Psychometric characteristics of the Functional Outcome Profile: a new measure of outcome following brain injury

by

Stacey L. Ross
B.A., McMaster University, 2004

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of

MASTER OF SCIENCE

in the Department of Psychology

Supervisory Committee

Psychometric characteristics of the Functional Outcome Profile: a new measure of outcome following brain injury

by

Stacey L. Ross
B.A., McMaster University, 2004

Supervisory Committee

Dr. Ronald W. Skelton, (Department of Psychology)
Supervisor

Dr. Fredrick M. E. Grouzet (Department of Psychology)
Departmental Member

Dr. Michael Joschko (Department of Psychology)
Departmental Member

Abstract

Supervisory Committee

Dr. Ronald W. Skelton, Psychology
Supervisor

Dr. Fredrick M. E. Grouzet, Psychology
Departmental Member

Dr. Michael Joschko, Psychology
Departmental Member

The Functional Outcome Profile (FOP, Joschko & Skelton, 2003) was developed to be a broad measure of functional outcome following brain injury, unique from existing outcome measures. The purpose of this thesis was to investigate the underlying assumptions and psychometric properties of aspects of the FOP. Data from 107 brain injury survivors and 20 matched significant others were utilized. Exploratory investigations of the items, the scoring and the subscales were undertaken first, followed by investigations of the internal consistency and test-retest reliability of the outcome scores and the survivor-proxy agreement of the frequency ratings. Finally, multiple lines of evidence for the validity of the FOP were put forth based on the content, its relation to external, demographic variables and its sensitivity to change over time. In general, the FOP was found to be a reliable measure of outcome following brain injury with the subscale scores providing the most useful information. The implications for these findings are discussed.

Table of Contents

Supervisory Committee ii
Abstract iii
Table of Contents iv
List of Tables vi
List of Figures vii
Chapter 1 1
  Recovery and Outcome Following Brain Injury 4
  Brain Injury Rehabilitation 8
  How Should Post-Acute Outcome Be Measured? 10
Outcome Measures 11
  Reliability 12
  Validity 13
  Floor and Ceiling Effects 15
  Existing Outcome Measures 15
  General Limitations of Current Outcome Measures 20
  The Functional Outcome Profile (FOP) 22
Research Plan 33
Chapter Two: Methods 34
  Participants 34
  Measures 37
  Procedure 37
  Analytic Strategy 38
Chapter Three: Results 44
  Exploratory Investigations 44
  Reliability 58
  Validity 75
Chapter Four: Discussion 87
  Exploratory Investigations 87
  Reliability 90
  Validity 92
  Comparison to Previous Psychometric Investigation of the FOP 96
  Comparison to Existing Outcome Measures 97
  A Brief Discussion of Problematic Findings 99
  Limitations 104
  Future Directions 108
  Conclusions 109
References 112
Appendix A Detailed description of existing outcome measures 120
  Glasgow Outcome Scale (GOS) and Extended Glasgow Outcome Scale (GOSE) 120
  Disability Rating Scale (DRS) 123
  Functional Independence Measure (FIM™) and Functional Assessment Measure (FAM) 125
  Community Integration Questionnaire (CIQ) 129

  Mayo-Portland Adaptability Inventory (MPAI-4) 132
  European Brain Injury Questionnaire (EBIQ) 134
  Brain Injury Community Rehabilitation Outcome Scale (BICRO-39) 137
Appendix B The Functional Outcome Profile (FOP) 140

List of Tables

Table 1. 35
Table 2. 46
Table 3. 46
Table 4. 47
Table 5. 49
Table 6. 51
Table 7. 53
Table 8. 55
Table 9. 57
Table 10. 58
Table 11. 58
Table 12. 60
Table 13. 62
Table 14. 66
Table 15. 67
Table 16. 68
Table 17. 71
Table 18. 73
Table 19. 76
Table 20. 78
Table 21. 83
Table 22. 85

List of Figures

Figure 1. Visual-analogue scales used to help survivors and significant others provide ratings......25
Figure 2. An example of the graphic profile provided by the FOP. The dark grey bars indicate the outcome score for each item, so the higher the bar, the better the functioning as perceived by the rater. 28
Figure 3. Item-Composite outcome score correlations for the Vancouver group (n = 27) and the Gorge group (n = 80). Black bars indicate that the Vancouver group had a higher correlation between the given item and the composite outcome score; white bars indicate that the Gorge group had a higher correlation between the given item and the composite outcome score. 65
Figure 4. Number of survivors reporting each frequency rating for Item #61 (Other Problems). 81
Figure 5. Number of significant others reporting each frequency rating for Item #61 (Other Problems)......81
Figure 6. Percent change in FOP outcome scores for the short and long interval groups separately, with those changes larger than the SDD indicated. 86

Chapter 1

Psychologists rely heavily on a multitude of different tests and instruments to gain information from clients that is otherwise difficult to obtain. Tests of intelligence, personality and symptomatology, for example, are utilized in both clinical and research settings. Before such tests are put to use, however, they must first be shown to psychometrically reliable and have sufficient evidence for the validity of the information they provide. In this thesis, I will investigate the psychometric properties of specific aspects of a new measure, the Functional Outcome Profile (FOP). The FOP was designed to collect information regarding the daily functioning of survivors of brain injury and their significant others for both research and clinical purposes.

Brain injury is a serious public-health problem. However, the scope of the crisis is under-recognized by the general population (Harris Interactive Inc., 2000); as such, brain injury is commonly referred to as a “silent epidemic” (Langlois, Marr, Mitchko, & Johnson, 2005). However, available statistics highlight the extent of the problem: the Canadian Institute for Health Information (2006) estimates that approximately 15,300 Canadians were admitted to hospital in the 2003-2004 fiscal year with a brain injury, and in the United States it has been estimated that someone sustains a traumatic brain injury (TBI) every 21 seconds (Binder, Corrigan, & Langlois, 2005).

The problem is exacerbated by the large number of brain injury survivors who are left with disabilities. These disabilities often result in an increased need for health services (Angus et al., 1998; Ghajar, 2000), an inability to return to work (Wehman, Targett, West, & Kregel, 2005), and many other persistent functional and psychosocial problems (Dikmen, Machamer, Powell, & Temkin, 2003). As a group, males between the ages of 15 and 24 consistently sustain more traumatic brain injuries per year than almost any other group of individuals (National Institutes of Health Consensus, 1999). Since the long-term effects of brain injury do not cause a significant decrease in life expectancy, the number of brain injury survivors grows every year resulting in an exorbitant number of individuals living with the long-term sequelae of brain injury.

Brain injuries are generally categorized as either a traumatic brain injury (TBI) or an acquired brain injury (ABI). TBI, which results from external forces, occurs when the head makes impact with a solid object at a high velocity and/or there is movement of the brain within the skull (known as acceleration/deceleration injury) (Richardson, 2000). The most frequent causes of TBI are motor vehicle accidents and falls (Baguley, Cooper, & Felmingham, 2006; Bernstein, 1999; Ghajar, 2000; Richardson, 2000). The more encompassing category of ABI includes TBI, as well as endogenous causes of damage such as stroke, aneurysm, tumor, anoxia and infection or disease (Richardson, 2000). The severity of the resulting damage to the brain varies widely between individuals and can range from a mild contusion to damage so severe that it results in death (Richardson, 2000).

The number and severity of symptoms experienced by an individual can be related to the severity of the brain damage, however symptom profiles are extremely variable (Ghajar, 2000). The symptoms experienced following brain injury can affect every aspect of the survivor’s life: deficits in physical functioning (such as paralysis or paresis, pain, and sensory deficits), cognitive functioning (problems with memory, language comprehension and expression, and attention), executive functioning (the inability to plan, make decisions, and inhibit inappropriate behaviour), affective expression (problems with mood and emotional reactivity), as well as problems functioning in the community (loss of contact with friends, isolation, inability to return to work) (Donnelly, Donnelly, & Grohman, 2005; Fork et al., 2005; Huang, Wartella, Kreutzer, Broaddus, & Lyckholm, 2001). However, it is important to emphasize that the precise symptoms that any one survivor will experience are impossible to predict based on the location and/or severity of their injury (van Baalen et al., 2003).

In addition, researchers and clinicians have recently become more aware that the survivor is not the only person who is affected by the brain injury. Brain injury and the impact that is has on the survivor also invariably affects those close to the survivor, such as his/her family. Some of the consequences can include changes in the required role of the survivor's significant other from equal partner to caregiver (McPherson, Pentland, & McNaughton, 2000; Wood, Liossi, & Wood, 2005), the experience of burden and psychological distress (such as symptoms of anxiety and depression) within the significant other (Anderson, Parmenter, & Mok, 2002; Hoofien, Gilboa, Vakil, & Donovick, 2001), and unhealthy family functioning in response to behavioural and/or cognitive problems experienced by the survivor (Anderson et al., 2002). Despite these consequences, recent research has also begun to show how important significant others and other support providers are to the psychosocial health of survivors (McCarthy et al., 2006). In fact, Tomberg, Toomela, Pulver and Tikk (2005) conclude that a high quality and quantity of social support are important for the survivor to achieve effective rehabilitation and maintain a sense of well-being. Thus, although the brain injury can induce situations that can strain those close to the survivor, the quality of social support that these people provide can be viewed as one crucial element to a brain injury survivor's recovery. For this reason, significant others and family should be involved in the recovery and rehabilitation process whenever possible.

Recovery and Outcome Following Brain Injury

The concept of recovery following brain injury is not clear-cut; it can be extremely protracted and is likely specific to the individual. The constellation of symptoms experienced by a survivor over the course of recovery is likely not fixed. One theory is that early in the post-injury period, the most distressing problems are of a physical nature such as paralysis or paresis. However, as time passes, individuals become more aware of the cognitive deficits that they are experiencing, such as the inability to remember new information (anterograde amnesia) and a lack of concentration. It is not until survivors have had a sufficient amount of time back in the community that they have the opportunity to become aware of problems in social and executive functioning, such as the inability to act appropriately in social situations and make appropriate decisions. However, this theory remains to be empirically investigated.

Different definitions of recovery exist within the brain injury literature, serving to further complicate the discussion. For instance, recovery can imply that the post-injury behaviour is identical to the pre-injury behaviour; the behaviour is performed in exactly the same way, served by exactly the same neural structures (Stein, 2000). Yet when others speak of recovery after brain injury their definition is one of *partial restitution*, or a gradual improvement in the ability to perform particular behaviours, which may or may not return to the same level of accuracy as before the injury (Kolb, 1995). Often partial restitution occurs over the first few months following the injury and likely reflects the remission of some secondary effect of the brain injury, such as swelling (Kolb, 1995). However, because the brain structures that are damaged show limited reorganization, they will not function exactly as they had before the injury. Therefore, a more useful characterization of recovery includes the phenomenon of *compensation* (Stein, 2000). This form of recovery is more accurately an adaptation to the loss of function, and may result from a change in the strategy used to complete the behaviour, or the substitution of a new behaviour for one that can no longer be performed reliably (Kolb, 1995). While compensation does not represent a return of the behaviour that was lost (Kolb, 1995), it is still a useful aspect of functional recovery because the individual will, by definition, be functioning better because of the compensatory strategy. Likely, this is what matters most to survivors as well – better functioning regardless of the means allowing this to occur. Thus, as long as an individual reports improved completion of a task, regardless of the means used to reach the goal, I will regard this as recovery.

The above description shows that the concept of recovery implies an ongoing process; a series of actions (e.g., spontaneous changes in neural connections or a compensatory strategy learned during the course of rehabilitation) that may eventually result in improved functioning. Outcome, however, is more static; it is a picture of one's functioning at a specific point in time. Thus, there are two aspects to the concept of outcome: one is the "when" of outcome, or when it is investigated; the other is the "what" of outcome, or what behaviours are of particular concern. These two aspects are likely to be strongly interrelated: *when* outcome is investigated will determine *what* behaviours are deemed important to look at. Both when and what type of outcome is investigated will further be determined by *who* is doing the investigation.

When referring to brain injury, the term 'outcome' may be more difficult to define than is typical with other medical conditions. For example, if an individual suffers from acute appendicitis and must have his appendix removed, outcome will be investigated by the surgeon and the in-patient care team. As such, the patient's outcome will likely be investigated over a very discrete window of time: following the surgery, prior to discharge from the hospital, and during a follow-up appointment a week or so after the surgery. At each of these points in time the outcome “behaviours” deemed important may also be different. Following the surgery, outcome may be based on behaviours such as stable heartbeat and respiration while at the follow-up appointment behaviours such as mobility and appetite may be more important. However, the patient’s final outcome from the surgery will most likely be clearly defined and determined at the follow-up appointment. Following a brain injury there are many individuals who will have interests in, and different approaches to, the investigation of outcome. For example, the neurologist will investigate the survivor’s outcome when he/she is initially discharged from the hospital. The primary focus of the neurologist will be the survivor’s neurological functioning including aspects such as muscle tone, reflexes, arousal, gait, and posture; if the survivor is doing well in these areas, the neurologist is likely to report that the survivor has shown a good outcome. Following rehabilitation, however, the rehabilitation team will also investigate the survivor’s outcome with a specific focus on the areas that were targeted during the program. For example, upon the completion of a cognitive rehabilitation program the survivor’s memory, attention, awareness or cognitive flexibility may be the focus of outcome investigation. Upon return to the community, governmental agencies (such as those responsible for providing disability benefits or worker’s compensation) will be interested in the survivor’s outcome, however the definition of outcome used by these entities will be very broad such as ability to return to work or live independently. Finally, the survivor will also have opinions about his/her own outcome. This interpretation of outcome may not be based on formal investigation or measurement but instead on reflections of previous functioning, expectations of recovery or comparison to others. What is important is that, unlike many other medical conditions, it is difficult to determine when a brain injury survivor has reached his/her “final” outcome, and opinions on the point at which this occurs will likely differ depending on who you ask – a neurologist, a occupational therapist, or the survivor him- or herself.

One thing that is clear is that outcome is generally viewed as how the individual is functioning. The World Health Organization’s (2001) *International Classification of Functioning, Disability and Health* reports that functioning is a combination of three specific aspects: 1) the *functional and structural integrity of the body*, or the physiological functions of body systems and the anatomical parts of the body; 2) *activities*, or the execution of a task or action by an individual; and 3) *participation*, or an individual’s involvement in a life situation. Disability, the opposite of functioning, is therefore the combination of impairment, defined as problems in body function or structure (e.g., brain injury), and the resulting activity limitations and participation restrictions. Consequently, the investigation and measurement of brain injury outcome is interested in the degree of functioning and/or the lack of disability present following the injury at specific points in time.

Outcome (i.e., functioning) has been found to be consistently related to a number of individual- and injury-characteristics. Outcome is commonly found to be influenced by age at the time of injury, with individuals who are older when they sustain an injury having poorer long-term outcome (Rothweiler, Temkin, & Dikman, 1998). Outcome may also be affected by gender, as it has been shown that women report better community integration and more independence in completing daily tasks than men (Schmidt, Garvin, Heinemann, & Kelly, 1995). Also, with all other variables being equal, outcome will be influenced by the length of time since the injury due to continued recovery that will occur over time (see, for example, Dikmen, et al., 2003). Outcome may also be influenced by the type of injury sustained; for example, TBI and cerebrovascular disease have been shown to display different sequelae and outcome profiles (Hellawell & Pentland, 2001). Finally, outcome is generally found to be negatively related to injury severity such that individuals with more severe injuries often have poorer outcome (see, for example, Cattelani, Tanzi, Lombardo, & Mazzucchi, 2002; Teasdale & Engberg, 2005). However, the relationship between outcome and injury severity is not necessarily linear. Mailhan, Azouvi and Dazord (2005) found that individuals with moderate disability following brain injury reported the lowest levels of life satisfaction, while individuals with severe or mild disability did not differ in their post-injury life satisfaction ratings. This relationship could reflect the possibility that survivors with moderate disabilities have a greater sense of their losses compared to those with mild disabilities who have mild losses, and those with severe disabilities who may have a greater sense of their recovery, a greater appreciation for simply surviving or less awareness of their deficits. However, it is also possible that differences such as these may reflect discrepancies in the definition of outcome (e.g. outcome defined as a score on the Glasgow Outcome Scale as opposed to outcome defined as life satisfaction). What has been conclusively shown is that rehabilitation improves outcome following brain injury (e.g., Cicerone, Mott, Azulay, & Friel, 2004; Sarajuuri et al., 2005; Svendsen & Teasdale, 2006).

Brain Injury Rehabilitation

There are generally three main phases in the rehabilitation process (Pietrapiana, Bronzino, Perino, & Rago, 1997). The first phase, or peri-coma phase, begins immediately following the injury and lasts until the brain injury survivor is medically stable (Pietrapiana, et al., 1997). While much of this phase involves secondary and/or tertiary healthcare intervention (e.g., intensive care and neurosurgery) a treatment plan may also be devised at this point.

Rehabilitation involving postural control, range of motion exercises, and controlled sensory stimulation is also included at this stage (Pietrapiana et al., 1997).

Once the patient is medically stable and has regained consciousness, the second phase, or acute phase, of rehabilitation can begin (Pietrapiana et al., 1997). This phase often requires the transfer of the survivor to a special rehabilitation unit as specific therapies require specific environments that are not readily available in other, more general, units of the hospital (Pietrapiana et al., 1997). This phase is commonly known as in-patient rehabilitation.

The third, or post-acute, phase of rehabilitation is the phase of social integration or community rehabilitation (Pietrapiana et al., 1997). While the temporal criteria for when to begin such treatment is dependent on the individual, it is often the case that this phase begins approximately 6 months after the injury (Pietrapiana et al., 1997). A large portion of this treatment may involve participation in a day, out-patient program involving organized activities at least 4 days a week (Pietrapiana et al., 1997). The aims of community rehabilitation are to maintain the functional improvements made during the previous rehabilitation interventions, as well as to enhance the survivor's adaptive functioning and work towards re-integration into society (Mateer, 1997; Pietrapiana et al., 1997).

The focus of rehabilitation should be on "real-life, functional problems" experienced by the survivor (B. A. Wilson, 2002). As such, Ben-Yishay (2000) briefly outlines some of the general aims of rehabilitation as: 1) preventing the occurrence of "catastrophic responses" (e.g., severe anxiety in response to the completion of required tasks); 2) helping the survivor overcome the tendency to avoid engaging in challenging learning situations; 3) assisting the individual in attaining, through learned compensation, modified capabilities that remain within that individual's potential; 4) moving the individual towards becoming productive again; and

5) helping the individual begin to feel “healthy” again. Thus, it is evident that the current goal of rehabilitation has become improving the survivor’s well-being and ability to function in society.

As mentioned, rehabilitation – specifically post-acute, community integrated, holistic rehabilitation – has been found to be efficacious in improving outcome following brain injury (Cicerone et al., 2004; Prigatano et al., 1984; Rice-Oxley & Turner-Stokes, 1999; Svendsen & Teasdale, 2006). Interestingly, the effectiveness of rehabilitation is not limited to the early post-acute stage following brain injury; rehabilitation has been shown to improve outcome in individuals more than 2-years post-injury (Coetzer & Rushe, 2005). Furthermore, Diller and Ben-Yishay (2003) have determined that, based on existing studies, comprehensive holistic rehabilitation meets the standard to be regarded as a recommended practice guideline, indicating that it should be regarded a first line treatment for survivors of brain injury. However, the specific types of rehabilitation that should be included in a specific individual’s program (e.g., cognitive remediation, psychotherapy, physical or speech therapy) depend on the needs of the survivor, which many experts agree should be determined with the input from the survivor him or herself (Prigatano, 1997; B. A. Wilson, 2002).

