	.37 [0.13, 0.995]	.04	.17 [0.07, 0.42]	<.001
Workplace Accident	4.60 [2.40, 8.81]	<.001	4.39 [2.64, 7.30]	<.001
Compensation-seeking	3.51 [1.25, 9.79]	.01	5.07 [1.99, 12.93]	.001
Constant	.09	.03		

Flor = behavioral floridity; TBI = traumatic brain injury; Pac.Isl = Pacific Island; Euro/W = European/White; IndoAs = Indo-Asian; V.Severe = very severe TBI. [1]Ethnicity was dummy coded with European/White as the reference group. [2]Severity was dummy coded with mild traumatic brain injury as the reference group.

Demographic variables

Age was significantly related to ETF in bivariate analysis but age did not make a significant contribution when adjusted for the effect of the other variables in the model. Years of education was a significant predictor of ETF such that a 1-year increase in education reduced the odds of ETF by .83 (or by 17%).

Compensation-seeking

Compensation-seeking was a significant statistical predictor of ETF in both preliminary bivariate analysis and when adjusted for the effect of the other variables in the model: The odds of ETF in the compensation-seeking sample are 3.5 times those for the non-compensation-seeking sample.

Ethnicity

Unadjusted odds ratios suggested that ethnicity was predictive of ETF, but ethnicity was not significantly related to ETF when included in the final model that included being foreign-born, which was a more powerful predictor of ETF. In comparison with being locally born, being foreign-born increased the odds of ETF by five times.

Mental state variables

Having a self-reported mood disorder, psychotic illness, and the display of BFlor were predictive of ETF, both when considering unadjusted odds ratios and in the final model. Self-reporting mood symptoms increased the odds of ETF by five times relative to those without a formally diagnosed mood disorder, and having a psychotic disorder increased the odds of ETF by about 13 times. The odds of ETF were almost six times higher in those displaying BFlor than those not displaying BFlor.

Injury-related variables

Severity of TBI was a significant predictor of ETF when considering the unadjusted odds such that individuals who had sustained severe or very severe injuries were less likely to display ETF than those with milder injuries, but TBI severity was not a significant predictor when adjusted for the effect of other variables in the model. Having sustained a workplace accident was predictive of ETF when examined in a bivariate logistic regression model and also in the multivariable model. Having a workplace accident increased the odds of ETF by about 4.5 times relative to those not having a workplace accident.

Goodness of fit

Goodness of fit of the final logistic regression model including all predictors was assessed via the Hosmer and Lemeshow test, $\chi^2(8, N = 463) = 12.19$, $p = .14$. The area under the receiver operating characteristic curve for the model was .87 (95% CI [.82, .91]), indicating excellent discrimination in predicting those with ETF and those without (Hosmer & Lemeshow, 2000) and Nagelkerke $R^2 = .48$.

DISCUSSION

The aim of this study was to examine potential correlates of ETF and to gain an impression of the relative importance of predictors of ETF toward identifying a model of ETF in a traumatic brain injury sample. It was hypothesized that compensation-seeking, demographic, psychological/behavioral, and injury variables would predict ETF. The predictive relationships identified are statistical in nature not causal.

This study found that 20% of the total sample displayed ETF. This is a relatively low proportion compared with the previous literature. This finding is likely due to this study including a higher proportion of participants with severe injuries than is typically seen in this literature and that ETF is more likely in those with mild injuries than severe injuries (West et al., 2011). Also, the proportion of compensation-seeking participants was relatively lower than that which is typically seen and there is a known relationship between compensation-seeking and ETF (Bianchini et al., 2006). In the compensation-seeking participants with mTBI 44% of participants failed one effort measure and 30% failed two or more. Additionally, although the measures employed here are among the most frequently employed measures of effort, other sensitive measures (e.g., Word Memory Test; Green, 2003) were not included and this may have contributed to the factors described above and account for this apparently low proportion of ETF. Finally, this study has required failure on more than one effort measure for classification of ETF; 37% of the sample failed one effort measure a rate that is similar to that seen in other effort studies in civil litigation settings (Larrabee, 2003a). However, Victor and colleagues (2009) showed that sensitivity to ETF appears to fall when increasing the number of effort measures employed, thus reducing the apparent proportion of those displaying ETF, but specificity and positive predictive power increases, making findings safer and more stable.

This study confirmed that compensation-seeking was significantly predictive of ETF. This finding was consistent with a number of previous studies that have found a significant main effect of compensation-seeking on effort and symptom exaggeration (e.g., Bianchini et al., 2006; Henry et al., 2011; Paniak et al., 2002). A number of previous researchers have independently reported that when incorporated into a predictive model with other variables, compensation-seeking is not significantly predictive of ETF (Donders & Boonstra, 2007; Ross, Putnam, & Adams, 2006; Stulemeijer et al., 2007) or only modestly predictive (Moore & Donders, 2004). Those previous findings in combination with the present results suggest that holding other potentially predictive variables constant by using a multivariable research strategy can help to clarify the predictive power of one variable (e.g., compensation-seeking). As new predictors become identified sample-dependent findings will become less problematic. Furthermore this emphasizes that the factors that underpin ETF may be multiple and that the search for other predictors may be a fruitful exercise.

The study found that having sustained a workplace accident was a predictor of ETF independent of compensation-seeking status. Having a workplace accident has previously been reported to be associated with adjustment difficulties and affective disturbance (Mason et al., 2002) but in this sample having a workplace accident was only weakly related to developing a self-reported mood disorder. This suggests that there are non-affective drivers of ETF in those who have suffered workplace accidents. Occupational variables such as low worker satisfaction, work monotony, work stress, and low levels of autonomy/control are associated with prolonged disability following injury and illness (Dragano & Schneider, 2011; Krokstad, Johnsen & Westin, 2002; Kuoppala, Lamminpaa, Vaanen-Tomppo, & Hinkka, 2011) and their relationship with ETF may prove worthy of further examination.

In contrast with some previous studies that failed to find a relationship between psychiatric disturbance and ETF (Schroeder & Marshall, 2011), this study found a significant relationship between self-reported mood disorder and ETF, and between displaying psychotic illness and ETF. This study supports previous findings with clinical samples, including non-compensation-seeking samples that have found a relationship between psychiatric disturbance and ETF (e.g., Dandachi-Fitzgerald, et al., 2011; Gorissen et al., 2005; Rohling et al., 2002; Stulemeijer et al., 2007). Schroeder and Marshall (2011) found no significant ETF in a non-compensation-seeking psychiatric sample, and suggested that previous findings of low effort in psychiatric samples might partially be an artifact of reliance on one

SVT to diagnose poor effort. This thesis is not supported by the present study, which required failure on two or more of three specific and embedded measures.

Reporting of mood symptoms and psychosis in this study might be considered an aspect of heightened symptom endorsement, which is known to be related to ETF (Greiffenstein & Baker, 2006). Boone and Lu (1999), Larrabee (2003b), Mooney, Speed and Sheppard (2005), Armistead-Jehle (2010), and Jones, Ingram, and Ben-Porath (2012) found heightened psychological symptom reporting but no significant over-reporting in those displaying ETF. Repeated factor-analytic studies using a variety of personality measures have found that emotional over-reporting and ETF represent independent constructs (Jones & Ingram, 2011; Nelson, Sweet, Berry, Bryant, & Granacher, 2007; Ruocco et al., 2005; Whiteside, Dunbar-Mayer, & Waters, 2009). These findings are further supported by the data of Sumanti, Boone, Savodnik, and Gorsuch (2006), Demakis, Gervais, and Rohling, (2008) and Dandachi-Fitzgerald et al. (2011). Thus, while some individuals might over-report both cognitive and psychological symptoms, the findings of Nelson et al. (2007) and others indicate that this is unlikely to occur throughout a large sample such as in the present study. In respect of those displaying psychotic illness in the current study, clinical file review shows that the psychotic individuals typically had long and well-documented histories of psychosis that pre-dated their injuries and/or they had sustained severe and unambiguous brain injuries and several had come under involuntary committal to receive treatment simple fabrication of their psychotic symptoms appeared unlikely.