How Should Post-Acute Outcome Be Measured?

Because of the variability in long-term sequelae following brain injury, it is very difficult to know what functional deficits a survivor will experience. Therefore the best measure of post-acute outcome needs to be comprehensive in its inclusion of *possible* functional deficits, and obtain the subjective experience from the survivor. Such a measure would benefit the acquisition of knowledge in three major ways: first, on a case-by-case basis, it would help determine what functional areas an individual is experiencing difficulty with in their everyday life and therefore assist in determining where to focus rehabilitation efforts, especially post-acute rehabilitation; second, it would be ideal to investigate the efficacy of rehabilitation programs or to compare different approaches to rehabilitation; and third, on a larger scale, it could be used to investigate the general patterns of functional recovery seen following brain injury.

Outcome Measures

Measures used specifically to investigate outcome following brain injury are different from traditional neuropsychological tests and/or batteries. The major difference is that neuropsychological tests are used to determine specific deficits (i.e., diagnosis) (Strauss, Sherman, & Spreen, 2006), while outcome measures should be used to describe levels of functioning (Hall, Bushnik, Lakisic-Kazazic, Wright, & Cantagallo, 2001) ideally in the real world. As such, standardized neuropsychological tests may not be the best way to determine outcome following brain injury. It is believed that the investigation of outcome should include assessments of functioning to compliment the information that is obtained from standardized assessments (B. A. Wilson, 2002).

Many outcome measures are currently available. However, each has shortcomings. Some of these shortcomings include: a lack of scope; overlooking the subjective experiences of the survivors and/or their significant others; and insufficient psychometric analyses. This last point is extremely important as a measure’s scores must be reliable, sufficiently sensitive in order to avoid floor and ceiling effects and able to detect change when change occurs (e.g., during the course of recovery). Overall, evidence must be collected to show that the proposed interpretations of a measure’s scores are valid. The next section will provide an overview of some of the common approaches taken to illustrate the validity of a measure’s score interpretation.

Reliability

Reliability refers to the accuracy of mental measurement. Accuracy is determined by the consistency of scores across many administrations of the same test (or multiple tests known to measure the same construct administered once each) (Osterlind, 2006). A reliable measure will provide a score that is close to an individual’s “true score” with little influence by different types of error. There are many different ways to test reliability. The different tests are meant to investigate the influence of the different types of error that can be encountered.

Internal consistency reflects item homogeneity, or the extent to which items measure the same domain or construct (Henderson, 2001). Internal reliability is a measure of the intercorrelation of test items which is commonly reported using the statistic *alpha* (more specifically, Cronbach’s alpha). As Strauss, Sherman and Spreen (2006) explain, “alpha is based on the average intercorrelation between test items, and is used for tests with items that have more than two possible responses”. An alpha value greater than 0.80 is generally desirable (Henderson, 2001).

Test-retest reliability provides an estimate of the stability of scores across time (Henderson, 2001) by determining the correlation between scores for a single measure given to the same individual at two separate times. Simple correlations are often used to measure test-retest reliability and commonly accepted standards state that an r between 0.10 and 0.30 indicates a weak relationship between scores, 0.30 to 0.50 a moderate relationship and 0.50 and above a strong relationship (Cohen, 1988). Tests that measure dynamic abilities will, by definition, produce lower test-retest reliabilities than tests that measure more trait-based and stable abilities (Strauss et al., 2006). Test-retest reliability is an important property of outcome measures, most of which are used to investigate improvement or recovery, because it allows one to be sure that any change found is not merely an artefact of an unreliable measure.

Inter-rater reliability provides an estimate of the consistency of the administration and scoring of the test across examiners. Another form of inter-rater reliability commonly utilized in the investigation of brain injury outcome measures is *Survivor-proxy reliability* where the amount of agreement between the survivor and an individual who knows the survivor well is measured. Kappa is a commonly used statistic to investigate inter-rater reliability due to the fact that its calculation takes into account chance agreement as well as the amount of disagreement between the two observers when determining the percentage of overall agreement (van Baalen et al., 2006). However, Kappa is only appropriate when the items are categorical (Garson, 2008). Therefore, simple correlations (i.e., Pearsons or Spearmans, depending on the scale of measurement of the data) can also be used to investigate inter-rater reliability. This approach may be less conservative, however, as these statistics do not take chance agreement into account.

Validity

Validity refers to the degree to which collected evidence, in combination with existing theory, support the interpretations of a test's scores as outlined in the proposed uses of a test (American Educational Research Association [AERA], American Psychological Association [APA] & National Council on Measurement in Education [NCME], 1999). In the past, a measure was said to possess different *types* of validity based on the results of different investigations. However, validity is now regarded as a unitary concept; there a multiple lines of evidence that can be collected to illuminate different *aspects* of validity (AERA, APA, & NCME, 1999).

For example*, validity evidence based on relations to other variables* (traditionally known as *concurrent* or *convergent validity*) is one such aspect of validity. This line of evidence is based on the relationship of a test's scores to other variables that are related to the concept the test purports to measure (AERA, APA & NCME, 1999; Reynolds, Livingston, & Willson, 2006). These variables can include other tests shown to measure the concept in question, or external variables empirically shown to relate to the concept.

Another line of validity evidence, *validity evidence based on test content,* (previously known as *content validity*) is obtained through analysis of the relationship between the test's content and the construct the test is intended to measure (AERA, APA & NCME, 1999; Reynolds et al., 2006). This evidence can come from many sources including logical or empirical analyses of the adequacy of the test's content compared to the construct, expert judgements on the relationships between components of the test and the construct, or the use of rules or algorithms to determine items that systematically differ on various components of the concept (AERA, APA & NCME, 1999).

For measures that are intended for program evaluation or investigating recovery and outcome over time, *sensitivity to change* is of utmost importance. It is the degree to which a measure is likely to reflect changes that occur following participation in some type of intervention (Vermeersch, Lambert, & Burlingame, 2000), such as therapy or neurorehabilitation. Vermeersch and colleagues provide 3 guidelines for the investigation and interpretation of sensitivity to change: first, the change in scores must occur in the theoretically proposed direction; second, participants receiving an intervention should change significantly more than participants under control conditions; and third, the change should not be attributable to measurement error, non-intervention factors, practice effects, and the like.

While this is not an exhaustive list of the aspects of validity that can be investigated, it helps illustrate that validity is mulitfacted; a test cannot “be valid”. Thus, multiple lines of evidence must be obtained in order to establish the validity of a test’s score interpretation.

Floor and Ceiling Effects

Floor and Ceiling effects indicate limitations in the range of item difficulty at either the upper end (ceiling) or lower end (floor) of the scale. A test is said to have a *low ceiling* or show *ceiling effects* when a high number of examinees obtain scores at or near the highest point possible, indicating that the test lacks sufficient range of more “difficult” items (Strauss et al., 2006). Conversely, measures are said to show a *floor effect* when a large number of individuals score at or near the lowest point possible, which indicates that such measures are lacking easier items. Therefore, when used with a population such as brain injury survivors, problems with a measure’s range will make it only applicable to individuals with a limited range of disability. Given the wide range of long-term deficits possible following brain injury, measures of outcome should be sensitive at all possible functional levels.

Existing Outcome Measures

There are many outcome measures commonly utilized at different points in recovery after brain injury. The psychometric characteristics of these measures are important, as they inform users of a measure’s appropriate applications and weaknesses (for more detail, see Appendix A).

Glasgow Outcome Scale (GOS) and Glasgow Outcome Scale Extended (GOSE). The GOS (Jennett & Bond, 1975) was designed to be a brief scale to define gross outcome following brain injury, especially in the peri-coma and acute phases. Ratings are given by a clinician based on observations of, and interviews with, the brain injury survivor. Possible scores range from 0 (dead) to 5 (good recovery). The GOSE (J. T. L. Wilson, Pettigrew, & Teasdale, 1998) was developed to increase the sensitivity of the measure by increasing the number of categories: moderate disability (GOS score of 4) and good recovery (GOS score of 5) were divided into upper and lower levels of outcome. Therefore, possible scores on the GOSE range from 0 (dead) to 8 (upper good recovery).

The GOS and GOSE have been subjected to a range of psychometric analyses with mostly satisfactory outcomes. However, the GOS has been shown to suffer from ceiling effects and be insensitive to recovery after the first 6 months post-injury (Richardson, 2000). Therefore, while the GOS and GOSE are appropriate measures to use in some situations – at hospital discharge or evaluation of individuals with more severe injuries – it is not an appropriate measure for use in community living survivors or those with more mild injuries.

Disability Rating Scale (DRS). The DRS (Rappaport, Hall, Hopkins, Belleza, & Cope, 1982) was designed to measure and describe a brain injury survivor's level of disability based on his/her ability to complete activities of daily living independently. The scale consists of 8 items grouped into 4 categories: 1) arousal, awareness, responsivity; 2) cognitive ability for self-care activities; 3) dependence on others; and 4) psychosocial adaptability.

Psychometric analyses done to investigate the DRS are inconsistent; while some studies have found good reliability across administrations or raters, others have found very poor reliability (see, for example, Rappaport et al., 1982; van Baalen et al., 2006). However, the DRS has been found to be more sensitive to changes occurring over the first post-injury year than the GOS (Rappaport et al., 1982). Yet, its usefulness is still limited to measuring gross outcome in the acute rehabilitation phase after brain injury.

Functional Independence Measure (FIM™) and Functional Assessment Measure (FAM). The FIM™ (Keith, Granger, Hamilton, & Sherwin, 1987) was developed to be utilized as a standardized measure of functioning for individuals during inpatient care (Gurka et al., 1999), not brain injury survivors specifically. This is evident in the areas that the FIM™ includes: motor and self-care skills involved in activities of daily living are the focus. The FAM was developed to increase the scope of the measure and include areas more relevant to brain injury that are under-emphasized in the FIM™ (Hall et al., 2001). In the context of brain injury research and rehabilitation, the FIM™ and FAM are generally administered together as one measure known as the FIM+FAM. The FIM+FAM contains 30 items scored on a 7-point scale for a possible range of scores from 30 to 210.

The FIM+FAM has been shown to be a very reliable measure of gross functioning. Also, the FIM+FAM is one of the only measures to be analyzed using modern psychometric techniques. However, Rasch analysis showed that both the FIM™ and FAM have numerous items that do not fit with the model implying that they do not measure the proposed underlying trait well (Hawley, Taylor, Hellawell, & Pentland, 1999; Linn et al., 1999). Also, it was found that the FAM items do not extend the difficulty of the measure beyond that of the FIM™ items (Linn et al., 1999). Therefore, the FIM+FAM also suffers from ceiling effects making it inapplicable for higher functioning survivors, especially those who have returned to the community.

Community Integration Questionnaire (CIQ). The CIQ (Willer, Rosenthal, Kreutzer, Gordon, & Rempel, 1993) was developed specifically to measure community integration following brain injury. The measure consists of 15 items grouped into 3 sub-scales: home competency, social integration, and productive activity. The CIQ has two versions; one completed by the brain injury survivor and one completed by a significant other, friend or family member. In addition, the CIQ can be administered face-to-face by a trained individual or via telephone (Sander et al., 1999).

The CIQ has been found to have good internal (Sander et al., 1999), test-retest (Cusick, Gerhart, & Mellick, 2000) and survivor-proxy reliability (Cusick et al., 2000; Sander et al., 1997). It has also been found to *not* be subject to floor or ceiling effects (Hall et al., 2001), indicating that it has a good breadth of item difficulty. However, because the CIQ focuses exclusively on community integration, there are many areas of functioning that are not included.

Mayo-Portland Adaptability Inventory, Fourth Edition (MPAI-4). The MPAI (Malec & Thompson, 1994) was developed to improve on the Portland Adaptability Inventory, a measure developed by Lezak in 1987 to document and measure behavioural and social problems commonly seen following brain injury (Malec et al., 2003). The most recent version of the MPAI, the MPAI-4, includes 35 items grouped into 3 sub-scales: abilities, adjustment, and participation. All items are rated on a 4-point scale, where 0 represents no impairment and 3 indicates complete or near complete loss of function in the area. The MPAI-4 has survivor, significant other and clinician versions available.

Limited psychometric analysis has been done with the MPAI-4, but what has been done has found the measure to have good internal reliability (Malec et al., 2003) and inter-rater reliability (Malec, 2004). A Rasch analysis was done using an earlier version of the MPAI with satisfactory results (Malec, 2004; Malec, Moessner, Kragness, & Lezak, 2000).

The MPAI has also been found to be sensitive to different levels of impairment (Malec, 2004). Due to the increased scope of the MPAI-4, it is a relatively good measure of functioning for use with community-living individuals. However, the measure focuses exclusively on the frequency of the problem and does not inquire about the impact of the problem on the lives of the survivor or significant other.

European Brain Injury Questionnaire (EBIQ). The EBIQ (Teasdale et al., 1997) is one of the most recently developed measures. It was designed specifically for use with brain injury survivors and their significant others to enquire about a wide range of difficulties common following brain injury. Both the survivor and significant other versions include 63 items that address problems experienced by the survivor and 3 items that focus on the impact of the injury on the significant other. The 63 survivor-focused items are grouped into 9 sub-scales: somatic, cognitive, motivation, impulsivity, depression, isolation, physical, communication, and core symptoms. All items are rated on a 3-point scale of gross symptom frequency where 1 represents “not at all” and 3 represents “a lot” (Teasdale et al., 1997).

While limited psychometric analysis has been done on the EBIQ, what has been done has shown the measure to have good internal reliability (Teasdale et al., 1997) and good survivor-proxy agreement (Deloche, Dellatolas, & Christensen, 2000). The EBIQ has also been shown to be sensitive to the effects of brain injury (Teasdale et al., 1997) and change over time (Svendsen, Teasdale, & Pinner, 2004). The most recent study of the EBIQ has shown it to have good test-retest reliability for both the survivor and the significant other version over a 4 week interval (Sopena, Dewar, Nannery, Teasdale, & Wilson, 2007). However, the psychometric investigation is incomplete making it difficult to determine if the

EBIQ is valid as a measure of outcome. Still, it is one of the most commonly used measures of brain injury outcome in Europe (Sopena et al., 2007).

Brain Injury Community Rehabilitation Outcome Scale (BICRO-39). The BICRO-39 (Powell, Beckers, & Greenwood, 1998) was developed to assess the progress and outcome of community re-integration following brain injury. There are 3 different forms of the BICRO-39: the patient pre-injury, which requires the survivor to retrospectively rate his or her pre-injury functioning; the patient post-injury, which asks the survivor to rate his or her current functioning; and the "carer" post-injury, which requires the survivor's significant other to rate the survivor's current functioning. All forms of the BICRO-39 include 39 items grouped into 7 sub-scales: personal care, psychological, socializing, self-organization, mobility, family contact, and productive employment.

Analyses have found some of the sub-scales to have good internal reliability, while others do not (Powell et al., 1998). However the measure has been shown to have good test-retest reliability, survivor-proxy agreement, and concurrent validity (Powell et al., 1998). While one of the more promising measures available to investigate outcome following brain injury, the BICRO-39 still focuses almost exclusively on the frequency of the problem and is limited in its comprehensiveness.

General Limitations of Current Outcome Measures

While the above is not an exhaustive list of the existing outcome measures available, it includes the original measures designed to investigate outcome after brain injury as well as some of the more recent measures designed to be improvements on these earlier measures. However, it can be seen that all of these measures have limitations that should be recognized in order to further improve the investigation of brain injury outcome.

A common problem is a limited range of difficulty within the measure. Issues such as ceiling effects (the GOSE and GOSE, and the FIM+FAM) and insensitivity to long-term recovery (the GOS and GOSE, the FIM+FAM and the CIQ) make measuring outcome in community living individuals less precise than it could, and should be. For example, while the GOS and GOSE are still commonly used in the acute stage of recovery (such as hospital discharge), research has shown that very few brain injury survivors exhibit a change of category on the GOS more than six months following their accident (see Richardson, 2000 for a review). These results imply that either recovery is complete by 6 months post-injury – which, as previously mentioned, the literature has consistently shown to not be the case – or that the GOS is insensitive to the changes that do occur after the first 6 months of recovery.

Additionally, the problem of limited comprehensiveness is pervasive in these measures. Many measures include only gross measures of functioning and exclude “higher order” functional abilities more relevant to community living survivors. For example, the DRS includes only very limited tasks as a measure of outcome including abilities such as eye opening and feeding ability (cognitive understanding only) (see Wright, 2000 for a brief description and link to the DRS form). Because of the diversity in number and severity of sequelae that can be experienced after brain injury, measures with limited comprehensiveness may overlook many deficits experienced by a survivor.

Another problem with many of the measures discussed above is that they fail to inquire about the impact of the brain injury on the survivor and his/her significant others. Many of the earlier developed measures are exclusively expert-rated (e.g. the GOS and GOSE, the DRS and the FIM+FAM) completely overlooking the input of survivors. Of those that are completed by the survivor, few have equivalent survivor and significant other forms thus overlooking the input of significant others. Only the CIQ, MPAI-4, EBIQ and BICRO-39 have significant other versions. Even fewer inquire about how the injury has affected the significant other. In fact, the EBIQ is the only existing measure that includes any items that investigate the impact of the brain injury on the significant other and it employs a mere 3 questions to address this issue. Thus, it is clear that all currently available outcome measures overlook at least some aspects of the impact of brain injury.

A final limitation is that none of these measures are able to provide information beyond the incidence and/or frequency of the problem. In other words, these measures ask only *if* a problem occurs, and if it does *how often* it occurs with more frequent problems automatically being indicative or poorer outcome. As such, a high frequency problem that is relatively unimportant to the survivor will be given more weight than a lower frequency problem that is of great importance to the survivor. For example, a survivor may have constant problems with his/her divided attention but they are not of great importance to him/her; he/she may be much more concerned with his/her infrequent but problematic temper.

Therefore, in an attempt to address these issues and provide clinicians and researchers with a measure that can be used to investigate outcome in community-living brain injury survivors, the Functional Outcome Profile was developed.

The Functional Outcome Profile (FOP)

Development. The development of the FOP (Joschko & Skelton, 2003) used a theoretically and clinically driven procedure with reiterative professional consultations to ensure that the content of the FOP was appropriate for the proposed population and uses. Initially, the developers investigated existing outcome measures to determine what was available. The developers then probed three main sources to determine the necessary breadth:

the WHO's ICIDH-2, existing ABI literature, and the expertise of clinicians and rehabilitation professionals. Based on the information acquired from these sources, general constructs and specific items were devised. These items were then brought to clinicians and rehabilitation professionals for their input, and revised. This process was repeated a number of times until the rehabilitation professionals were satisfied with the content and wording of the items.

Initially, the FOP consisted of 2 sections: 1) items from the Scales of Independent Behaviour, Revised (SIB-R) and 2) 48 newly generated items. These items were tested in a preliminary experiment investigating the item characteristics and reliability of the measure based on data collected from 12 survivor-significant other pairs each interviewed 3 times at approximately 6-month intervals (Skelton et al., 2005). At the end of each interview, participants were asked if they found any questions to be difficult to understand or whether there were any areas of their life affected by brain injury that they were not asked about. Based on the participants' reaction to the questions in this study (specifically, that the SIB-R portion was "tedious"), the wording of the items was revised and 15 new items were developed to cover the areas included in the SIB-R. The items were grouped into subscales based on their functional area, not statistical similarity, producing thematic clusters of items. This was done to improve interpretability and usefulness.