In the present study the display of exaggerated behavior (BFlor) was closely related to ETF. This finding indicates that some patients signal their likelihood to display ETF via exaggerated illness-related behaviors in the session and supports previously related research on the behaviors that are associated with ETF (Dandachi-Fitzgerald et al., 2011; Greiffenstein & Baker, 2006). This is an exploratory study only and has not attempted to tightly operationalize the abnormal behaviors of interest, but it is hoped that this study might spur future research, allowing better operational definition of the abnormal behaviors seen in this population. Ekman and co-workers have commented comprehensively on the facial behavioral displays seen in lying and in malingering specifically (Ekman & O'Sullivan, 2006) and this study supports their position that symptom exaggeration is detectable through overt behavior.

BFlor was found to be closely related to compensation-seeking status and to self-reporting a mood disorder. The shared variance between these variables raises the possibility of a common factor of symptom over-reporting being present in this group. A measure of psychological symptom over-reporting (e.g., MMPI-2 Fp) was not included in this study and future studies examining BFlor may well investigate this possibility.

Ethnicity has been examined relatively infrequently as a predictor of ETF. Most previous studies have found no effect of ethnicity (Armistead-Jehle, 2010; Inman et al., 1998; Lange et al., 2010; Meyers et al., 2011), although Salazar et al., (2007) reported a significant effect of ethnicity and suggested separate effort test cut-off scores for different ethnicities. The current study found that, with the exception of being foreign-born, ethnicity was not predictive of cognitive symptom exaggeration. These findings support previous results of no impact of ethnicity on effort and suggest that ETF is more attributable to the foreign-born status of an individual than their race. There is considerable social psychology experimental and survey evidence that shows that, as social distance increases and group identification decreases, self-interested behavior increases and ethical/fair behavior becomes less likely (e.g., Hoffman, McCabe, Shachat, & Smith, 1996; Wenzel, 2004). These findings of reduced test-taking effort in foreign-born participants may be a reflection of this broader social psychology phenomenon.

Previous studies have shown that advancing age is associated with prolonged disability from work following illness/injury (Bruusgaard, Smeby, & Claussen, 2010; Flach et al., 2008), and it was hypothesized that age would be associated with ETF. A relationship between age and ETF was evident when age was examined independently such that each year of age increased the odds of ETF by 1%; however age was not a significant predictor in the multivariable model.

As noted above, education was found to be inversely related to ETF, such that a 1-year increase in years of education decreased the odds of ETF by 17% and this supports the majority of previous findings in this area that show that lower education is associated with greater risk of ETF (Babikian et al., 2006; Greve et al., 2009; Mahdavi et al., 2011; Stulemeijer et al., 2007).

TBI severity has been previously reported to be related to ETF (Green & Iverson, 2001, Greiffenstein & Baker, 2006; West et al., 2011). The present study supported previous findings and showed that those with the most severe injuries are less likely than those with mTBI to display ETF. It is worth noting, however, that severity of injury was one of the weakest statistical predictors of effort in the entire model and that some 11% of those with severe or extremely severe injuries displayed ETF. As such, effort testing should not be reserved only for those with mild injuries.

A limitation of this study was the archival convenience sample. This methodology risks introducing sampling bias, in this case toward the compensation-seeking population. As such, caution should be adopted in extending the conclusions of this study beyond that population. Also, this sample may over-represent more severely injured and chronically disabled individuals. These findings, however, may extend generalizability beyond the malingering population of mTBI that has traditionally been examined.

Second, because New Zealand precludes litigation for damages following injury, this study cannot assess the relative importance of litigation versus compensation-seeking in predicting ETF. It may be that the adversarial nature of litigation is a factor that increases the risk of ETF. It may also be that the larger, more immediate, secondary gain that is achieved by successful litigation has relatively greater potential for increasing ETF than the temporally distant secondary gains that are seen in worker's compensation and social security claims. Such a relationship would be consistent with the findings of the behavioral economics literature on temporal discounting of rewards (e.g., Killeen, 2009). An analysis of this distinction between different forms of secondary gains would be a useful addition to the preliminary model of ETF described here.

Another limitation was the use of clinical interview alone to judge self-report of mood disorder and the presence of BFlor, without the use of formal psychometrics. DSM diagnostic criteria were strictly adhered to in determining self-reported mood disorders, and there is no accepted operational definition of

BFlor, nor are there psychometrically sound behavioral measures of illness display. Behavioral observation techniques are required to further this direction of study, and would add a new category of tools for the detection of symptom exaggeration.

Although this study employed a large sample, logistic regression techniques are demanding of a high ratio of cases to variables and this is calculated based on the smaller of the two groups (Harrell, 2001). This study employed unbalanced groups (VE versus ETF), somewhat intrinsic to the subject matter, and consequently the case-to-variable ratio is somewhat low. Diagnostic procedures have not detected any sign of marked over-fitting but a degree of over-fitting is a possibility, particularly given the relatively low numbers of participants displaying BFlor and psychotic symptomatology.

Although the predictive model was statistically and clinically significant and showed excellent discriminative power, the model did not account for all the observed variance in effort in this sample. This means that there are substantial, to-date unexamined, predictors of effort that are not being detected by this model or by the extant literature. As noted above, workplace and occupational variables present as potentially important variables for future investigation. The pain and health psychology literature has examined such variables as catastrophization (Sullivan, Bishop, & Pivik, 1995), self-efficacy (O'Leary, 1985), external health locus of control (Torres et al., 2009), fear of pain and fear of re-injury (Waddell, 1993), as predictors of disability similar variables may also prove important predictors of effort in the neuropsychological domain. Low resilience (Seligman & Csikszentmihalyi, 2000) has also been posited as a driver of disability following brain injury (White, Driver, & Warren, 2008).

The findings of the present study indicate that compensation-related, injury-related, demographic, psychological, and behavioral factors are statistical predictors or correlates of ETF. The picture of ETF appears more complex than has been seen previously. There is a need to investigate other variables that are potentially associated with low test-taker effort.

REFERENCES

American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev. ed.). Washington, DC: American Psychiatric Association.

Armistead-Jehle, P. (2010). Symptom validity test performance in U.S. Veterans referred for evaluation of Mild TBI. *Applied Neuropsychology, 17*, 52–59.

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*, 125–130.

Babikian, T., Boone, K. B., Lu, P., & Arnold, G. (2006). Sensitivity and specificity of various Digit Span scores in the detection of suspect effort. *The Clinical Neuropsychologist, 20*, 145–159.

Baker, R., Donders, J., & Thompson, E. (2000). Assessment of incomplete effort with the California Verbal Learning Test. *Applied Neuropsychology, 7*, 111–114.

Belsley, D. A., Kuh, E., & Welsch, R. E. (1980). *Regression diagnostics: Identifying influential data and sources of collinearity*. New York: Wiley.

Bianchini, K. J., Curtis, K. L., & Greve, K. W. (2006). Compensation and malingering in traumatic brain injury: A dose–response relationship? *The Clinical Neuropsychologist, 20*, 831–847.

Bierley, R. A., Drake, A. I., Ahmed, A., Date, E. S., Rosner, M., Warden, D., ... The Defence and Veterans Head injury Program (DVHIP) Study Group (2001). Biased responding: A case series demonstrating a relationship between somatic symptoms and impaired recognition memory performance for traumatic brain injured individuals. *Brain Injury, 15*, 697–714.

Boone, K. B., & Lu, P. H. (1999). Impact of somatoform symptomatology on credibility of cognitive performance. *The Clinical Neuropsychologist, 13*, 414–419.

Boone, K. B., Salazar, X., Lu, P., Warner-Chacon, K., & Razani, J. (2002). The Rey 15-Item Recognition Trial: A technique to enhance sensitivity of the Rey 15-Item Memorization Test. *Journal of Clinical and Experimental Neuropsychology, 24*, 561–573.

Bruusgaard, D., Smeby, L., & Claussen, B. (2010). Education and disability pension: A stronger association than previously found. *Scandinavian Journal of Public Health, 38*, 686–690.

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*, 960–971.

Carroll, L. J., Cassidy, D., Holm, L., Kraus, J., & Coronado, V. G. (2004). Methodological issues and research recommendations for mild traumatic brain injury: The WHO Collaborating Centre Task Force on Mild Traumatic Brain Injury. *Journal of Rehabilitation Medicine, 43*, 113–125.

Chaftez, M., & Prentkowski, E. (2011). A case of malingering by proxy in a social security disability psychological consultative examination. *Applied Neuropsychology, 18*, 143–149.

Dandachi-Fitzgerald, B., Pronds, R. W. H. M., Peters, M. J. V., & Merckelbach, H. (2011). Cognitive underperformance and symptom over-reporting in a mixed psychiatric sample. *The Clinical Neuropsychologist, 25*, 812–828.