In the study conducted by Price (2007) using the most recent version of the FOP, the interview was administered by 4 occupational therapists, 3 physical therapists and a neuropsychologist. These professionals were all asked to report if they found any of the items to be superfluous, ambiguous or if any important areas were missing. None of these professionals reported any problems substantive enough to warrant further changes to the interview.

Description. The FOP (Joschko & Skelton, 2003, see Appendix B) was designed to be an inventory of functional outcome following brain injury. The FOP consists of 63 items assessing a wide range of functional areas with an emphasis on quality of life. Each item collects information about a different construct area and includes multiple questions. The items are grouped into 8 theoretically-derived scales: Physical Functioning/Activities of Daily Living, Health, Cognitive Functioning, Executive Functioning, Emotional/Behavioural Functioning, Social Functioning, Activities, and Overall Well-Being.

The FOP is a self-report instrument with both a survivor version (FOP) and a significant other version (FOP SO). A clinician or trained individual reads the questions for each item and the survivor or significant other provides ratings from 0 to 10.

Survivors are asked to make a number of different ratings with respect to their current functioning including: Impact ratings (How much does your problem with X affect your day-to-day functioning?); Satisfaction ratings (How satisfied are you with X?); Importance ratings (How important is X to you?); Time estimation ratings (How much time do you spend engaged in X?); and Frequency ratings (How often does your problem with X cause you difficulty in your day-to-day life?) (see Figure 1).

Significant others are asked to make slightly different ratings; while significant others rate the frequency and time estimation questions with respect to the survivor, they rate the impact, satisfaction and importance questions with respect to themselves. In this way, the FOP SO examines the influence of the survivor's brain injury on the significant other's day-to-day functioning and allows for investigation of the impact that the brain injury has on the survivor's social network (Price, 2007).

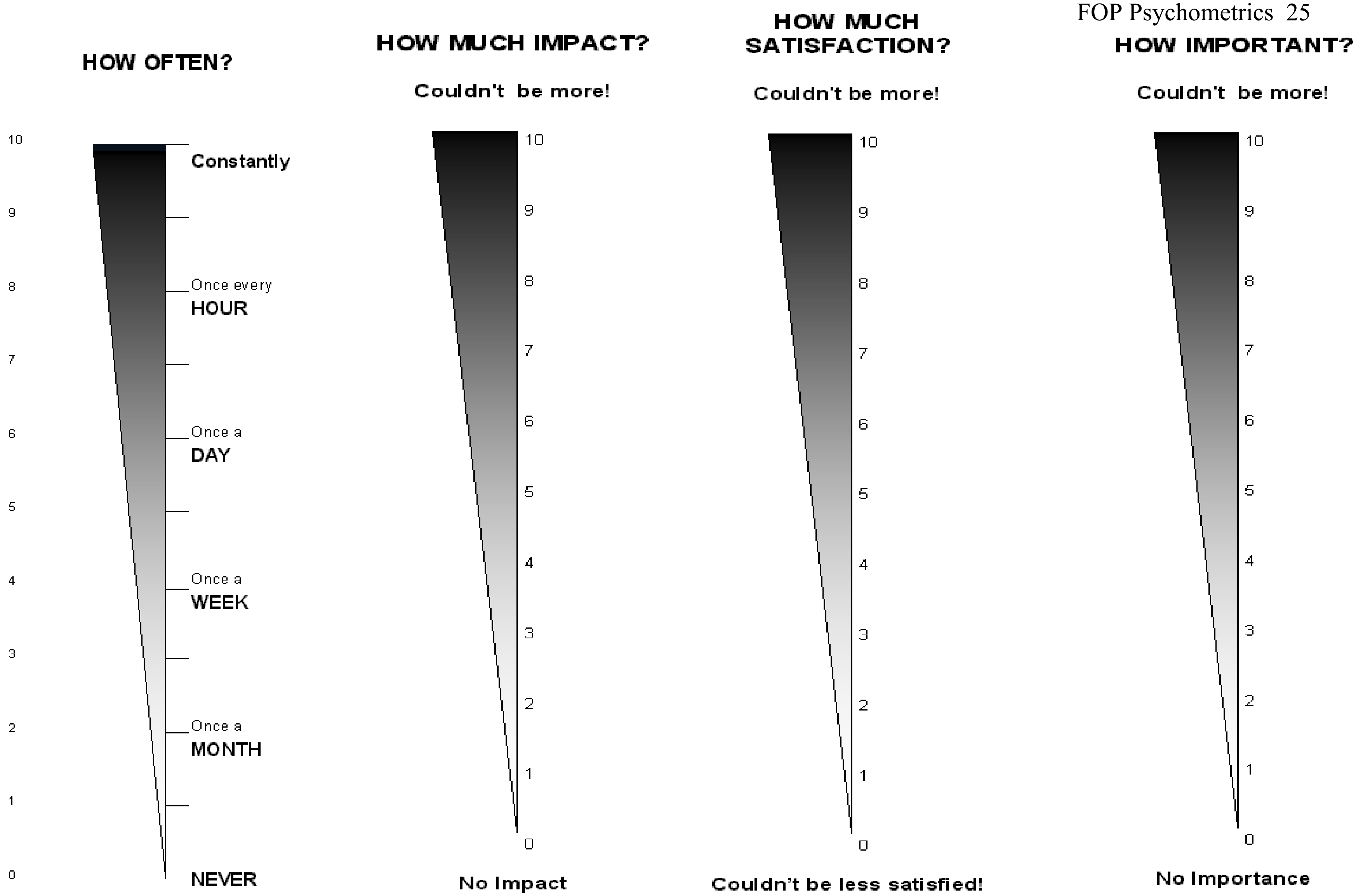


Figure 1. Visual-analogue scales used to help survivors and significant others provide ratings.

Responses to the impact, satisfaction, importance and time estimation questions are converted from ratings of 0 to 10 into a scale from 0 to 1 to ease interpretation and the impact and frequency ratings are inverted so that higher scores indicate better functioning. An outcome score is reported for each item based on the average of the ratings provided for the questions that make up the item, and calculated for each of the 8 subscales and the whole measure (the composite outcome score) based on the average of the item outcome scores. The outcome scores are based on the impact, satisfaction, importance and time estimation ratings. A secondary variable generated by the FOP is the frequency score(s) which provides information as to how often a problem is interfering with the survivor's daily functioning from the point of view of the survivor or his/her significant other.

The FOP is a unique measure of brain injury outcome for four reasons: first, the FOP is more appropriately regarded as an inventory than as a measure. One of the goals during the development of the FOP was to make it as comprehensive as possible ensuring that almost every problem that could be experienced as a result of brain injury was included. Although some items may be less frequently endorsed, their inclusion is important in order to assess the impact of the problem for those that do endorse them. Second, the FOP collects considerably more information than existing outcome measures by obtaining information about problem frequency, its impact on functioning, as well as the rater's satisfaction with the domain in question and his/her subjective rating of the importance of the domain in question. Most existing outcome measures obtain information about problem frequency only. Third, the FOP focuses on the survivor's functioning and not the survivor's deficits like most other outcome measures. This perspective shift is evident in the interaction between the inclusion of impact questions and diminished focus on problem frequency; the FOP includes problem frequency almost exclusively as a means of providing an equal and objective rating between survivors and significant others. The focus on functioning is in accordance with the definition of recovery discussed earlier. It is very difficult, and in fact sometimes impossible, to overcome a deficit (e.g., a memory deficit may show little to no true recovery). However, compensatory mechanisms can help improve an individual's functioning (e.g., using a personal organizer to compensate for a memory deficit). Thus, the FOP items inquire purely about functioning and not whether that functioning involves any compensation for a deficit. Finally, the FOP provides a profile of functioning. In other words, the FOP can be interpreted at the item level to provide information about specific domains that are problematic (e.g., Fatigue), or at the subscale level to provide more general information about construct areas that are problematic (e.g., Social Functioning). The profile (see Figure 2 for an example) provides a depiction of an individual's functioning in each domain in relation to other domains. This profile will be unique to the individual, like a fingerprint. In addition, looking at how the profile changes over time can illustrate domains where an individual has felt functional improvement. Clients may benefit from being shown this information.

Currently, the FOP has two proposed clinical applications: the first is to augment the development of individual rehabilitation programs by providing clinicians with the subjective experiences of the survivor and significant other; the second is to investigate the effectiveness of rehabilitation programs by determining which areas the survivor and significant other feel have improved. Possible experimental applications of the FOP include program evaluation on a larger scale, the longitudinal investigation of recovery after brain injury focusing on profile changes over time and the long term effects of brain injury on significant others including if and how the impact on their lives changes over time.

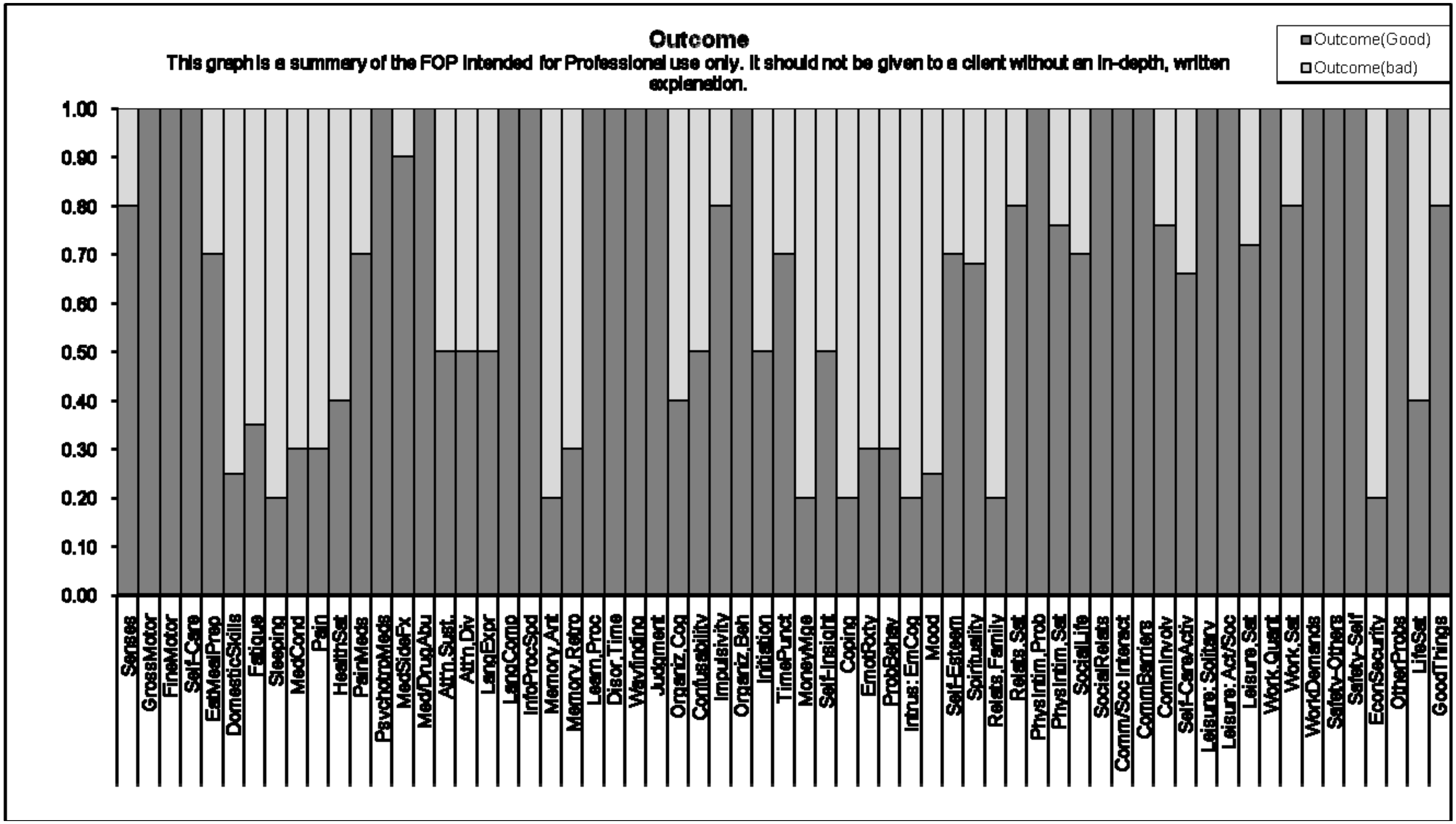


Figure 2. An example of the graphic profile provided by the FOP. The dark grey bars indicate the outcome score for each item, so the higher the bar, the better the functioning as perceived by the rater.

Common practice. There are a number of conventions of the FOP that are not transparent to a non-user, but which are important to understand. These conventions were followed in both the current and all previous studies of the FOP unless otherwise stated. First, all of the FOP items first include a dichotomous question inquiring about whether the domain in question is a problem for the survivor to which he/she and his/her significant other can answer either “yes” or “no”. If the rater reports that the area is problematic, they will be asked to provide the required 0-10 rating(s). If the rater reports that the area is *not* problematic he/she is not asked to provide any further ratings and the interviewer will move on to the next item. When this happens, the item is automatically given an outcome score of 10 (which is converted to 1.0) which indicates optimal functioning in the domain in question. While this scoring method makes clinical sense (i.e., if the person feels that there is no problem, they must be functioning optimally in the domain in question thus the score of 1.0 is fitting) it may cause problems psychometrically. Especially for problems that are less common, the high number of “no problem”/1.0 scores may lead to negatively skewed data.

Second, the FOP questions that compose each item in a subscale often ask for different types of ratings (i.e., all items in a subscale will generally not include only impact, satisfaction or activity questions). For example, the Emotional/Behavioural Functioning subscale (Items #36 – #41) includes 3 items that has Frequency and Impact questions, 1 item that has Frequency, Impact and Satisfaction questions, 1 item that has only a Satisfaction question and 1 item that has Satisfaction and Importance questions. The ratings provided for the Impact, Satisfaction and Importance questions become the outcome scores for the respective items and then the outcome scores get averaged as the Emotional/Behavioural Functioning subscale score. Thus, each subscale contains multiple types of information. Important to note again is that the Frequency ratings are *not* included in the outcome scores. These ratings provide information at the item level only (i.e., the Frequency scores are not grouped into subscales or used to compute a composite outcome score) that is kept independent of the other ratings provided.

Previous Psychometric Investigation. For a new measure, a considerable amount of psychometric analysis has been done with the FOP. All of the analyses completed to this point were carried out by Price (2007) and comprised his doctoral dissertation.

Analysis of FOP responses from community living ABI survivors ($N = 113$) has shown it to have good item properties. All 63 items were endorsed as problematic by at least one survivor, and 62 of the items had scores that spanned the entire range of the scale (0 – 1.0). All of the 49 FOP frequency items were also endorsed as problematic and 47 of the items had scores that spanned the entire 0-1.0 range. However, 25 of the 49 frequency items had median scores at ceiling level (1.0) (Price, 2007).

The FOP was found to have satisfactory internal reliability. The total outcome score was found to have satisfactory internal reliability ($\alpha = 0.94$) as were five of the 8 sub-scales (physical /ADL, cognitive, executive, emotional/behavioural, and social). Only 1 subscale (activities) appears to have weak, unsatisfactory internal reliability ($\alpha = 0.31$). The range for the remaining 7 subscales is from $\alpha = 0.61$ to 0.87 (Price, 2007).

The FOP was found to show good test-retest reliability in a subset of the original sample ($n = 25$) investigated by Price (2007). This group of survivors completed the FOP on two occasions separated by one week. The Pearson correlation was found to be very high for the composite outcome score ($r = 0.97$, $p < 0.001$), and correlations ranged from 0.68 to 0.96 for the 8 composite scores.

The FOP and FOP-SO scores from 22 survivor-proxy pairs were compared in order to investigate survivor-proxy agreement. This analysis was limited to the frequency scores as these are the only ratings that are made in reference to the same individual (the survivor) by both the survivor and the significant other. Spearman ρ correlation coefficients were calculated based on the ordinal scale of measure of the frequency scores. Survivor-proxy agreement for individual items ranged from $\rho = 0.03$ to 0.88. The correlations revealed that most items showed reasonably good agreement between survivors and significant others in regards to the areas that are problematic for survivors (Price, 2007).

Evidence for validity based on relations to other variables was obtained by comparing the FOP composite outcome score to the total outcome score of the MPAI-4. A moderately strong correlation in the expected direction was found between the two outcome scores ($r = -0.75$, $p < 0.001$). Therefore, individuals who have higher scores on the MPAI-4 (indicating higher levels of brain injury related problems) will most likely have lower scores on the FOP (indicating poorer functioning) (Price, 2007).

A Rasch analysis was also conducted on the FOP. Item-fit analysis found that only one item, Item #55 (Work quantity), did not fit the model. The FOP was found to have good person and item reliability indicating that the items have good discriminability and consistency. The person and item separation were also found to be good (Price, 2007).

Evaluation. The FOP has been submitted to some aspects of traditional psychometric analyses as well as Rasch analysis. To this point, it has been shown to be a reliable measure of functional outcome following brain injury. However it is important to note the unsatisfactory internal reliability of the Activities subscale, the poor survivor-proxy agreement on a number of items and the finding that Item #55 does not fit the Rasch model. In addition, the use of

Rasch in investigating the FOP is questionable. There are two main expectations of the Rasch model. The first is that an individual higher on the spectrum of the concept being measured (i.e., of higher ability) should consistently have a greater level of success on (or be more likely to endorse) any item than an individual lower on the spectrum (or, of lower ability). The second, related, expectation is that any individual should be more likely to endorse an easier, as opposed to a more difficult, item (Bezruczko, 2005). Unfortunately, these expectations do not apply to the FOP. Although the FOP items go from more "basic" abilities (e.g., gross motor functioning) to more "higher order" abilities (e.g., judgement) it is not an expectation that an ABI survivor who has problems with his/her gross motor movements will necessarily have problems with his/her judgement, for example. Therefore, due to the nature of brain injury and the information collected with the FOP, these underlying expectations of the Rasch model are not met.

However, the FOP does improve on many of the limitations found in existing outcome measures. First, it was designed to have a very broad range of item difficulty. As it was intended to be used with survivors who have returned to the community, and are thus "higher functioning" than those who may remain in long-term care facilities, it includes items taping domains relevant to this population of survivors. Also, the FOP is much more comprehensive not only including more "difficult" items, but also a good range of lower-level items as well. Obviously, the FOP is able to provide more information than any existing measure regarding the impact of the injury on the daily functioning of the survivor as well as the survivor's significant other(s). Yet, before the FOP is put to use in clinical or research settings other important aspects of reliability as well as evidence of its validity as a measure of outcome following brain injury must be investigated and reported. Therefore, the current study aims to provide this needed information and evidence.

Research Plan

The preliminary psychometric analyses conducted on the FOP have shown it to be a reliable measure with good psychometric properties. However a number of the underlying characteristics of the FOP have never been investigated. As such, the current analyses begin with exploratory investigations of the item distributions, the scoring of the items (i.e., the formulae used to create the outcome scores for the individual items) and the grouping of the items (i.e., the appropriateness of the subscales).

The study then investigates the psychometric characteristics of aspects of the FOP and provides additional evidence for the validity of the FOP to be used as a measure of functional outcome following brain injury. To conduct these analyses, a portion of the data utilized in a previous study (Price, 2007) was re-analyzed. The current study investigated three specific types of reliability in order to answer three questions:

1) Do the FOP outcome scores show acceptable internal reliability?
2) Do the FOP outcome scores show acceptable test-retest reliability?
3) Do the FOP and FOP-SO frequency ratings show acceptable survivor-proxy agreement?

Finally, in order for the FOP to be utilized in applied settings, it is necessary to demonstrate that it is a valid measure of functional outcome following brain injury. Therefore, different types of formal validity evidence were collected. Evidence based on the test content and external variables were collected and proof that the FOP outcome scores are sensitive to change over time was presented.

Chapter Two: Methods

Participants

The populations of interest for this study were community-dwelling, adult survivors of brain injury and their significant others. The data that were analyzed in the study were from a portion of the sample included in a previous study investigating the psychometric properties of the FOP (Price, 2007). A total of 107 survivors were included in the current study. These participants were recruited from: i) the Gorge Road Rehabilitation Hospital (GRH) in Victoria, BC ($N = 80$); and ii) advertisements in Vancouver, BC area newspapers ($N = 27$). The inclusion criteria, which were the same for the participants recruited from GRH and the advertisements, were that the participant had to be community dwelling, suffer from no addictions (alcohol or drugs), suffer from no psychoses, suffer from no conduct disorder, have some awareness of deficits, and have participated in rehabilitation secondary to their ABI (Price, 2007). The participants recruited from GRH were engaged in an out-patient rehabilitation program during the course of the study and were thus attending sessions during weekdays. The survivors recruited through GRH did not receive any incentives for their participation as the completion of the FOP comprised a portion of their rehabilitation. The survivors recruited through the newspapers received a $25 honorarium for each interview session.

The 107 participants’ whose data were included in the current study are of mixed injury severity (mild, moderate and severe injury) and mixed injury etiology. Table 1 includes available demographic information for the full sample. Unfortunately, demographic information for one of the 107 survivors was not recorded due to clerical error, and 5 others have incomplete demographic information as follows: two individuals do not have any – or

Table 1.

Demographic Information for Full Sample and Sub-samples.

		% Male	Age at injury	Time since injury (months)	Etiology – General	PTA (hours)	Injury Severity
			Mean (*SD*)	Mean (*SD*)	% (*n*)	Mean (*SD*)	% (*n*)
Full Sample N = 107		54.2	42.9 *(16.9)*	153.2 *(1062.7)*	ABI[a] – 41.3% (*43*) TBI – 53.8% (*56*) Other – 4.8% (*5*)	127.6 (*234.25*)	Mild – 58.4% (*59*) Moderate – 9.9% (*10*) Severe – 31.7% (*32*)
		(*n* = 106)	(*n* = 104)	(*n* = 103)	(*n* = 104)	(*n* = 101)	(*n* = 101)
Sub-sample: Survivor-proxy reliability (N = 20)		70%	43.3 (*13.5)*	53.7 (*61.5*)	ABI[a] – 75.0% (5) TBI – 25.0% (15)	121.7 (*207.6*)	Mild – 40.0% (8) Moderate – 15.0% (3) Severe – 45.0% (*9*)
		(*n* = 20)	(*n* = 20)	(*n* = 20)	(*n* = 20)	(*n* = 20)	(*n* = 20)
Sub-sample: Sensitivity to change (N = 33)	Short Interval (*n* = 18)	44.4%	38.2 (*11.7*)	82.4 (*66.6*)	ABI[a] – 27.8% (5) TBI – 72.2% (*13*)	121.0 (*186.5*)	Mild – 38.9% (*7*) Moderate – 11.1% (*2*) Severe – 50.0% (*9*)
		(*n* = 18)	(*n* = 18)	(*n* = 18)	(*n* = 18)	(*n* = 18)	(*n* = 18)
	Long Interval (*n* = 15)	66.7%	44.1 (*20.0*)	77.9 (*126.2*)	ABI[a] – 60.0% (9) TBI – 26.7% (*4*) Other – 26.7% (2)	180.0 (*283.1*)	Mild – 64.3% (*9*) Moderate – 0% (*0*) Severe – 35.7% (*5*)
		(*n* = 15)	(*n* = 15)	(*n* = 15)	(*n* = 15)	(*n* = 14)	(*n* = 14)

[a] *ABI without TBI*

have inaccurate – information recorded regarding their age at time of injury (i.e., age at time of injury and year of injury do not match); two individuals do not have a clear etiology recorded; for three participants the length of time post injury (specifically, the number of months) could not be determined; and 5 subjects did not know if they experienced any post-traumatic amnesia (PTA).

Twenty significant others who had also participated in the previous study (Price, 2007) were included in the current study. The inclusion criteria for the significant others was that they spend a minimum of 10 hours per week with the survivor. All significant others were interviewed within one week of the survivor's first FOP administration. The significant others recruited through the GRH did not receive any incentive for participation and completed the FOP-SO as a part of the survivors' rehabilitation program. However, the significant others recruited through the newspaper advertisements each received a $25 honorarium for their interview. Demographic information for the significant others was not collected.

Three sub-samples of the 107 participants were used for more specific analyses. First, the data from 20 survivor-significant other pairs were used to investigate the survivor-proxy reliability of the FOP. Table 1 contains demographic information for this subsample of survivors. Second, 18 survivors were administered the FOP twice within one week. This sub-sample will be referred to as the short interval group. These data were used to investigate test-retest reliability. Third, 15 survivors were administered the FOP twice with a 3 to 5 month interval (*M* = 3.67, *SD* = 0.724) during which time they were engaged in an out-patient rehabilitation program at GRH. This sub-sample will be referred to as the long interval group. These data, in conjunction with that from the short interval group, were used to investigate sensitivity to change. For these sub-samples, demographic information (see Table 1) was complete for all subjects except one who was unsure if he had experienced any PTA. Based on the information available, no significant differences were found between the short and long interval groups for gender ($\chi^2 = 1.63, p > 0.20$), age at injury ($t = 0.99, p > 0.30$), time since injury ($t = 0.13, p > 0.85$), or etiology ($\chi^2 = 7.39, p > 0.10$).

Measures

Demographic Questionnaire. All participants completed a demographic questionnaire in order to collect basic information about their age (in years), the time since injury (in years), the type of injury (TBI, ABI without TBI, or other), and length of post-traumatic amnesia (none or don't know, <1 hour, 1-24 hours, or >24 hours). Demographic information was not collected from the significant others.

Functional Outcome Profile (Survivor Version). All of the survivors completed the FOP. Analyses were done using individual item outcome scores for all 63 items, as well as the 8 subscale scores and composite outcome score. The frequency ratings from the 49 items that collect this information were also utilized.

Functional Outcome Profile (Significant Other Version). The 20 significant others who participated in the study completed the FOP-SO. Analyses were done using only the individual frequency ratings from the 49 items that collect this information; no analyses were done using the impact, satisfaction or importance ratings in the current study.

Procedure

The full sample ($N = 107$) was probed to determine which participants would be included in the three subsamples. As mentioned, data were available from 20 survivor-significant other pairs. These data were utilized in the investigation of survivor-proxy agreement. Also, from the full sample two groups of participants had been administered the

FOP twice. One group included 18 participants who were interviewed twice no more than one week apart. These participants composed the short interval group. The second group included 15 participants who were interviewed twice 3 to 5 months apart during which time they were engaged in a rehabilitation program. These participants composed the long interval group. Once the groups for all of the analyses were determined, the exploratory investigations and analyses were carried out on the data. The data from the short interval group were used to investigate test-retest reliability and the data from the short and long interval groups were compared to investigate sensitivity to change. The remainder of the analyses utilized the full sample unless otherwise indicated.

Analytic Strategy

Exploratory Investigations. The analyses began with exploratory investigations of specific aspects of the FOP to determine if any characteristics of the original FOP were inappropriate. First the item distributions were investigated, followed by the scoring used to create the item outcome scores, and then the item groupings (or subscales).

The first exploratory investigation was merely an examination of the distributions of the outcome scores for the items and scales (subscales and composite outcome score) from the first FOP administration for the full sample ($N = 107$). Means, standard deviations and skew were calculated and the later was investigated for normality. The full first administration data were then probed and a new data set was created that included only scores related to areas rated as problematic, which will be referred to as the “problem only” data. In other words, if a survivor reported that the domain in question was not a problem, he/she would automatically be assigned a score of 1.0 for that item – all of these scores (always 1.0) were removed from the data. These data points were then treated as missing data for the individual. As previously mentioned, it is likely that automatically giving an item outcome score of 1.0 to items that are not problematic (as is the common practice) may skew the data. Therefore, the skew of the item, subscale and composite outcome scores based on the "problem only" data were also investigated.

The second analysis focused on the scoring of the items to determine if the current score calculations were the most suitable. The FOP is composed of 3 broad question types: Impact questions, Satisfaction questions and Activity questions. The ratings provided for each of these question types are then converted into the outcome scores for the individual items using formulae created by the authors of the FOP. To determine if the original formulae provide the "best" outcome scores they were compared to newly developed formulae based on different weightings and interactions between the different ratings. The scores produced by the original and newly created formulae were compared to "intuitive outcome scores". The intuitive outcome scores were created using 3 x 3 matrices of the interaction between low (1.0), moderate (5.0), and high (9.0) ratings based on my interpretation of how they should interact. The intuitive outcome scores were created with the goals of the FOP (i.e., the focus on functioning) in mind. The objective of these analyses was to see if a formula could be derived that would improve the outcome score (i.e., get the score to more closely reflect the intuitive matrix and thus the goals of the FOP) for each question type.

The third exploratory analysis was aimed at investigating the item groupings, or subscales, to determine if more appropriate groupings were available. Because the FOP is both multidimensional and multi-methodological, an investigation needed to be done to ensure that the most appropriate information was obtained when using the original subscale scores. In other words, do the existing subscales present the most informative summary of the information provided, or could the items be grouped in a more revealing manner? The author devised 3 options for grouping the items which were then investigated using the first FOP administration data for the full sample ($N = 107$). The first grouping was based on theory and/or convention and represented the existing subscales that were originally created during the development of the FOP. The second grouping method was based on type of information obtained (i.e., type of question) in an attempt to address the multi-methodological aspect of the FOP. The final grouping was based on specific psychological constructs in order to address the multidimensional aspect of the FOP. Once the groupings were determined, the outcome score for each item in the new group/subscale was correlated with the outcome score for the relevant grouping/subscale, and finally each grouping/subscale outcome score was to be correlated with the composite outcome score.

Reliability. The planned investigation of three types of reliability of two aspects of the FOP was completed: internal consistency and test-retest reliability of the outcome scores and survivor-proxy agreement of the frequency ratings was checked. Standard psychometric approaches were used to investigate each type of reliability. To investigate the internal consistency of the FOP outcome scores, each item outcome score was correlated (using Pearson product-moment correlations) with its related subscale score and with the composite outcome score to investigate the coherence of the items. Each subscale score was then also correlated (using Pearson correlations) with the composite outcome score. For all of these analyses, correlations were categorized as strong ($r > 0.51$), moderate ($0.26 < r < 0.50$), weak ($0.11 < r < 0.25$), as having no relationship ($0.00 < r < 0.10$), or has having a negative relationship ($r < 0.00$). As another investigation of the internal consistency of the outcome scores, Cronbach’s alpha was calculated for the composite outcome score and each of the subscales (based on the results of the exploratory analyses). To determine if the high number of items per survivor rated as not problematic (and thus automatically given an item outcome score of 1.0) were inflating the internal consistency statistics, Cronbach's alpha was re-calculated using the "problem only" data (as described above).

To establish test-retest reliability of the outcome scores, Pearson product-moment correlations were calculated between Time 1 and Time 2 data from the short-interval group (*n* = 18), for each item, subscale, and for the composite outcome score. This same analysis was then re-run using the "problem only" data to again see if the individuals reporting "no problem" on items were inflating the correlations.

Finally, the survivor-proxy agreement of the frequency ratings provided by the survivors and their significant others was investigated using 3 analyses. First, Spearman correlations were calculated using the frequency ratings from the 20 survivors and 20 significant others. Second, the original frequency ratings were re-coded on a three point scale such that 0 represented "no problem", 0.5 represented "sometimes a problem" and 1 represented "always a problem" (a scale which is more similar to that used by existing outcome measures than the 11-point scale utilized by the FOP) and associations were determined using Cramer's V for nominal variables. Finally, the percent agreement was calculated between the survivors' and the significant others' reports of whether the area was a problem (frequency rating from 0.1 to 1.0) or not (frequency rating of 0.0).

Validity. Finally, evidence for the valid utilization of the FOP as a measure of outcome following brain injury will be presented. First, evidence was obtained based on the test content. This evidence was based on: a) development-based information; b) an objective investigation of item usage (specifically, the number of items reported as problematic by at least one survivor and significant other); and c) a probe of Item #61 which inquires about any Other Problems being experienced that the survivor or significant other feel were not discussed. It was expected that the evidence based on test content would demonstrate that the FOP items are relevant to brain injury survivors and their significant others and that the FOP has sufficient breadth as it includes items that collect information about most problematic domains following brain injury.

Second, evidence for the validity of the FOP outcome scores based on external variables was investigated. As previously mentioned, many person- and injury-characteristics have been consistently shown to be related to outcome following ABI. Therefore, the influence of specific demographic variables on the FOP subscale and composite outcome scores were investigated. All continuous independent variables were converted to discrete variables (ordinal scale of measurement) for ease of interpretation. The independent variables were gender, etiology (three levels: TBI, ABI not including TBI, and other), PTA severity (three levels: mild, moderate and severe), age at time of injury (three levels: under 25, 25-50, and over 50), and time since injury (five levels: under 6 months, 6-12 months, 13-24 months, 25-48 months, over 48 months). The effect of these variables on the FOP composite outcome score was investigated using a MANOVA. It is expected that these variables will be found to differentially impact the FOP subscale and composite outcome scores, suggesting that the FOP is measuring outcome as the developers intended it to.

Third, sensitivity to change over time of the outcome scores was investigated with a 2 x 2 x 9 Repeated Measures ANOVA. This investigation compared the subscale and composite outcome scores from Time 1 and Time 2 for the short and long interval groups ($N = 33$). A

Time x Interval interaction is expected, indicating that a difference between Time 1 and Time 2 scores depends on interval group.

In addition, the smallest detectable difference (SDD) was calculated. The SDD is based on the standard error of the mean (SEM), and is calculated with 95% confidence when using the formula:

$$1.96 \times \sqrt{2} \times \text{SEM}$$

where 1.96 is the *z*-score associated with an area under the curve of .05 (i.e., when $p = .05$). Only mean differences between the two test administrations that are larger than the SDD can be interpreted with 95% certainty as real change (van Baalen et al., 2006). Therefore, SDDs were calculated for each subscale and the composite outcome score and the mean differences from Time 1 to Time 2 for the short and long interval groups were then compared to the SDD for each respective scale.

Finally, change was also investigated using only the data where change was possible. For this analysis, data points were removed and regarded as missing data if the survivor had no problem in an area at both Time 1 and Time 2 (i.e., scores of 1.0 on an item at both administrations). Using only this “change possible” data, the average change was investigated for the short and long interval groups using an Independent Samples t-test with the subscale and composite outcome mean difference scores as the dependent variables and the interval group as the grouping variable.

Chapter Three: Results

Exploratory Investigations

First, the characteristics of the outcome scores for the items, subscales and composite scale (i.e., the full measure) were investigated by calculating descriptive statistics including the skew for all items and scales. The means of the outcome scores for the 63 items ranged from .463 (Item 21, Anterograde Memory) to .983 (Item 15, Drug Abuse) while Standard Deviations ranged from .107 (Item 15, Drug Abuse) to .440 (Item 13, Psychotropic Medication). The means of the outcome scores for the subscales had a much more limited range with values between .621 (Physical/ADL Subscale) and .791 (Social Subscale) with Standard Deviations ranging from .124 (Activities Subscale) to .246 (Executive Subscale). Most of the items were negatively skewed as the responses were clustered above the mid-point, and 19 of the items were found to have a negative skew greater than -1.00 (but less than -2.00). However, 4 items were found to be severely negatively skewed, greater than -2.00 with Item 15 (Drug Abuse) having a skew of -7.890. Only one of the subscales was found to have a skew greater than -1.00 (Social Subscale, skew = -1.146). The composite outcome score had a mean of 0.690, a standard deviation of 0.156 and a skew of -0.446.

A reinvestigation of the distributions of the item outcome scores using the “problem only” data found that for the 63 items the skew ranged from -1.553 (Item 15, Drug Abuse) to 1.437 (Item 31, Psychotropic medications). However, the number of skewed items was greatly reduced; only 9 items were found to have skews greater than +/- 1.00. Similarly, all of the subscales had skews between +1.00 and -1.00, and the composite outcome score was normally distributed with a skew of 0.04 when the “problem only” data was used.

Second, the formulas used to calculate the individual item outcome scores were investigated. As previously mentioned, the items were divided into three groups based on the type of questions each included: Impact questions ($n = 46$), Satisfaction questions ($n = 13$), and Activity questions ($n = 4$). Originally the impact items were scored as such:

$$1 - (\text{Impact}/10)$$

such that an impact rating of 10 (i.e., “couldn’t have more negative impact on my daily life”) is converted to an outcome score of 0 (indicating poor functioning) for that item. Because the Impact questions also inquire about problem frequency (which is not currently included in the computation for the impact items) a 3 x 3 matrix was created to display how I intuitively thought problem frequency and impact should interact, and what the resulting outcome scores should be (Table 2). Two new formulas, which included the problem frequency ratings, were then created. The first formula (which created “Outcome 2”) was:

$$[1 - (\text{Frequency}/10)] \times [1 - (\text{Impact}/10)]$$

so that frequency and impact are equally important in the item outcome score. The second newly derived formula (which created “Outcome 3”) was:

$$[1 - (\text{Frequency}/10)^2] \times [1 - (\text{Impact}/10)]$$

so that the main contributor to the item outcome score is impact, but frequency (which, because working with decimals will be decreased by squaring) will still contribute to the score. Table 3 shows the results of the different equations based on low, medium, or high frequency and impact scores. These results show that none of the equations match the conceived interaction between these elements. However, the original equation (“Outcome 1”) comes the closest, especially for high frequency – low impact items.

Table 2.

Conceived item scores based on interaction between problem frequency and impact

		Impact		
		Low	Medium	High
Frequency	Low	1.00	0.70	0.30
	Medium	0.70	0.50	0.10
	High	0.70	0.30	0.00

Table 3.

Comparison of possible Impact formulas to conceived interaction between problem frequency and impact.

Type of Interaction	Frequency Rating	Impact Rating	Conceived Outcome	Outcome 1[a]	Outcome 2[b]	Outcome 3[c]
Low/Low	1.0	1.0	1.00	0.90	0.81	0.89
Low/Medium	1.0	5.0	0.70	0.50	0.45	0.50
Low/High	1.0	9.0	0.30	0.10	0.09	0.10
Medium/Low	5.0	1.0	0.70	0.90	0.45	0.68
Medium/Medium	5.0	5.0	0.50	0.50	0.25	0.38
Medium High	5.0	9.0	0.10	0.10	0.05	0.08
High/Low	9.0	1.0	0.70	0.90	0.09	0.17
High/Medium	9.0	5.0	0.30	0.50	0.05	0.10
High/High	9.0	9.0	0.00	0.10	0.01	0.02

a *Outcome 1* = 1 – (Impact/10)

b *Outcome 2* = [1 – (Frequency/10)] x [1 – (Impact/10)]

c *Outcome 3* = $[1 - (\text{Frequency}/10)^2] \times [1 - (\text{Impact}/10)]$

The original equation for the Satisfaction questions was:

$$1 - [(1 - \text{Satisfaction}/10) \times (\text{Importance}/10)]$$

such that the amount of dissatisfaction is moderated by the degree of importance. Therefore, when there is a high degree of dissatisfaction with a domain of high importance, the outcome score will be low. On the other hand, when there is a high degree of dissatisfaction with a domain of low importance, the outcome score will be higher. In other words, dissatisfaction with something ‘means more’ when that thing is felt to be important. Based on this logic, the equation makes intuitive sense and seems to account for all of the information being provided. Thus, I created another 3 x 3 matrix for the interaction between satisfaction and importance and compared these conceived scores to what would actually be obtained given low, medium, and high ratings of satisfaction and importance (Table 4). Again, the conceived scores are relatively close to the actual scores produced by the original formula.