Demakis, G. J., Gervais, R. O., & Rohling, M. L. (2008). The effect of failure on cognitive and psychological symptom validity tests in litigants with symptoms of post-traumatic stress disorder. *The Clinical Neuropsychologist, 22*, 879–895.

Donders, J., & Boonstra, T. (2007). Correlates of invalid neuropsychological test performance after traumatic brain injury. *Brain Injury, 21*, 319–326.

Dragano, N., & Schneider, L. (2011). Work related psychosocial factors and the risk of early disability pensioning: A contribution to assessing the need for rehabilitation. *Rehabilitation, 50*, 28–36.

Ekman, P., & O'Sullivan, M. (2006). From flawed self-assessment to blatant whoppers: The utility of voluntary and involuntary behavior in detecting deception. *Behavioral Sciences and the Law, 24*, 673–686.

Faul, M., Xu, L., Wald, M. M., & Coronado, V. G. (2010). *Traumatic brain injury in the United States: Emergency department visits, hospitalizations and deaths 2002–2006*. Atlanta, GA: Centers for Disease Control and Prevention.

Flach, P. A., Krol, B., & Groothoff, J. W. (2008). Determinants of sick-leave duration: A tool for managers? *Scandinavian Journal of Public Health, 36*, 713–719.

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 (Ed.), *Assessment of feigned cognitive impairment* (pp. 281–309). New York: The Guilford Press.

Gorissen, M., Sanz, J. C., & Schmand, B. (2005). Effort and cognition in schizophrenia patients. *Schizophrenia Research, 78*, 199–208.

Green, P. (2003). *Green's Word Memory Test for windows user's manual*. Edmonton, Alberta, Canada: Green's Publishing.
Green, P. (2009). Questioning common assumptions about depression. In J. E. Morgan & J. J. Sweet (Eds.), *Neuropsychology of malingering casebook* (pp. 132–144). New York: Psychology Press.
Green, P., & Iverson, G. L. (2001). Validation of the Computerized Assessment of Response Bias in litigating patients with head injuries. *The Clinical Neuropsychologist, 15*, 492–497.
Green, P., Iverson, G. L., & Allen, L. M. (1999). Detecting malingering in head injury litigation with the Word Memory Test. *Brain Injury, 13*, 813–819.
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*, 1045–1060.
Greiffenstein, M. F., & Baker, W. J. (2006). Miller was (mostly) right: Head injury severity inversely related to simulation. *Legal and Criminological Psychology, 11*, 131–145.
Greiffenstein, M. F., Baker, W. J., & Gola, T. (1994). Validation of malingered amnesia measures with a large clinical sample. *Psychological Assessment, 6*, 218–224.
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*, 1176–1190.
Greve, K. W., Etherton, J. L., Ord, J., Bianchini, K. J., & Curtis, K. L. (2009). Detecting malingered pain-related disability: Classification accuracy of the Test of Memory malingering. *The Clinical Neuropsychologist, 23*, 1250–1271.
Gronwall, D., & Wrightson, P. (1980). Duration of post-traumatic amnesia after mild head injury. *Journal of Clinical Neuropsychology, 2*, 51–60.
Grote, C. L., Kooker, E. K., Garron, D. C., Nyehuis, 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*, 709–719.
Hardin, J. W. (1995). Variance inflation factors and variance-decomposition proportions. *Stata Technical Bulletin (STB 24)*. Retrieved from: http://www.stata.com/bookstore/individual-stata-technical-bulletin-issues/
Harrell, F. E. (2001). *Regression modeling strategies: With applications to linear models, logistic regression, and survival analysis*. New York: Springer-Verlag.
Henry, G. K., Heilbronner, R. L., Mittenberg, W., Enders, C., Stevens, A., & Dux, M. (2011). Noncredible performance individuals with external incentives: Empirical derivation and cross-validation of the Psychosocial Distress Scale (PDS). *Applied Neuropsychology, 18*, 47–53.
Hoffman, E., McCabe, K., Shachat, K., & Smith, V. (1996). Social distance and other regarding behavior in dictator games. *American Economic Review, 86*, 653–660.
Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression* (2nd ed. ed.). New York: John Wiley & Sons, Inc.
Inman, T. H., Vickery, C. D., Berry, D. T. R., Lamb, D. G., Edwards, C. L., & Smith, G. T. (1998). Development and initial validation of a new procedure for evaluating adequacy of effort given during neuropsychological testing: The Letter Memory Test. *Psychological Assessment, 10*, 128–139.
Jones, A., & Ingram, M. V. (2011). A comparison of selected MMPI-2 and MMPI-2-RF Validity Scales in assessing effort on cognitive tests in a military sample. *The Clinical Neuropsychologist, 25*, 1207–1227.
Jones, A., Ingram, M. V., & Ben-Porath, Y. (2012). Scores on the MMPI-2 RF scales as a function of increasing levels of failure on cognitive symptom validity tests in a military sample. *The Clinical Neuropsychologist, 26*, 790–815.