Table 4.

Comparison of Satisfaction formula to conceived interaction between issue satisfaction and importance.

Type of Interaction	Satisfaction Rating	Importance Rating	Conceived Outcome	Outcome 1[a]
Low/Low	1.0	1.0	0.7	0.91
Low/Medium	1.0	5.0	0.3	0.55
Low/High	1.0	9.0	0.0	0.19
Medium/Low	5.0	1.0	0.8	0.95
Medium/Medium	5.0	5.0	0.5	0.75
Medium/High	5.0	9.0	0.3	0.55
High/Low	9.0	1.0	1.0	0.99
High/Medium	9.0	5.0	1.0	0.95
High/High	9.0	9.0	1.0	0.91

a *Outcome 1* = 1 – [(1 – Satisfaction/10) x (Importance/10)]

The formula used to calculate the scores for the Activity questions is much less straightforward than those for the Impact or Satisfaction questions. The first component of the outcome scores based on the Activity questions is the "Activity Credit Score". The Activity Credit Score is based on the principle of giving survivors "full credit" for activities completed independently and "partial credit" for activities completed with help. Included in this principle is the assumption that, if an individual is getting help with an activity 100% of the time, they are contributing 50% of the total time spent completing the activity. For example, if a survivor requires no help for an activity then he/she gets 1 hour credit for each hour per day spent completing the activity. If, on the other hand, an individual requires help 100% of the time to complete a particular activity, he/she will get ½ an hour credit for each hour per day spent completing the activity. Therefore, the Activity Credit Score is computed using the following formula:

$$\frac{1}{2}(T \times H\%) + [T - (T \times H\%)]$$

based on the concept of receiving half-credit for time that help is required $(T*H\%)$ and full credit for the remainder of the time (i.e., time that help is not required) $(T-T*H\%)$ where T is the hours per day spent completing the activity, and $H\%$ is the percent of time that help is required. Therefore, as the amount of help required decreases the Activity Credit Score will increase. For example, an individual who spends 3 hours a day completing an activity with no help (i.e., complete independence) would get an Activity Credit Score of 3 hours per day. An individual who spends 3 hours a day completing an activity with 100% help (i.e., complete dependence) would get an Activity Credit Score of 1.5 hours per day. Finally, an individual who spends 3 hours completing an activity with help 25% of the time (i.e., a large degree of independence, but some help is still required) would get an Activity Credit Score of 2.625

hours per day. The activity outcome score is then based on the obtained Activity Credit Score as outlined in Table 5. In the absence of sufficiently detailed norms in the literature regarding the "appropriate" amount of time spent engaged in particular types of activities, the developers chose ranges that seemed "reasonable and generous" (Skelton et al., 2005). The logic that underlies the activity outcome scores seems reasonable to the author and so the calculation was not altered. Overall, the original formulae for the three question types were retained and have been shown to provide the most logical information given the purpose of the FOP (i.e., to be a measure of functioning and not deficit).

Table 5.

Calculation of Activity Outcome Scores based on Activity Credit Scores

	Outcome Score			
Activity Type	0.00	0.33	0.66	1.00
Self-Care	0 hours OR More than 8 hours	0-1 hours OR 6-8 hours	1-2 hours OR 4-6 hours	2-4 hours
Leisure Solitary	More than 14 hours	5-14 hours	3-5 hours	0-3 hours
Leisure Active/Social	0 hours OR More than 14 hours	0-1 hours OR 11-14 hours	1-2 hours OR 9-11 hours	2-9 hours
Work Quantity	0 hours OR More than 14 hours	0-1 hours OR 11-14 hours	1-2 hours OR 9-11 hours	2-9 hours

The final exploratory analyses investigated the item groupings, or subscales. Three grouping options were derived as previously described. These 3 grouping options divided the items in the following manners: the first option (based on theory and/or convention) was the original 8 subscales -- Physical/ADL, Health, Cognitive, Executive, Emotional/Behavioural, Social, Activities, and Overall. The second option (based on type of question) resulted in 3 groups: Impact, Satisfaction and Activity. The final grouping option (based on specific constructs) resulted in 18 groups: Physical, ADL, Sleeping, Health, Medication, Attention, Language, Memory, Planning/Acting, Inhibition, Daily Independence, Awareness, Emotion, Relationships, Leisure, Work, Safety, and Satisfaction.

Pearson correlations were calculated for the outcome score of each item with the outcome score of its relevant group/subscale using the first-administration data for the full sample ($N = 107$). The correlations between the outcome scores for items and the relevant original subscales revealed correlations ranging from -.28 to .81 (Table 6) with 53 of the items having a strong correlation with their relevant subscales ($r > .51$), 5 items having a moderate relationship ($.26 < r < .50$), 1 item having a weak relationship $0.11 < r < .25$), 1 item having no relationship ($r < .10$), and 3 items having negative correlations with their relevant original subscale. The same process was repeated with the outcome scores for the items and the relevant new question-type groups (Table 7). These correlations ranged from .12 to .82 with 36 of the items having a strong relationship, 24 of the items having a moderate relationship, and 3 of the items having a weak relationship. Finally, when each the outcome scores for the items were correlated with the score for the relevant construct-based group (Table 8), the correlations ranged from .11 to .88 with 59 items having a strong relationship with their relevant construct group score, 3 items having

Table 6.

Pearson Correlations between each item and relevant subscale score

Subscale	Item		Pearson correlation
Physical, ADL	1	Senses	0.64
	2	Gross Motor	0.65
	3	Fine Motor	0.49
	4	Self-care	0.65
	5	Eating & Meal Preparation	0.72
	6	Domestic Skills	0.65
	7	Fatigue	0.59
	8	Sleeping	0.49
Health	9	Medical Conditions	0.52
	10	Pain	0.62
	11	Health Satisfaction	0.55
	12	Pain Medications	0.71
	13	Psychotropic Medications	0.64
	14	Medication Side Effects	0.58
	15	Drug Abuse	0.18
Cognitive	16	Sustained Attention	0.66
	17	Divided Attention	0.76
	18	Language Expression	0.74
	19	Language Comprehension	0.62
	20	Information Processing	0.72
	21	Anterograde Memory	0.81
	22	Retrograde Memory	0.59
	23	Procedural Memory	0.67
	24	Time Disorientation	0.62
	25	Wayfinding	0.53
Executive	26	Judgement	0.65
	27	Cognitive Organization	0.80
	28	Confusability	0.76
	29	Impulsivity	0.72
	30	Behavioural Organization	0.75
	31	Initiation	0.68
	32	Time & Punctuality	0.62
	33	Money Management	0.59
	34	Self-Insight	0.69
	35	Coping	0.56
Emotional, Behavioural	36	Emotional Reactivity	0.75
	37	Problem Behaviour	0.43
	38	Emotional/Cognitive Intrusions	0.78
	39	Mood	0.81
	40	Self Esteem	0.72
	41	Spirituality	0.58

Social	42	Family Relationships	0.59
	43	Relationship Satisfaction	0.64
	44	Physical Intimacy	0.55
	45	Physical Intimacy Satisfaction	0.57
	46	Social Life	0.63
	47	Social Relationships	0.61
	48	Community/Social Interactions	0.69
	49	Community Barriers	0.41
	50	Community Involvement	0.55
Activities	51	Self-Care Activities	0.63
	52	Leisure: Solitary	0.41
	53	Leisure: Active/Social	0.68
	54	Leisure Satisfaction	0.03
	55	Work, Quantity	0.61
	56	Work Satisfaction	-0.10
	57	Work Demands	-0.28
	58	Safety Concerns by Others	0.08
	59	Safety Concerns by Self	-0.13
Overall	60	Economic Security	0.71
	61	Other Problems	0.65
	62	Life Satisfaction	0.79
	63	Good Things	0.58

Table 7.

Pearson Correlations between each item and relevant question-type grouping score

Subscale	Item		Pearson correlation
Impact	1	Senses	0.49
	2	Gross Motor	0.35
	3	Fine Motor	0.23
	4	Self-care	0.40
	5	Eating & Meal Preparation	0.48
	6	Domestic Skills	0.47
	7	Fatigue	0.60
	9	Medical Conditions	0.23
	10	Pain	0.45
	12	Pain Medications	0.38
	13	Psychotropic Medications	0.41
	14	Medication Side Effects	0.34
	15	Drug Abuse	0.12
	16	Sustained Attention	0.66
	17	Divided Attention	0.71
	18	Language Expression	0.58
	19	Language Comprehension	0.44
	20	Information Processing	0.50
	21	Anterograde Memory	0.68
	22	Retrograde Memory	0.44
	23	Procedural Memory	0.50
	24	Time Disorientation	0.65
	25	Wayfinding	0.52
	26	Judgement	0.61
	27	Cognitive Organization	0.79
	28	Confusability	0.71
	29	Impulsivity	0.62
	30	Behavioural Organization	0.71
	31	Initiation	0.61
	32	Time & Punctuality	0.56
	33	Money Management	0.57
	34	Self-Insight	0.57
	35	Coping	0.55
	36	Emotional Reactivity	0.58
	37	Problem Behaviour	0.29
	38	Emotional/Cognitive Intrusions	0.67
	41	Spirituality	0.48
	42	Family Relationships	0.52
	44	Physical Intimacy	0.31
	47	Social Relationships	0.45
	48	Community/Social Interactions	0.57
	49	Community Barriers	0.29

	57	Work Demands	0.37
	58	Safety Concerns by Others	0.28
	59	Safety Concerns by Self	0.39
	61	Other Problems	0.44
	63	Good Things	0.27
Satisfaction	8	Sleeping	0.60
	11	Health Satisfaction	0.56
	39	Mood	0.75
	40	Self Esteem	0.76
	43	Relationship Satisfaction	0.60
	45	Physical Intimacy Satisfaction	0.56
	46	Social Life	0.58
	50	Community Involvement	0.53
	54	Leisure Satisfaction	0.69
	56	Work Satisfaction	0.70
	60	Economic Security	0.61
	62	Life Satisfaction	0.82
Activity	51	Self-Care Activities	0.63
	52	Leisure: Solitary	0.41
	53	Leisure: Active/Social	0.68
	55	Work, Quantity	0.61

Table 8.

Pearson Correlations between each item and relevant construct-based grouping score

Grouping	Item		Pearson Correlation
Physical	1	Senses	0.71
	2	Gross Motor	0.78
	3	Fine Motor	0.72
ADL	4	Self-Care	0.67
	5	Eating & Meal Preparation	0.80
	6	Domestic Skills	0.72
	51	Self-Care Activities	0.40
Sleeping	7	Fatigue	0.79
	8	Sleeping	0.83
Health	9	Medical Conditions	0.76
	10	Pain	0.75
Medication	12	Pain Medications	0.71
	13	Psychotropic Meds	0.79
	14	Med Side Effects	0.73
Attention	16	Attention, Sustained	0.77
	17	Attention, Divided	0.87
	20	Information Processing Speed	0.77
Language	18	Language Expression	0.88
	19	Language Comprehension	0.86
Memory	21	Memory, Anterograde	0.83
	22	Memory, Retrograde	0.78
	23	Procedural Learning	0.75
Planning/Acting	26	Judgement	0.68
	28	Confusability	0.83
	30	Organization, Behavioural	0.79
	31	Initiation	0.74
Inhibition	15	Drug Abuse	0.33
	27	Organization, Cognitive	0.82
	29	Impulsivity	0.82
	37	Problem Behaviour	0.61
Independence	24	Disorientation, Time	0.80
	25	Way finding	0.68
	32	Time and Punctuality	0.75
	33	Money Management	0.73
	49	Community Barriers	0.44
Awareness	34	Self-Insight	0.85
	61	Other Problems	0.11
Emotion	35	Coping	0.83
	36	Emotional Reactivity	0.74
	38	Intrusions: Emotional and Cognitive	0.81
	39	Mood (Satisfaction)	0.83

Relationships	42	Relationships, Family	0.70
	44	Physical Intimacy Problems	0.57
	47	Social Relationships	0.77
	48	Community/Social Interactions	0.69
Leisure	52	Leisure: Solitary	0.63
	53	Leisure: Active/Social	0.82
Work	55	Work, Quantity	0.64
	57	Work Demands	0.56
Safety	58	Safety concerns, by Others	0.78
	59	Safety Concerns, by Self	0.82
Satisfaction	11	Health Satisfaction	0.59
	40	Self-Esteem	0.75
	41	Spirituality	0.61
	43	Relationships, Satisfaction	0.61
	45	Physical Intimacy Satisfaction	0.59
	46	Social Life	0.58
	50	Community Involvement	0.54
	54	Leisure Satisfaction	0.69
	56	Work	0.74
	60	Economic Security	0.58
	62	Life Satisfaction	0.81
	63	Good Things	0.52

moderate relationships, and 1 item having a weak relationship. These results imply that the construct-based grouping is the best option and the question-type grouping is the worst.

The score for each group/subscale was then correlated with the composite outcome score. For the original subscales (Table 9), the correlations with the composite outcome score ranged from .60 to .90 indicating that all of the subscales have a strong correlation with the composite outcome score. The correlations between the question-type groupings and the composite outcome score (Table 10) ranged from .03 to .99. Lastly, the correlations between the construct-based groupings and the composite outcome score (Table 11) ranged from .08 to .84 with 13 of the 18 groups having a strong relationship with the composite outcome score. These results imply that the original subscales are the best option and that the question-type groupings are the worst. Overall, no grouping option was consistently the best choice. However, the scores for the original subscales generally showed good relationships within the scales (i.e., the relationships between the item scores and the subscale scores) as well as good relationships to the composite outcome score. Furthermore, the original subscales are easily understood, highly conventional and similar to those found in existing outcome measures. For these reasons, I decided that the original subscales would be retained.

Table 9.

Pearson Correlations for the original subscale score with the composite outcome score

Subscale	Pearson Correlation
Physical, ADL	0.72
Health	0.62
Cognition	0.81
Executive	0.90
Emotional/Behavioural	0.84
Social	0.78
Active	0.60
Overall	0.74

Table 10.

Pearson correlations between question-type groupings and composite outcome score

Question-type grouping	Pearson correlation
Activities	0.03
Impact	0.99
Satisfaction	0.85

Table 11.

Pearson correlations between construct-based groupings and composite outcome score

Construct-based grouping	Pearson correlation
Physical	0.48
ADL	0.57
Sleeping	0.69
Health	0.46
Medication	0.51
Attention	0.76
Language	0.57
Memory	0.65
Planning/Acting	0.84
Inhibition	0.74
Independence	0.74
Awareness	0.60
Emotion	0.78
Relationships	0.69
Leisure	0.08
Work	0.18
Safety	0.42
Satisfaction	0.82

Reliability

A portion of the psychometric analyses previously completed by Price (2007) were reinvestigated to ensure that similar results were obtained and to reaffirm the reliability of FOP scores. First, to determine the internal consistency of the FOP outcome scores, the coherence of the item scores was investigated by correlating the score for each item with the outcome score from the first administration of the FOP for the full sample ($N = 107$) (see

Table 12). The correlations ranged from -.13 to 0.78 with 28 items having a strong relationship with the composite outcome score ($r > .51$), 28 items having a moderate relationship ($.26 < r < .51$), 4 items having a weak relationship ($.11 < r < .25$) 2 items appearing to have no relationship with the composite outcome score ($r < .10$) and 1 item with a negative correlation. Pearson correlations for the scores of each item with its relevant subscale score were calculated initially for the exploratory investigation of the item groupings, but these results also relate to the FOP scores' internal consistency. As previously reported, it was found that 1 item had a weak relationship with its subscale and another three were found to have negative correlations with their subscales (refer back to Table 6 for the full results).

Because of the rather disappointing results, an additional, post hoc analysis was run to further investigate the appropriateness of the individual items. Based on the assumption that recovery from brain injury is a heterogeneous process, such that individual survivors experience individual symptom constellations, item-composite score correlations were calculated for the Vancouver and Gorge samples separately. These samples were not thought to, as groups, differ greatly. However, one important difference is that at the time of the first administration, the Gorge group was beginning a rehabilitation program while the Vancouver group was not. For this reason it was thought that the groups may differ on the problems experienced, the perceived severity or importance of certain problems, or even the awareness of problems. As Table 13 illustrates, there does seem to be a difference between these two groups in the degree of relatedness between the outcome scores for individual items and the composite outcome score. For the Vancouver group, the correlations ranged from -.28 to .82; for the Gorge group the correlations ranged from -.28 to .80. Four of the items were within the

Table 12.

Pearson correlations between each item and composite outcome score

FOP Item		Pearson r	Significance
1	Senses	0.49	< 0.001
2	Gross Motor	0.36	< 0.001
3	Fine Motor	0.21	0.027
4	Self-care	0.39	< 0.001
5	Eating & Meal Preparation	0.48	< 0.001
6	Domestic Skills	0.47	< 0.001
7	Fatigue	0.60	< 0.001
8	Sleeping	0.53	< 0.001
9	Medical Conditions	0.23	0.019
10	Pain	0.47	< 0.001
11	Health Satisfaction	0.44	< 0.001
12	Pain Medications	0.38	< 0.001
13	Psychotropic Medications	0.41	< 0.001
14	Medication Side Effects	0.33	< 0.001
15	Drug Abuse	0.11	0.262
16	Sustained Attention	0.66	< 0.001
17	Divided Attention	0.69	< 0.001
18	Language Expression	0.57	< 0.001
19	Language Comprehension	0.42	< 0.001
20	Information Processing	0.48	< 0.001
21	Anterograde Memory	0.66	< 0.001
22	Retrograde Memory	0.40	< 0.001
23	Procedural Memory	0.47	< 0.001
24	Time Disorientation	0.63	< 0.001
25	Wayfinding	0.49	< 0.001
26	Judgement	0.58	< 0.001
27	Cognitive Organization	0.78	< 0.001
28	Confusability	0.67	< 0.001
29	Impulsivity	0.59	< 0.001
30	Behavioural Organization	0.70	< 0.001
31	Initiation	0.59	< 0.001
32	Time & Punctuality	0.55	< 0.001
33	Money Management	0.53	< 0.001
34	Self-Insight	0.55	< 0.001
35	Coping	0.57	< 0.001
36	Emotional Reactivity	0.58	< 0.001
37	Problem Behaviour	0.27	< 0.001
38	Emotional/Cognitive Intrusions	0.68	< 0.001
39	Mood	0.68	< 0.001
40	Self Esteem	0.70	< 0.001
41	Spirituality	0.52	< 0.001
42	Family Relationships	0.50	< 0.001

43	Relationship Satisfaction	0.55	< 0.001
44	Physical Intimacy	0.36	< 0.001
45	Physical Intimacy Satisfaction	0.38	< 0.001
46	Social Life	0.49	< 0.001
47	Social Relationships	0.46	< 0.001
48	Community/Social Interactions	0.58	< 0.001
49	Community Barriers	0.33	< 0.001
50	Community Involvement	0.41	< 0.001
51	Self-Care Activities	0.11	0.244
52	Leisure: Solitary	0.08	0.414
53	Leisure: Active/Social	0.04	0.655
54	Leisure Satisfaction	0.55	< 0.001
55	Work, Quantity	-0.13	0.187
56	Work Satisfaction	0.57	< 0.001
57	Work Demands	0.36	< 0.001
58	Safety Concerns by Others	0.28	0.003
59	Safety Concerns by Self	0.38	< 0.001
60	Economic Security	0.58	< 0.001
61	Other Problems	0.45	< 0.001
62	Life Satisfaction	0.66	< 0.001
63	Good Things	0.32	< 0.001

Table 13.

Item-composite outcome score correlations for the Gorge sample and the Vancouver sample separately

	Gorge Sample		Vancouver Sample	
Item	Pearson r	Significance	Pearson r	Significance
1 Senses	0.42	< 0.001	0.72*	< 0.001
2 Gross Motor	0.32§	0.004	0.50	0.008
3 Fine Motor	0.15§	0.175	0.40	0.037
4 Self-care	0.38	0.001	0.45	0.018
5 Eating & Meal Preparation	0.47	< 0.001	0.52	0.006
6 Domestic Skills	0.45	< 0.001	0.55	0.003
7 Fatigue	0.61	< 0.001	0.56	0.003
8 Sleeping	0.54	< 0.001	0.48	0.012
9 Medical Conditions	0.27§	0.016	0.01§	0.957
10 Pain	0.42	< 0.001	0.63*	< 0.001
11 Health Satisfaction	0.55	< 0.001	-0.01§	0.953
12 Pain Medications	0.42	< 0.001	0.28	0.152
13 Psychotropic Medications	0.48	< 0.001	0.18	0.373
14 Medication Side Effects	0.33§	0.003	0.39	0.045
15 Drug Abuse	0.18§	0.111	0.11	0.599
16 Sustained Attention	0.69*	< 0.001	0.57	0.002
17 Divided Attention	0.73*	< 0.001	0.59	0.001
18 Language Expression	0.58	< 0.001	0.54	0.004
19 Language Comprehension	0.50	< 0.001	0.01§	0.975
20 Information Processing	0.50	< 0.001	0.41	0.035
21 Anterograde Memory	0.65	< 0.001	0.74*	< 0.001
22 Retrograde Memory	0.37	0.001	0.52	0.005
23 Procedural Memory	0.53	< 0.001	0.28	0.163
24 Time Disorientation	0.64	< 0.001	0.59	0.001
25 Wayfinding	0.48	< 0.001	0.54	0.004
26 Judgement	0.55	< 0.001	0.73*	< 0.001
27 Cognitive Organization	0.80*	< 0.001	0.82*	< 0.001
28 Confusability	0.69*	< 0.001	0.67*	< 0.001
29 Impulsivity	0.66	< 0.001	0.51	0.006
30 Behavioural Organization	0.75*	< 0.001	0.57	0.002
31 Initiation	0.58	< 0.001	0.68*	< 0.001
32 Time & Punctuality	0.62	< 0.001	0.33	0.091
33 Money Management	0.57	< 0.001	0.38	0.052
34 Self-Insight	0.50	< 0.001	0.73*	< 0.001
35 Coping	0.59	< 0.001	0.54	0.004
36 Emotional Reactivity	0.61	< 0.001	0.49	0.009
37 Problem Behaviour	0.36	0.001	0.07§	0.726
38 Emotional/Cognitive Intrusions	0.69*	< 0.001	0.69*	< 0.001
39 Mood	0.68*	< 0.001	0.69*	< 0.001
40 Self Esteem	0.76*	< 0.001	0.55	0.003

41	Spirituality	0.59	< 0.001	0.09§	0.644
42	Family Relationships	0.54	< 0.001	0.57	0.002
43	Relationship Satisfaction	0.68*	< 0.001	0.21	0.297
44	Physical Intimacy	0.42	< 0.001	0.02§	0.940
45	Physical Intimacy Satisfaction	0.49	< 0.001	-0.10§	0.617
46	Social Life	0.56	< 0.001	0.18	0.357
47	Social Relationships	0.47	< 0.001	0.48	0.011
48	Community/Social Interactions	0.63	< 0.001	0.34	0.086
49	Community Barriers	0.34	0.002	0.36	0.069
50	Community Involvement	0.44	< 0.001	0.29	0.147
51	Self-Care Activities	0.08§	0.486	0.28	0.156
52	Leisure: Solitary	0.03§	0.789	0.28	0.160
53	Leisure: Active/Social	0.01§	0.958	0.21	0.287
54	Leisure Satisfaction	0.64	< 0.001	0.29	0.141
55	Work, Quantity	-0.23§	0.042	0.32	0.102
56	Work Satisfaction	0.61	< 0.001	0.45	0.020
57	Work Demands	0.44	< 0.001	0.03§	0.889
58	Safety Concerns by Others	0.33§	0.003	0.08§	0.696
59	Safety Concerns by Self	0.38	< 0.001	0.38	0.049
60	Economic Security	0.58	< 0.001	0.58	0.001
61	Other Problems	0.52	< 0.001	-0.28§	0.157
62	Life Satisfaction	0.68*	< 0.001	0.61	0.001
63	Good Things	0.37	0.001	0.12	0.566

* the 10 highest correlations

- § the 10 lowest correlations

10 strongest correlations for both groups, and for both groups the strongest correlation was between the score for Item #27 (Cognitive Organization) and the composite outcome score. However, as Figure 3 illustrates, considerable differences can be seen between the two groups when the item-composite outcome score correlations are compared. For example, Item #61 (Other Problems) had the poorest correlation with the composite score within the Vancouver sample (*r* = -.28), while the correlation for the Gorge sample was strong (*r* = .52). Thus, despite the poor correlations for some of the items within the whole sample, all of the items were retained.

The reader is also reminded of the correlations between the subscale scores and the composite outcome score previously reported as a portion of the exploratory investigations. As Table 9 shows, all of the subscales were found to have a strong correlation with the composite outcome score.

To further investigate the internal consistency of the FOP outcome scores, Cronbach's alpha was calculated using the full sample first administration data for each of the subscales' and the composite outcome's scores. The values ranged from .29 to .94 with 2 of the subscales and the composite outcome score having alpha values above the desired value of .80 (Table 14). A further 3 subscales had alpha values greater than .70 but less than the .80 desired. Together, these results speak to the internal reliability of the FOP, especially at the subscale score and composite outcome score level.

Finally, Cronbach's alpha was recalculated using the "problem only" data for the full first administration sample (*n* = 107). Unfortunately, because SPSS removes cases listwise by default when calculating alpha, reliability could not be determined for the composite outcome score because every survivor had at least one area that was not a problem (i.e., removed from

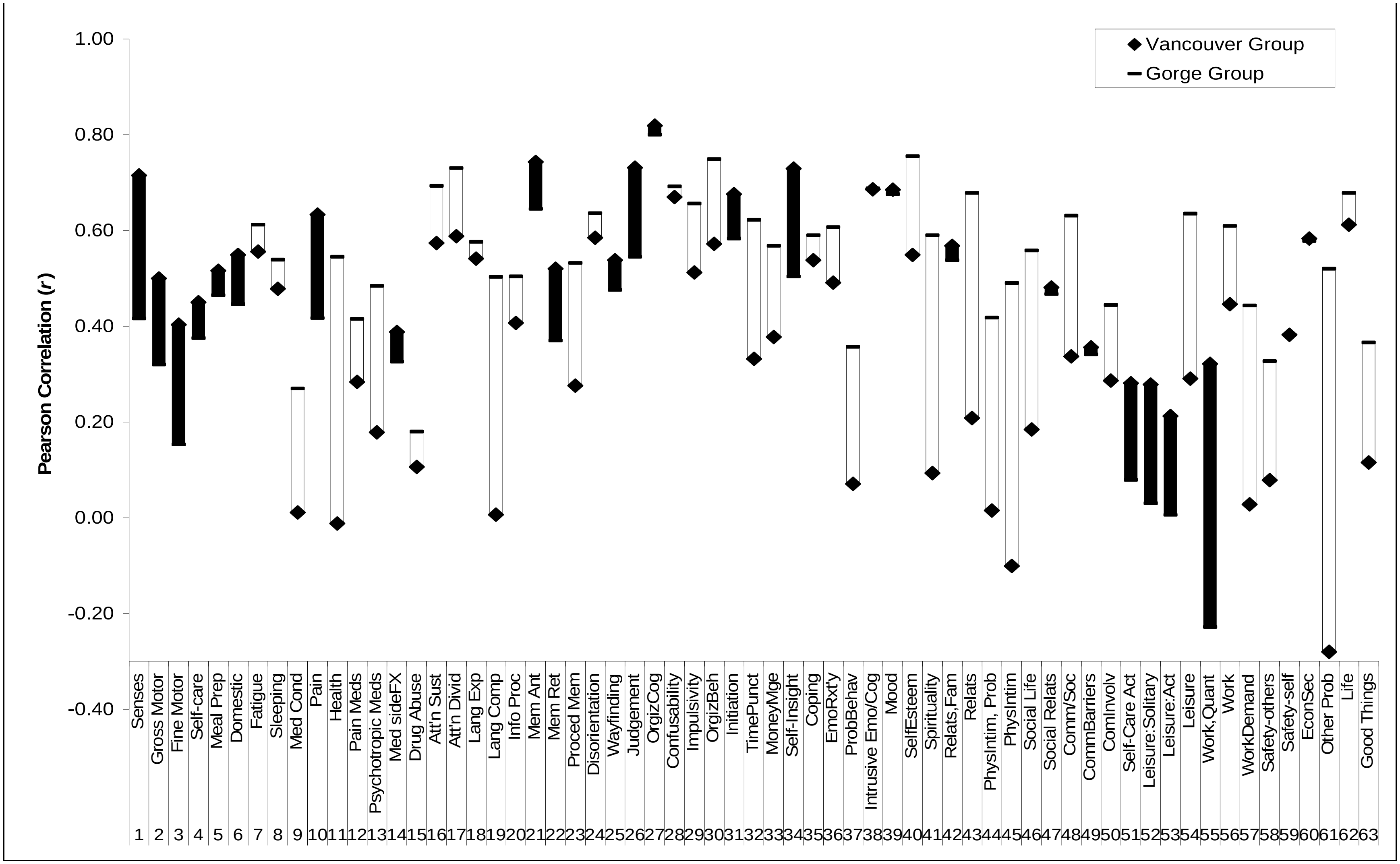


Figure 3. Item-Composite outcome score correlations for the Vancouver group (n = 27) and the Gorge group (n = 80). Black bars indicate that the Vancouver group had a higher correlation between the given item and the composite outcome score; white bars indicate that the Gorge group had a higher correlation between the given item and the composite outcome score.

Table 14.

Cronbach's Alpha calculated for the FOP composite outcome score and the 8 subscales

Scale	Cronbach's Alpha	Number of items
Composite outcome score	0.94	63
Physical, ADL	0.76	8
Health	0.64	7
Cognitive	0.87	10
Executive	0.87	10
Emotional/Behavioural	0.77	6
Social	0.74	9
Activities	0.29	9
Overall	0.61	4

the analysis). Thus, listwise deletion left no data that could be used to calculated alpha. However, 6 of the 8 subscales could have alpha calculated using the "problem only" data (Table 15). For these subscale scores, alpha ranged from .058 to .877 with 3 of the subscales having alpha values greater than the desired value of .80, and 1 more having alpha > .70; the remaining 2 subscales had alpha values < .70.

Table 15.

Cronbach's α calculated for the FOP subscales using the "problem only" data.

Scale	Cronbach's α	Number of items	Number of valid participants*
Physical, ADL	0.755	8	13
Health	-	7	0
Cognitive	0.877	10	7
Executive	0.852	10	5
Emotional/Behavioural	0.827	6	12
Social	-	9	1
Activities	0.058	9	18
Overall	0.651	4	32

* Valid Participants are those who had all items in the subscale rated as problematic thus enabling their scores to be included in the calculation. This data was based on the original full sample first administration data ($N = 107$).

The second reliability investigation was of the test-retest reliability of the outcome scores, which was investigated by correlating the time 1 and time 2 outcome scores for the short-interval group (i.e., 1 week interval). Pearson correlations were calculated for each item, each subscale and the composite outcome score (Table 16). The composite outcome score showed excellent test-retest reliability ($r = .97$), as did 5 of the 8 subscales ($r > .80$). The remaining 3 subscales had correlations between .70 and .79. The correlations for the individual items ranged from .24 to 1.00 with 30 of the 63 items having strong test-retest correlations ($r > .80$). A further 28 items had moderate correlations ($.50 < r < .79$) and only 3 items had weak

Table 16.

Test-retest correlations for short interval group (n = 18)

Scale/Item		Pearson Correlation
Composite Outcome scale		0.97
Physical, ADL		0.89
Health		0.93
Cognitive		0.90
Executive		0.92
Emotional/Behavioural		0.90
Social		0.75
Activities		0.70
Overall		0.76
1	Senses	0.97
2	Gross Motor	0.96
3	Fine Motor	0.85
4	Self-care	1.00
5	Eating & Meal Preparation	0.81
6	Domestic Skills	0.65
7	Fatigue	0.67
8	Sleeping	0.87
9	Medical Conditions	0.81
10	Pain	0.69
11	Health Satisfaction	0.65
12	Pain Medications	0.96
13	Psychotropic Medications	0.97
14	Medication Side Effects	0.82
15	Drug Abuse	1.00
16	Sustained Attention	0.59
17	Divided Attention	0.81
18	Language Expression	0.72
19	Language Comprehension	0.63
20	Information Processing	0.72
21	Anterograde Memory	0.82
22	Retrograde Memory	0.98
23	Procedural Memory	0.72
24	Time Disorientation	0.79
25	Wayfinding	0.68
26	Judgement	0.90
27	Cognitive Organization	0.68
28	Confusability	0.67
29	Impulsivity	0.79
30	Behavioural Organization	0.53
31	Initiation	0.84

32	Time & Punctuality	0.91
33	Money Management	0.99
34	Self-Insight	0.73
35	Coping	0.72
36	Emotional Reactivity	0.91
37	Problem Behaviour	0.56
38	Emotional/Cognitive Intrusions	0.67
39	Mood	0.77
40	Self Esteem	0.88
41	Spirituality	0.90
42	Family Relationships	0.94
43	Relationship Satisfaction	0.79
44	Physical Intimacy	1.00
45	Physical Intimacy Satisfaction	0.61
46	Social Life	0.24
47	Social Relationships	0.68
48	Community/Social Interactions	0.82
49	Community Barriers*	-
50	Community Involvement	0.42
51	Self-Care Activities	0.26
52	Leisure: Solitary	0.70
53	Leisure: Active/Social	0.93
54	Leisure Satisfaction	0.65
55	Work, Quantity	0.86
56	Work Satisfaction	0.75
57	Work Demands	0.84
58	Safety Concerns by Others	0.59
59	Safety Concerns by Self	0.82
60	Economic Security	0.83
61	Other Problems*	-
62	Life Satisfaction	0.84
63	Good Things	0.59

Note. All *ps* < 0.0001

* correlation could not be calculated due to a lack of variance

correlations (*r* < .49). Correlations for two of the items could not be calculated due to an absence of variability in the data for time 1 and/or time 2.

The test-retest correlations were then re-run using the “problem only” data for the 18 participants in the short interval group (Table 17). The composite outcome score still showed excellent test-retest reliability (*r* = .89) as did the scores for 5 of the 8 subscales (*r* > .80). Only 1 subscale score had unacceptable test-retest reliability (Activities subscale, *r* = .48). The correlations for the individual item scores ranged from -.50 to 1.00 with 21 of the 57 items for which correlations could be calculated having excellent test-retest reliability (*r* > .80). Correlations could not be calculated for the scores of 6 items because of either a lack of variability or too few participants (i.e., 1 or less).

Finally, the survivor-proxy agreement of the frequency ratings was investigated. First, Spearman correlations were calculated between the item frequency ratings provided by the survivors and their significant others. Spearman correlations were chosen based on the ordinal scale of measurement of the frequency ratings. The correlations ranged from -0.03 to 0.94 (see Table 18). Of the 49 items for which frequency ratings are made, only 3 had strong correlations (i.e., greater than 0.80) and 18 items had moderate correlations (i.e., between 0.50 and 0.79). The remaining 28 items had weak correlations (i.e., below 0.50), and 2 of these items had weakly negative correlations. Second, the frequency ratings provided by the survivors and significant others were re-coded to produce a 3-point scale and the associations were investigated via the calculation of Cramer’s V. The re-coding of the data was done so that original frequency ratings of 0 remained a score of 0 indicating “no problem”, original frequency ratings of 1.0 remained 1 indicating “always a problem”, and any rating in between (from 0.1 to 0.99) was re-coded as 0.5 indicating “sometimes a problem”. The Cramer’s V

Table 17.