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*, 904–911.

Killeen, P. R. (2009). An addictive-utility model of delay discounting. *Psychological Review, 116*, 602–619.

Krokstad, S., Johnsen, R., & Westin, S. (2002). Social determinants of disability pension: A 10-year follow-up of 62 000 people in a Norwegian county population. *International Journal of Epidemiology, 31*, 1183–1191.

Kuoppala, J., Lamminpaa, A., Vaananen-Tomppo, I., & Hinkka, K. (2011). Employee well-being and sick leave, occupational accident, and disability pension a cohort study of civil servants. *Journal of Occupational and Environmental Medicine, 53*, 633–640.

Lange, R. T., Iverson, G. L., Brooks, B. L., & Rennison, V. 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*, 961–972.

Larrabee, G. J. (2003a). Detection of malingering using atypical performance patterns on standard neuropsychological tests. *The Clinical Neuropsychologist, 17*, 410–425.

Larrabee, G. J. (2003b). Exaggerated MMPI-2 symptom report in personal injury litigants with malingered neurocognitive deficit. *Archives of Neuropsychology, 18*, 673–686.

Larrabee, G. J. (2007). Malingering, research designs, and base rates. In G. J. Larrabee (Ed.), *Assessment of malingered neuropsychological deficits* (pp. 3–13). Oxford, UK: Oxford University Press.

Larrabee, G. J., Greiffenstein, M. F., Greve, K. W., & Bianchini, K. (2007). Refining diagnostic criteria for malingering. In G. J. Larrabee (Ed.), *Assessment of malingered neuropsychological deficits* (pp. 334–371). Oxford, UK: Oxford University Press.

Mahdavi, M. -E., Mokari, N., & Amiri, Z. (2011). Educational level and pseudohypacusis in medico-lega compensation claims: A retrospective study. *Archives of Iranian Medicine, 14*, 58–60.

Mason, S., Wardrope, J., Turpin, G., & Rowlands, A. (2002). Outcomes after injury: A comparison of workplace and nonworkplace injury. *The Journal of Trauma: Injury, Infection, and Critical Care, 53*, 98–103.

Meyers, J. E., Volbrecht, M., Axelrod, B. N., & Reinsch-Boothby, L. (2011). Symptom validity tests and overall neuropsychological test performance. *Archives of Clinical Neuropsychology, 26*, 8–15.

Miller, H. (1961a). Accident neurosis: Lecture I. *British Medical Journal, 1*, 919–925.

Miller, H. (1961b). Accident neurosis: Lecture II. *British Medical Journal, 1*, 992–998.

Miller, H., & Cartlidge, N. (1972). Simulation and malingering after injuries to the brain and spinal cord. *The Lancet, 299*, 580–585.

Mooney, G., Speed, J., & Sheppard, S. (2005). Factors related to recovery after mild traumatic brain injury. *Brain Injury, 19*, 975–987.

Moore, B. A., & Donders, J. (2004). Predictors of invalid neuropsychological test performance after traumatic brain injury. *Brain Injury, 18*, 975–984.

Nelson, N. W., Sweet, J. J., Berry, D. T., Bryant, F. B., & Granacher, R. P. (2007). Response validity in forensic neuropsychology: Exploratory factor analytic evidence of distinct cognitive and psychological constructs. *Journal of the International Neuropsychological Society, 13*, 440–449.

O'Leary, A. (1985). Self-efficacy and health. *Behaviour Research and Therapy, 23*, 437–451.