Test-retest correlations for short interval group (n = 18) using "problem only" data

Scale/Item		Pearson Correlation	Significance	# of participants in analysis
Composite Outcome scale		0.893	< 0.001	18
Physical, ADL		0.833	< 0.001	16
Health		0.930	< 0.001	16
Cognitive		0.800	< 0.001	18
Executive		0.873	< 0.001	18
Emotional/Behavioural		0.950	< 0.001	17
Social		0.702	0.001	18
Activities		0.482	0.043	18
Overall		0.741	< 0.001	18
1	Senses	0.680	0.063	8
2	Gross Motor	0.944	0.005	6
3	Fine Motor	0.820	0.046	6
4	Self-care	1.000		2
5	Eating & Meal Preparation	-0.500	0.667	3
6	Domestic Skills	0.132	0.916	3
7	Fatigue	0.708	0.005	14
8	Sleeping	0.387	0.214	12
9	Medical Conditions	0.680	0.093	7
10	Pain	0.562	0.057	12
11	Health Satisfaction	0.640	0.008	16
12	Pain Medications	0.841	0.009	8
13	Psychotropic Medications	-0.200	0.704	6
14	Medication Side Effects	0.990	< 0.001	6
15	Drug Abuse	-	-	1
16	Sustained Attention	0.213	0.613	8
17	Divided Attention	0.788	0.001	14
18	Language Expression	0.592	0.161	7
19	Language Comprehension	0.752	0.051	7
20	Information Processing	-0.091	0.829	8
21	Anterograde Memory	0.615	0.015	15
22	Retrograde Memory	0.744	0.055	7
23	Procedural Memory	0.313	0.378	10
24	Time Disorientation	-0.201	0.799	4
25	Wayfinding	0.996	0.055	3
26	Judgement	0.880	0.002	9
27	Cognitive Organization	0.239	0.454	12
28	Confusability	0.803	0.003	11
29	Impulsivity	0.884	0.004	8
30	Behavioural Organization	0.712	0.048	8
31	Initiation	0.541	0.346	5
32	Time & Punctuality	-	-	2

33	Money Management	1.000	-	2
34	Self-Insight	0.379	0.250	11
35	Coping	0.862	0.013	7
36	Emotional Reactivity	0.938	< 0.001	11
37	Problem Behaviour	-	-	2
38	Emotional/Cognitive Intrusions	0.778	0.023	8
39	Mood	0.792	0.004	11
40	Self Esteem	0.890	< 0.001	16
41	Spirituality	0.959	< 0.001	11
42	Family Relationships	0.837	0.003	10
43	Relationship Satisfaction	0.832	< 0.001	17
44	Physical Intimacy	-	-	1
45	Physical Intimacy Satisfaction	0.633	0.006	17
46	Social Life	0.213	0.412	17
47	Social Relationships	0.390	0.444	6
48	Community/Social Interactions	1.000		2
49	Community Barriers	-	-	0
50	Community Involvement	0.771	0.005	11
51	Self-Care Activities	0.363	0.152	17
52	Leisure: Solitary	0.718	0.001	17
53	Leisure: Active/Social	0.913	< 0.001	17
54	Leisure Satisfaction	0.658	0.004	17
55	Work, Quantity	0.844	< 0.001	17
56	Work Satisfaction	0.776	< 0.001	17
57	Work Demands	0.500	0.667	3
58	Safety Concerns by Others	0.327	0.788	3
59	Safety Concerns by Self	0.000	1.000	4
60	Economic Security	0.837	< 0.001	17
61	Other Problems	-	-	0
62	Life Satisfaction	0.829	< 0.001	17
63	Good Things	0.583	0.011	18

Table 18.
Survivor-proxy agreement on problem frequency ratings

Item		Spearman correlation	Significance
1	Senses	0.48	0.031
2	Gross Motor	0.72	< 0.001
3	Fine Motor	0.70	< 0.001
4	Self-Care	0.77	< 0.001
5	Eating & Meal Preparation	0.80	< 0.001
6	Domestic Skills	0.31	0.188
7	Fatigue	0.60	0.005
8	Sleeping	0.36	0.117
9	Medical Conditions	0.61	0.004
10	Pain	0.75	< 0.001
12	Pain Medications	0.94	< 0.001
13	Psychotropic Medications	0.55	0.013
14	Medication Side Effects	0.40	0.084
15	Drug Abuse*	-	-
16	Sustained Attention	0.55	0.012
17	Divided Attention	0.31	0.185
18	Language Expression	0.83	< 0.001
19	Language Comprehension	0.71	< 0.001
20	Information Processing Speed	0.26	0.260
21	Anterograde Memory	0.43	0.059
22	Retrograde Memory	0.19	0.433
23	Procedural Learning	0.08	0.753
24	Time Disorientation	0.36	0.118
25	Wayfinding	0.70	0.001
26	Judgment	0.36	0.116
27	Cognitive Organization	0.41	0.070
28	Confusability	0.59	0.006
29	Impulsivity	0.47	0.037
30	Behavioural Organization	0.24	0.308
31	Initiation	0.65	0.002
32	Time & Punctuality	0.67	0.001
33	Money Management	0.34	0.138
34	Self-Insight	0.49	0.030
35	Coping	0.49	0.028
36	Emotional Reactivity	0.42	0.063
37	Problem Behaviour	0.70	0.001
38	Emotional/Cognitive Intrusions	0.49	0.029
39	Mood	0.55	0.012
42	Family Relationships	-0.02	0.946
44	Physical Intimacy	0.72	< 0.001
47	Social Relationships	0.24	0.303
48	Community/Social Interactions	0.63	0.003
49	Community Barriers	0.73	< 0.001

57	Work Demands	0.14	0.570
58	Safety Concerns by Others	-0.03	0.907
59	Safety Concerns by Self	0.32	0.167
60	Economic Security	0.32	0.173
61	Other Problems	0.20	0.390
63	Good Things	0.25	0.296

** correlation could not be calculated due to an absence of variability* associations for the re-coded frequency ratings ranged from 0.126 to 0.899 (Table 19). Associations could not be calculated for 2 of the items. Survivors' ratings on Item #15 were a constant (all rated as "not a problem") and significant others' ratings on Item #63 were a constant (all rated as "sometimes a problem"). Of the frequency items that associations were calculated for (47 of the 49 items), 19 were found to be significant. Finally, the percent agreement was calculated based on survivors and significant others both rating an item as not a problem (i.e., a frequency rating of 0) or as a problem (i.e., a frequency rating of 0.1 or greater) (Table 20). Agreement ranged from 45% to 95% with 40 of the 49 items being agreed upon by at least 70% of the survivor-significant other pairs.

All together, the results show that the FOP outcome scores have good internal consistency and test-retest reliability, but the frequency ratings have questionable survivor-proxy agreement.

Validity

Evidence based on test content. Multiple lines of evidence were collected for the validity of the FOP based on test content. First, the development of the FOP (as outlined in the introduction) followed the steps outlined in the *Standards for educational and psychological testing* for test development (AERA, APA & NCME, 1999). For example, the FOP was developed based on existing literature, knowledge of other outcome measures and professional expertise, and was revised based on feedback from professionals and brain injury survivors. These test development procedures are well accepted and lay the groundwork for a measure that has content which is relevant to the test objectives and the population of interest.

Table 19.

Re-coded Survivor-proxy agreement on problem frequency ratings

Item		Cramer's V	Significance
1	Senses	0.433	0.112
2	Gross Motor	0.637	0.003
3	Fine Motor	0.508	0.035
4	Self-Care	0.858	< 0.001
5	Eating & Meal Preparation	0.561	0.014
6	Domestic Skills	0.382	0.233
7	Fatigue	0.498	0.042
8	Sleeping	0.467	0.113
9	Medical Conditions	0.630	0.003
10	Pain	0.585	0.008
12	Pain Medications	0.899	< 0.001
13	Psychotropic Medications	0.545	0.015
14	Medication Side Effects	0.509	0.075
15	Drug Abuse*	-	-
16	Sustained Attention	0.538	0.016
17	Divided Attention	0.510	0.034
18	Language Expression	0.605	0.006
19	Language Comprehension	0.518	0.068
20	Information Processing Speed	0.415	0.179
21	Anterograde Memory	0.392	0.188
22	Retrograde Memory	0.126	0.853
23	Procedural Learning	0.154	0.492
24	Time Disorientation	0.425	0.165
25	Wayfinding	0.814	< 0.001
26	Judgment	0.244	0.552
27	Cognitive Organization	0.488	0.092
28	Confusability	0.391	0.216
29	Impulsivity	0.343	0.319
30	Behavioural Organization	0.298	0.468
31	Initiation	0.451	0.087
32	Time & Punctuality	0.612	0.005
33	Money Management	0.355	0.283
34	Self-Insight	0.372	0.236
35	Coping	0.504	0.038
36	Emotional Reactivity	0.241	0.678
37	Problem Behaviour	0.814	< 0.001
38	Emotional/Cognitive Intrusions	0.354	0.285
39	Mood	0.443	0.140
42	Family Relationships	0.134	0.949
44	Physical Intimacy	0.728	0.001
47	Social Relationships	0.358	0.275
48	Community/Social Interactions	0.639	0.017

49	Community Barriers	0.688	0.002
57	Work Demands	0.140	0.531
58	Safety Concerns by Others	0.392	0.216
59	Safety Concerns by Self	0.415	0.179
60	Economic Security	0.446	0.094
61	Other Problems	0.330	0.337
63	Good Things**	-	-

* association could not be calculated due to a constant survivor rating

** association could not be calculated due to a constant SO rating

Table 20.

Survivor-proxy percent agreement on dichotomized problem frequency ratings

Item		Percent agreement
1	Senses	75
2	Gross Motor	85
3	Fine Motor	75
4	Self-Care	95
5	Eating & Meal Preparation	85
6	Domestic Skills	75
7	Fatigue	95
8	Sleeping	70
9	Medical Cond	80
10	Pain	80
12	Pain Medications	95
13	Psychotropic Medications	80
14	Medication Side Effects	75
15	Drug Abuse	95
16	Sustained Attention	70
17	Divided Attention	85
18	Language Expression	85
19	Language Comprehension	75
20	Information Processing Speed	70
21	Anterograde Memory	85
22	Retrograde Memory	65
23	Procedural Learning	55
24	Time Disorientation	70
25	Wayfinding	85
26	Judgment	60
27	Cognitive Organization	75
28	Confusability	75
29	Impulsivity	60
30	Behavioural Organization	55
31	Initiation	75
32	Time & Punctuality	85
33	Money Management	70
34	Self-Insight	75
35	Coping	85
36	Emotional Reactivity	65
37	Problem Behaviour	85
38	Emotional/Cognitive Intrusions	75
39	Mood	70
42	Family Relationships	55
44	Physical Intimacy	90
47	Social Relationships	75
48	Community/Social Interactions	85
49	Community Barriers	95

57	Work Demands	75
58	Safety Concerns by Others	45
59	Safety Concerns by Self	75
60	Economic Security	60
61	Other Problems	80
63	Good Things	95

Second, more objective evidence for the validity of the FOP based on test content was found by probing all 63 items for utility. This probe revealed that problems were reported in all 63 of the functional areas by at least one survivor ($N = 107$) and at least one significant other ($N = 20$).

Third, a survey of FOP item 61 (Other problems) revealed that that 73% (78 of the 107) survivors and 90% (18 of the 20) significant others reported no “other problems”; closer analysis of the other problems reported showed many of these to be specific instances of general categories already covered by FOP items. In other words, for most survivors and significant others all of the problems currently being experienced were included in the FOP. In addition, Figure 4 and Figure 5 illustrate the frequency ratings provided by survivors and significant others respectively for item 61, revealing that those who were experiencing other problems were experiencing them relatively infrequently. Together, these lines of evidence show that the FOP is both relevant to survivors and significant others and sufficiently broad to include almost all problems experienced by brain injury survivors and their significant others.

Evidence based on external variables. As previously mentioned, many person- and injury-characteristics have been consistently shown to be related to outcome following ABI. Therefore, the relationship between 5 specific demographic variables and the FOP subscale and composite outcome scores were investigated using a MANOVA. All post-hoc pairwise comparisons were done using Dunnett’s test as the assumption of homogeneity of variances was violated (all $ps < 0.005$). The demographic variables included in the model were: gender (male or female), etiology (TBI, ABI, or other), injury severity (mild, moderate or severe as based on length of PTA reported), age group at time of injury (under 25 years, 26-50 years, or over 50

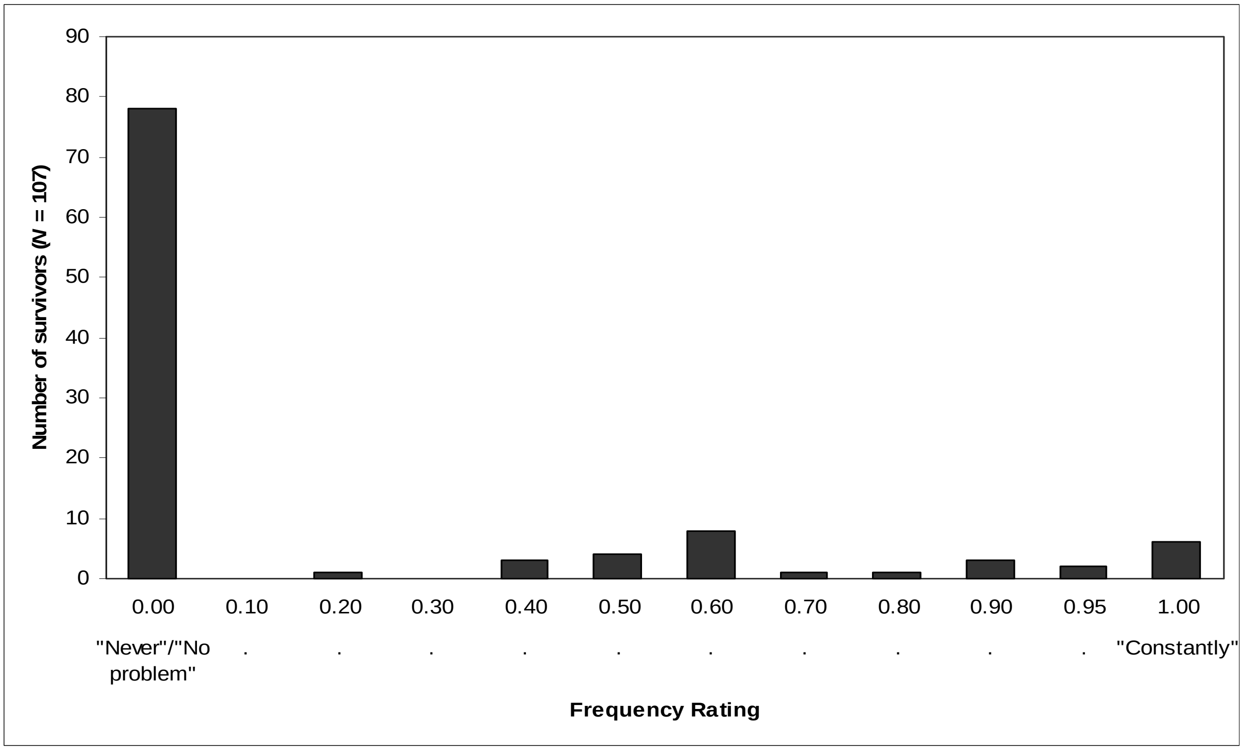


*Figure 4***.** Number of survivors reporting each frequency rating for Item #61 (Other Problems).

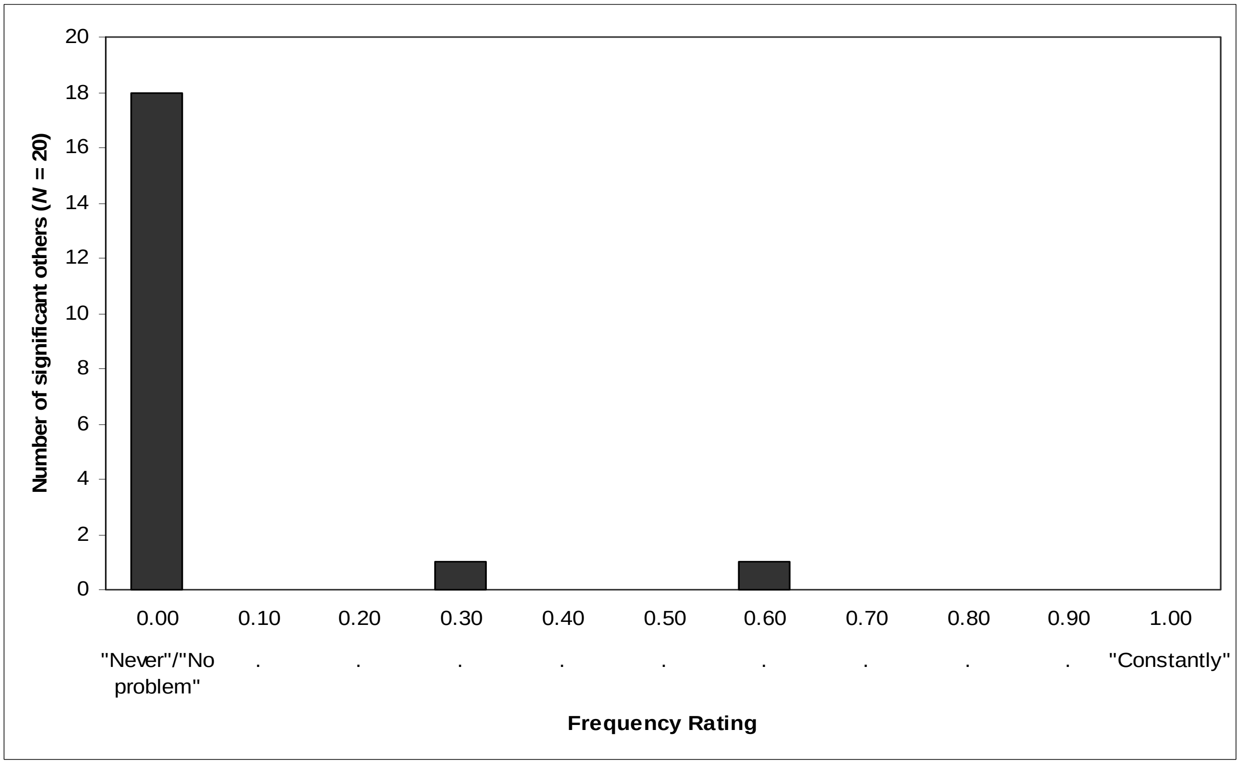


Figure 5. Number of significant others reporting each frequency rating for Item #61 (Other Problems).

years), and length of time post-injury (under 6 months, 6-12 months, 13-24 months, 25-48 months, or over 48 months).

When the effects of the demographic variables on the composite outcome score were investigated, significant results were found for the main effects of age at injury ($F = 5.60$, $p = 0.007$, $\eta^2 = .223$) and etiology ($F = 3.921$, $p = 0.028$, $\eta^2 = .167$). When the effects of the demographic variables on the subscale scores were investigated, the following main effects were found. Length of PTA was found to affect the Physical/ADL subscale ($F = 4.836$, $p = 0.013$, $\eta^2 = 0.199$) and the Executive subscale ($F = 3.494$, $p = 0.040$, $\eta^2 = 0.152$). Etiology was found to affect the Cognition Subscale ($F = 4.084$, $p = 0.025$, $\eta^2 = 0.173$), the Executive subscale ($F = 4.604$, $p = 0.016$, $\eta^2 = 0.191$), the Emotion subscale ($F = 3.770$, $p = 0.032$, $\eta^2 = 0.162$) and the Social subscale ($F = 6.673$, $p = 0.003$, $\eta^2 = 0.255$). Age at the time of injury was found to affect the Physical/ADL subscale ($F = 4.916$, $p = 0.012$, $\eta^2 = 0.201$), the Health subscale ($F = 5.064$, $p = 0.011$, $\eta^2 = 0.206$), the Cognition subscale ($F = 4.816$, $p = 0.014$, $\eta^2 = 0.198$), the Executive subscale ($F = 4.140$, $p = 0.023$, $\eta^2 = 0.175$), and the Emotion subscale ($F = 3.502$, $p = 0.040$, $\eta^2 = 0.152$). Finally, the length of time post-injury was found to affect only the Activity subscale ($F = 3.006$, $p =0.030$, $\eta^2 = 0.236$). In short, the demographic variables were found to have an influence on the outcomes scores obtained on the FOP.

I then further investigated the significant effects by computing pairwise comparisons. The results revealed that only comparisons of Etiology (TBI vs. ABI not including TBI) and Age at time of injury (25-50 years old vs. over 50 years old) were significant. While complete results can be found in Table 21, the comparisons consistently revealed that those with TBI had significantly lower outcome scores than those with ABI not including TBI and those who

Table 21.