Ord, J. S., Boettcher, A. C., Greve, K. W., & Bianchini, K. J. (2010). Detection of malingering in mild traumatic brain injury with the Conners' Continuous Performance Test–II. *Journal of Clinical and Experimental Neuropsychology, 32*, 380–387.

Paniak, C., Reynolds, S., Toller-Lobe, G., Melnyk, A., Nagy, J., & Schmidt, D. (2002). A longitudinal study of the relationship between financial compensation and symptoms after treated mild traumatic brain injury. *Journal of Clinical and Experimental Neuropsychology, 24*, 187–193.

Rees, L. M., Tombaugh, T. N., & Boulay, L. (2001). Depression and the Test of Memory Malingering. *Archives of Clinical Neuropsychology, 16*, 501–506.

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

Rohling, M. L., Green, P., Allen, L. M., & Iverson, G. L. (2002). Depressive symptoms and neurocognitive test scores in patients passing symptom validity tests. *Archives of Clinical Neuropsychology, 17*, 205–222.

Ross, S. R., Putnam, S. H., Millis, S. R., Adams, K. M., & Krukowski, R. A. (2006). Detecting insufficient effort using the Seashore Rhythm and Speech-Sounds Perception Tests in head injury. *The Clinical Neuropsychologist, 20*, 798–815.

Ruff, R. M., Iverson, G. L., Barth, J. T., Bush, S. S., Broshek, D. K., & the NAN Policy, Planning Committee (2009). Recommendations for diagnosing a mild traumatic brain injury: A National Academy of Neuropsychology education paper. *Archives of Clinical Neuropsychology, 24*, 3–10.

Ruocco, A. C., Swirsky-Sacchetti, T., Daines, D., Chute, D. L., Platek, S. M., Mandel, S., et al. (2005). Dissociation between neuropsychological and psychiatric forms of malingering [abstract]. *Archives of Clinical Neuropsychology, 20*, 883–883.

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 (Ed.), *Assessment of feigned cognitive impairment* (pp. 405–427). New York: The Guilford Press.

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*, 437–453.

Seligman, M. E. P., & Csikszentmihalyi, M. (2000). Positive psychology: An introduction. *American Psychologist, 55*, 5–14.

Shores, E. A., Marosszeky, J. E., Sandanam, J., & Batchelor, J. (1986). Preliminary validation of a clinical scale for measuring the duration of post-traumatic amnesia. *Medical Journal of Australia, 144*, 569–572.

Slick, D. J., Sherman, E. M. S., & Iverson, G. L. (1999). Diagnostic criteria for malingered neurocognitive dysfunction: Proposed standards for clinical practice and research. *Clinical Neuropsychologist, 13*, 545–561.

Statacorp (2011). *Stata Release 12. Statistical Software*. College Station, TX: Statacorp LP.

Stulemeijer, M., Andriessen, T. M. J. C., Brauer, J. M. P., Vos, P. E., & Van der Werf, S. (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*, 309–318.

Suhr, J. A., & Barrash, J. (2007). Performance on standard attention, memory, and psychomotor speed tasks as indicators of malingering. In G. J. Larrabee (Ed.), *Assessment of malingered neuropsychological deficits* (pp. 334–371). Oxford, UK: Oxford University Press.

Suhr, J., Tranel, D., Wefel, J., & Barrash, J. (1997). Memory performance after head injury: Contributions of malingering, litigation status, psychological factors, and medication use. *Journal of Clinical and Experimental Neuropsychology, 19*, 500–514.

Sullivan, M. J. L., Bishop, S. R., & Pivik, J. (1995). The Pain Catastrophizing Scale: Development and validation. *Psychological Assessment, 7*, 524–532.

Sumanti, M., Boone, K. B., Savodnik, I., & Gorsuch, R. (2006). Noncredible psychiatric and cognitive symptoms in a worker's compensation "stress" claim sample. *The Clinical Neuropsychologist, 20*, 754–765.

Teasdale, G. M. (1995). Head injury. *Journal of Neurology, Neurosurgery and Psychiatry, 58*, 526–539.

Teasdale, G., & Jennett, B. (1974). Assessment of coma and impaired consciousness: A practical scale. *The Lancet, 304*, 81–84.

Tombaugh, T. (1996). *Test of Memory Malingering (TOMM)*. North Tonawanda, NY: Multi-Health Systems.

Tombaugh, T. N. (1997). The Test of Memory Malingering (TOMM): Normative data from cognitively intact and cognitively impaired individuals. *Psychological Assessment, 9*, 260–268.

Torres, V., Collado, A., Arias, A., Peri, J. M., Bailles, E., Salamero, M., & Valdes, M. (2009). Pain locus of control predicts return to work among Spanish fibromyalgia patients after completion of a multidisciplinary pain program. *General Hospital Psychiatry, 31*, 137–145.

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*, 297–313.

Waddell, G. (1993). A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low-back pain and disability. *Pain, 52*, 157–168.

Wenzel, M. (2004). An analysis of norm processes in tax compliance. *Journal of Economic Psychology, 25*, 213–228.

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*, 114–125.

White, B., Driver, S., & Warren, A. (2008). Considering resilience in the rehabilitation of people with traumatic disabilities. *Rehabilitation Psychology, 53*, 9–17.

Whiteside, D. M., Dunbar-Mayer, P., & Waters, D. P. (2009). Relationship between TOMM performance and PAI validity scales in a mixed clinical sample. *The Clinical Neuropsychologist, 23*, 523–533.

Williams, D. H., Levin, H. S., & Eisenberg, H. M. (1990). Mild head injury classification. *Neurosurgery, 27*, 422–428.

APPENDIX C

Letter of Final Approval by the Executive of the Ethics Review Committee (Human Research)

MACQUARIE UNIVERSITY

Research Office
Research Hub, Building C5C East
MACQUARIE UNIVERSITY NSW 2109

Phone **+61 (0)2 9850 8612**
Fax +61 (0)2 9850 4465
Email ro@vc.mq.edu.au

Ethics
Phone **+61 (0)2 9850 6848**
Email ethics.secretariat@vc.mq.edu.au

8 July 2009

Mr James Webb
School of Behavioural Cognitive and Social Sciences
University of New England
Armidale
NSW 2351

Reference: HE26JUN2009-D00005

Dear Mr Webb,

FINAL APPROVAL

Title of project: Symptom exaggeration following traumatic brain injury: prevalence, factor structure, and associated demographic features

The above application was granted Interim Approval by the Executive of the Ethics Review Committee (Human Research) on 9 June 2009. This Interim Approval was considered by the Committee at its meeting on 26 June 2009 and was ratified.

Please note the following standard requirements of approval:

1. Approval will be for a period of twelve (12) months. At the end of this period, if the project has been completed, abandoned, discontinued or not commenced for any reason, you are required to submit a Final Report on the project. If you complete the work earlier than you had planned you must submit a Final Report as soon as the work is completed. The Final Report is available at: http://www.research.mq.edu.au/researchers/ethics/human_ethics/forms

2. However, at the end of the 12 month period if the project is still current you should instead submit an application for renewal of the approval if the project has run for less than five (5) years. This form is available at http://www.research.mq.edu.au/researchers/ethics/human_ethics/forms. If the project has run for more than five (5) years you cannot renew approval for the project. You will need to complete and submit a Final Report (see Point 1 above) and submit a new application for the project. (The five year limit on renewal of approvals allows the Committee to fully re-review research in an environment where legislation, guidelines and requirements are continually changing, for example, new child protection and privacy laws).

3. Please remember the Committee must be notified of any alteration to the project.

4. You must notify the Committee immediately in the event of any adverse effects on participants or of any unforeseen events that might affect continued ethical acceptability of the project.

5. At all times you are responsible for the ethical conduct of your research in accordance with the guidelines established by the University http://www.research.mq.edu.au/researchers/ethics/human_ethics/policy

If you will be applying for or have applied for internal or external funding for the above project it is your responsibility to provide Macquarie University's Research Grants Officer with a copy of this letter as soon as possible. The Research Grants Officer will not inform external funding agencies that you have final approval for your project and funds will not be released until the Research Grants Officer has received a copy of this final approval letter.

Yours sincerely

Ms Karolyn White
Director of Research Ethics
Chair, Ethics Review Committee (Human Research)

Cc: Dr Jennifer Batchelor, Department of Psychology, Macquarie University