Post-hoc pairwise comparisons using Dunnett`s: Descriptive statistics and significance of difference for outcome scores as a function of demographic variables

FOP Scale (DV)	Demographic Variable				Significance of Difference	Std. Error
	Name	Level	Mean	Std. Dev		
Composite Outcome Score	Etiology	TBI	0.624	0.151	$p < 0.001$	0.027
		ABI not incl TBI	0.779	0.111		
	Age at time of injury	25 – 50 years	0.650	0.166	$p = 0.001$	0.029
		Over 50 years	0.770	0.103		
Physical/ADL Subscale	Age at time of injury	25 – 50 years	0.576	0.234	$p = 0.010$	0.044
		Over 50 years	0.719	0.171		
Health Subscale	Age at time of injury	25 – 50 years	0.635	0.209	$p = 0.005$	0.041
		Over 50 years	0.779	0.154		
Cognitive Subscale	Etiology	TBI	0.569	0.211	$p < 0.001$	0.046
		ABI not incl TBI	0.768	0.264		
	Age at time of injury	25 – 50 years	0.597	0.244	$p = 0.005$	0.051
		Over 50 years	0.748	0.207		
Executive Subscale	Etiology	TBI	0.569	0.243	$p < 0.001$	0.039
		ABI not incl TBI	0.831	0.141		
	Age at time of injury	25 – 50 years	0.636	0.258	$p = 0.001$	0.045
		Over 50 years	0.808	0.157		
Emotion Subscale	Etiology	TBI	0.615	0.218	$p < 0.001$	0.040
		ABI not incl TBI	0.825	0.170		
	Age at time of injury	25 – 50 years	0.649	0.231	$p < 0.001$	0.043
		Over 50 years	0.823	0.158		
Social Subscale	Etiology	TBI	0.735	0.167	$p < 0.001$	0.027
		ABI not incl TBI	0.871	0.096		

sustained their injuries between 25 and 50 years of age had significantly lower outcome scores than those who were over 50 years of age when they sustained their injury.

Together, these results indicate that the composite outcome score and subscale scores of the FOP are influenced by demographic variables that have been consistently shown to influence outcome following brain injury. The effect of the demographic variables on the FOP outcome scores provides evidence that the FOP is measuring outcome as it was intended to do.

Change over time. To investigate the ability of the FOP to detect change over time, a 2 x 2 x 9 Repeated Measures ANOVA was used. The variables of interest were Interval (short vs. long), Time (first vs. second administration) and Scale (the 8 sub-scale scores plus the composite outcome score[1]). Of primary interest was a Time x Interval interaction, which was found to be significant ($F = 12.42$, $p < 0.001$, $R^2 = .286$). Therefore, a difference between time 1 and time 2 scores (i.e., change over time) is dependent on interval group, indicating that the FOP is sensitive to change over time. Other significant, but less relevant, results included a main effect for Time ($F = 13.85$, $p < 0.001$, $R^2 = .309$), a main effect for Scales ($F = 2.66$, $p < 0.030$, $R^2 = .470$), as well as a Time x Scales interaction ($F = 3.60$, $p < .005$, $R^2 = .545$) which indicates that some of the scales changed more from Time 1 to Time 2 across (or independent of) the two interval groups.

The SDD was calculated for the 8 subscales and the composite outcome score as another determinant of change over time. It was found that the short interval group had no subscale differences between time 1 and time 2 greater than the SDD (Table 22). On the other hand, in

---

[1] To ensure that including the composite outcome score along with the other 8 subscales did not bias the results, a 2 x 2 x 8 Repeated Measures ANOVA was also run. Significant findings remained the same with or without the inclusion of the composite outcome score, although power was slightly higher when the composite outcome score was not included.

Table 22.

Smallest Detectable Difference (SDD) and mean T1-T2 differences for the short and long interval groups for the subscales and composite outcome score.

Scale	SDD	Short interval mean difference	Long interval mean difference
ADL	0.100	0.013	0.093
Health	0.097	-0.007	**0.100**
Cognition	0.116	0.020	**0.127**
Executive	0.108	-0.007	0.080
Emotional	0.089	-0.020	0.040
Social	0.069	-0.013	0.020
Activities	0.044	0.013	**0.080**
Overall	0.067	0.013	**0.100**
OUTCOME	0.067	0.001	**0.073**

Bold indicates differences greater than the SDD

the long interval group 4 subscales and the composite outcome score had mean differences larger than the SDD (as can be seen in Figure 5).

The final analysis was an investigation of change when change was possible, where only data points indicating problematic functioning at Time 1 and their corresponding Time 2 scores were included (i.e., scores that were not 1.0 at Time 1 and 2). The Independent samples t-test was run investigating the effect of interval group on the mean difference scores for the subscales and composite outcome. Interestingly, the only significant difference between the short and long interval groups was found on the composite outcome score ($t(31) = 3.33$, $p = .002$, equal variances assumed), although a near significant result was found for the Health subscale ($t(31) = 2.00$, $p = .059$, equal variances not assumed).

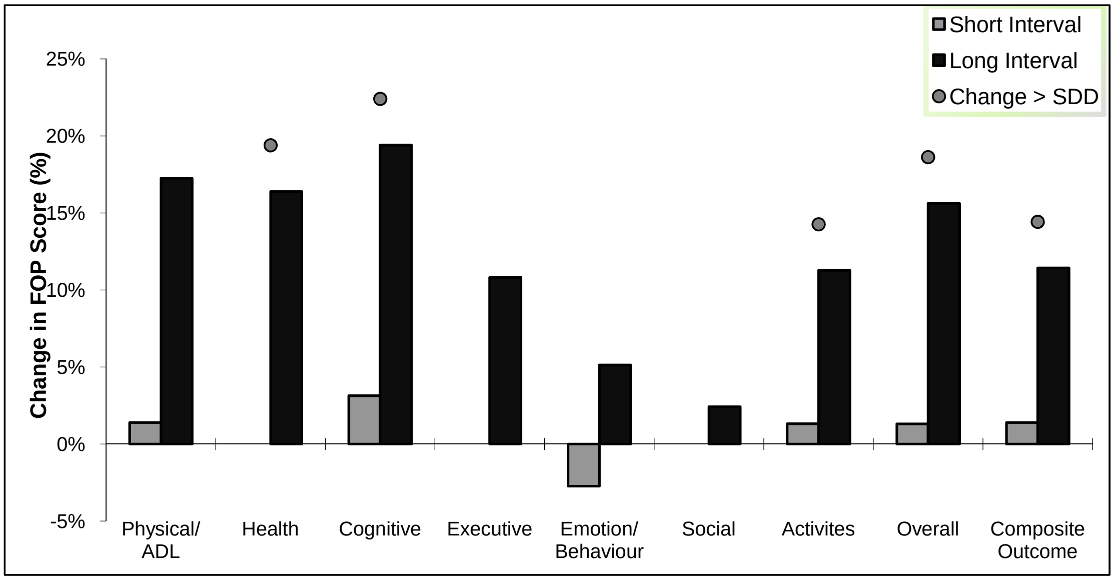


Figure 6. Percent change in FOP outcome scores for the short and long interval groups separately, with those changes larger than the SDD indicated.

Chapter Four: Discussion

The main objective of the present study was to investigate the FOP's appropriateness as a measure of outcome following brain injury. Through both traditional and more unconventional approaches, the psychometric characteristics of aspects of the FOP were investigated, and multiple lines of evidence for the validity of the FOP were put forth including its ability to detect change over time. Overall, the results indicate that the FOP is a psychometrically sound measure that will be useful for the investigation of functional outcome after brain injury from the perspective of survivors and their significant others.

Exploratory Investigations

Although the FOP had previously undergone some psychometric investigation, the underlying assumptions incorporated in the development of the FOP were never probed. Therefore, this task was undertaken first.

The investigation of the distribution of the items, subscales, and composite outcome scores revealed that some items and one subscale had noteworthy negative skews. This finding was not surprising given the common practice of giving a score of 1.0 (optimal functioning) for items rated as “not a problem”. As such, the results of the investigation of the “problem only” data, which included only items that the survivor rated as problematic, revealed that for a majority of the items the initial skew was related to the number of participants who had rated the items as not problematic and were thus automatically given scores of 1.0. This common practice of the FOP leads to a large number of scores being above the mid-point (and specifically, at the ceiling) which resulted in many of the items having data that were negatively skewed. However, because the rating scales and outcome scores are not arbitrary (i.e., the end-points do have meaning) transformation of the data was not an option. Therefore, whenever possible or applicable, the subsequent analyses were run with both the full data and the problem only data to ensure that the impact of the negative skew on the results was clear.

The scoring of the items was investigated next. The three types of questions included in the FOP require the survivor or significant other to rate different aspects of functioning. The Impact questions require the survivor or significant other to rate the problem frequency as well as the impact of the problem, despite the fact that these frequency ratings are not included in the calculation of the Impact scores. During the development of the equation for the Impact questions, the thought was that the rater (survivor or significant other) would take into consideration the frequency of the problem when providing the impact rating (R. W. Skelton, personal communication, September 18, 2008). In addition to this logic, it appears that when the frequency ratings are included in the formula, especially when problem frequency is high, the score is a reflection of the individual's deficit as opposed to their functioning. In other words, simply because a problem is frequent does not necessarily imply that there is a deterioration of functioning. The impact rating, which asks the individual to rate how much the problem impacts his/her daily life, is a more accurate measure of the individual's functioning, such that infrequent high-impact problems interfere with functioning more than frequent low-impact problems. Therefore, the formula for calculating the Impact scores was not changed.

The satisfaction questions also include two ratings to be made by the survivor or significant other: a rating of the individual's satisfaction with the domain in question and a rating of the importance of the domain to the individual. The original calculation of the Satisfaction questions includes both of these ratings such that the level of dissatisfaction with an issue is moderated by the importance of the issue to the individual. For example, if an individual is unsatisfied with his/her relationships with family and these relationships are very important to him/her, a low score will be obtained for this item which corresponds to the poor functioning in this area. On the other hand, if another individual is equally dissatisfied with his/her relationships with family, but these relationships are unimportant to him/her (due, perhaps, to a close circle of friends) a higher score will be obtained corresponding to the 'neutral' functioning in this area. Since it was found that the original Satisfaction formula reflected this relationship well, it was retained.

The last question type inquires about different activities and the level of independence with which these activities are completed. As mentioned, the logic underlying the formula seems reasonable and a more appropriate formula that included all of the aspects included in the original could not be created. Therefore, the formula was retained. Nevertheless, it is important to keep in mind that the calculation of these scores does require both an accurate estimate of how many hours per day are spent involved in a number of different activities, as well as the amount of help that is required to complete each activity. Common deficits after brain injury, such as memory impairments or a lack of awareness, could negatively impact these estimates. However, this must be considered in light of a survivor's overall ability to provide relatively consistent ratings on the rest of the items.

The final exploratory investigation was aimed at determining the appropriateness of the 8 subscales. The method used to determine which group of subscales was the "best" (the 8 original subscales, the 3 question-type subscales, or the 18 construct-based subscales) involved computing Pearson correlations between items and their representative subscale score, and between each of subscale score and the composite outcome score. This method has obvious limitations, the main one being a bias related to the number of items included in each subscale.

When correlating each item with its representative subscale, a subscale score that is composed of only 2 items is more likely to have strong or moderate correlations with each of its items than subscales containing of a larger number of items (e.g., 10 or more). This is because each subscale score is an average of the outcome scores for the items included the subscale with each item given equal weight. Thus each item in a subscale made up of only 2 items encompasses much more of the subscale score than each item in a subscale made up of 10 items.

The opposite is true when considering the correlations between the subscale scores and the composite outcome score. Because each subscale reflects the average of a group of items and the composite outcome score reflects the average of all of the items, the more items a subscale is composed of the more composite outcome score it accounts for. Therefore, it is more likely that subscales consisting of a large number of items will have a strong correlation with the composite outcome score. In both analyses, the correlations in question are looking at the relationship of smaller parts of the FOP with larger parts of the FOP, the correlations may be regarded as being of little value, or at least providing little unbiased information. A consideration was made to compare the alpha coefficients for each subscale, but a similar problem would have arisen: because alpha is strongly influenced by the number of items used in the calculation (Osterlind, 2006), the scales with more items would inevitably have a higher alpha. Overall, the possible changes considered did not improve the FOP above the original aspects of the measure, so no changes were made.

Reliability

This study investigated three standard types of reliability. The first was internal consistency which was investigated through both correlational analyses as well as Cronbach's alpha – a more traditional approach to internal consistency. The results of the correlations revealed that, for the full sample, a number of items were found to have little to no relationship to the composite outcome score. This result was somewhat surprising and for some may be grounds for removal of the items (an issue that will be discussed in more detail later). However, in the current study these results were believed to be related (in part) to the heterogeneity of the population of brain injury survivors in general, such that the symptom constellation experienced is individual to each survivor. This theory was supported by the finding that the item-composite correlations for the Gorge and Vancouver samples, when looked at separately, were quite different from each other and from the results for the full sample. This theory is also supported by the general finding that every survivor in the sample reported at least 1 area as not problematic as reflected in the ability to calculate an alpha value for the composite outcome measure using the “problem only” data. Therefore, it is believed that these findings do not necessarily imply that the FOP has poor internal consistency, but that item-composite correlations are an inappropriate way to investigate this type of reliability. As such, the values for Cronbach’s alpha (based on both the full and the “problem only” data) were generally acceptable although some may have been inflated when based on full first administration data.

The test-retest reliability analyses were completed next. As expected, the FOP outcome scores showed excellent test-retest stability especially at the subscale and composite outcome score levels. At the item level, however, a number of items were found to have very poor and even negative correlations when based on the “problem only” data, indicating that those with problems in specific domains provide less reliable accounts of the impact and importance of, and their satisfaction with, the domain in question. Nevertheless, these results should be interpreted with caution as many of the correlations were based on very few data points. Three of the correlations are based on only 2 pairs of scores, and a further 5 correlations are based on only 3 pairs of scores. Overall however, the results indicate that the FOP is not subject to significant practice effects (nor was it expected to be, given what it is intended to measure) and it has been shown to be a relatively stable measure over a short time interval.

The final reliability analyses looked at survivor-proxy agreement of the frequency ratings of the FOP and FOP SO. The level of agreement on a number of items was found to be disappointing when the full 11-point scale was utilized. Interestingly, when the ratings were re-coded on a 3-point scale, little improvement in the agreement was found. However, after the scores were dichotomized to either agree or disagree, a satisfactory percent agreement was found. Therefore, survivors and significant others seem to agree if an area is or is not problematic, but are less able to agree on how often a problem occurs. This issue will be discussed in more detail later.

Validity

Three lines of evidence were presented for the validity of the FOP to be used as a measure of outcome following brain injury. The evidence, both theoretical and empirical, provides grounds upon which to deem the FOP a valid measure of outcome following brain injury. The evidence collected based on test content showed that the FOP items are relevant to the current sample of brain injury survivors and significant others; the items appear to be a valid reflection of the problems that these groups were experiencing. Also, the items appear to have sufficient breadth so that few problems were reported outside of those inquired about in the FOP as evidenced by very few participants reporting “other problems” on item #61. Of the “other problems” that were reported most were more specific examples of FOP items. For example, some of the other problems reported by survivors included the “inability to plan and adjust to changing situations”, “relationships with wife and son”, and a “lack of motivation”. Only two of 29 problems reported as “other problems” were not included in the FOP; these were “getting stressed out and frustrated both with other people and with myself” and “driving restrictions”. Thus, the evidence shows that the FOP is applicable to the population of community living survivors and has sufficient breadth to collect information about almost all domains experienced as problematic.

The validity evidence collected based on external variables was less straightforward. It was found that all of the injury- and demographic-characteristics investigated except Gender (i.e., Etiology, Injury severity, Age at time of injury, and Time since injury) were related to the FOP scale scores. Interestingly, gender is the variable that is much less frequently linked to outcome in the literature when compared to the other, more injury-related, variables. Because the FOP outcome scores were found to be related to these variables which have been previously shown to influence outcome, it is concluded that the FOP is measuring functional outcome.

The follow-up, pairwise comparisons of the effects of the demographic variables on FOP scores showed that only age at time of injury and type of injury had significant between group differences. In general, it was found that individuals who sustained a brain injury during young-middle adulthood (i.e., between the ages of 25 and 50) had the poorest functional outcomes, significantly poorer than those who were older than 50 when they sustained their injury. A similar trend was found by Stambrook, Moore, Lubusko, Peters, and Blumenschein (1993) when investigating the relationship between current age and psychosocial outcome. Their results showed that middle-age individuals (mean 41.5 years) reported worse mood disturbances and were rated as suffering from higher levels of psychopathology than both younger (mean 26.63 years) and older (mean 73.42 years) individuals. The present study also found that individuals who sustained traumatic brain injuries had the poorest functional outcomes, significantly poorer than those who sustained acquired brain injuries not including TBI. There seems to be no available literature comparing those with TBI to those with ABI in any areas of functioning or outcome after injury, and as such any speculation into why this was found would be speculative. The finding of a significant interaction between these variables, however, implies that the relationship between outcome and age is influenced by the type of injury sustained. It is also likely that those who were 25 to 50 years of age are more likely to have sustained a TBI while those over 50 are more likely to have sustained an ABI (e.g., a stroke). This relationship and its implications for recovery are definitely worth future investigation.

A key feature of the current study was the investigation of the FOP's sensitivity to change over time. Sensitivity to change over time is an important, and necessary, characteristic if the FOP will be used to evaluate rehabilitation programs or the progress made by individuals over the course of recovery. The results of these analyses met most of the guidelines put forth by Vermeersch and colleagues (2000) for the investigation and interpretation of sensitivity to change. The first guideline is that the change in scores should occur in the theoretically proposed direction. In this study, the change in FOP scores was in the theoretically proposed direction; the long-interval group showed an improvement in their functioning from Time 1 to Time 2. The second guideline is that the participants who received the intervention should show more change than the control group. This guideline is problematic for this study because there is no true control group. The short interval group was not receiving the intervention, had a different between-test interval, and was located in a different city than the long interval group. While these differences make the attribution of change impossible (e.g., did change occur because of the intervention, because of the time interval, both, or for some other unknown factor), it does not negate that change did occur for the long interval group and not for the short interval group. The third guideline is that the change should not be attributable to measurement error, practice effects, and other such sources of error. In this study the results of the test-retest reliability and the investigation of the SDD show that the FOP outcome scores are stable over short intervals and the change found is not attributable to issues such as practice effects or measurement error.

Overall, the results demonstrate that the FOP is able to detect change over time, even with a relatively short interval between administrations. The sample whose members had a 3-5 month interval between the two administrations of the FOP in conjunction with rehabilitation had appreciably better outcome scores at Time 2 compared to Time 1, while the group whose members were interviewed twice within only one week showed no significant change from Time 1 to Time 2 in all three analyses. Despite its significance, the composite outcome score's mean difference between Time 1 and Time 2 for the long interval group was relatively small (.08 for the full data, and .07 for the change possible data). However, it should be kept in mind that the 3-5 month interval between administrations is short when compared to other studies; for example, van Baalen and colleagues (2006) investigated change with a 1 year interval between test administrations. Another factor that may have influenced the size of the difference is the average time since injury; the long-interval group was an average of 6½ years post-injury, a point by which recovery is commonly thought to have reached a plateau

(Richardson, 2000). That the FOP was able to detect change despite these disadvantages speaks to its sensitivity.

Comparison to Previous Psychometric Investigation of the FOP

While many of the investigations completed in this study were unique (i.e., the investigations using “problem only” data), a number of the analyses were done in order to re-investigate the psychometric properties of the FOP following from the work done by Price (2007). However, even when the same approach was taken, a slightly different sample was used; some of the participants included in the earlier study were not included in the current study. As such, the results were expected to show minor differences across studies.

Three types of reliability that were previously looked at by Price (2007) - internal consistency, test-retest reliability and survivor-proxy agreement - were re-investigated. A comparison of the past and present results for Cronbach’s alpha used as an analysis of internal consistency revealed that six of the alpha values (5 subscales and the composite outcome score) found in the current study were identical to those found by Price; of those alpha values that differed, the differences were very small (.02 or less). Thus, both investigations found that the FOP composite outcome score and the individual subscales have very good internal reliability.

Similarly, when comparing the test-retest results, the two studies found very similar correlations between Time 1 and Time 2 outcome scores for 60 of the 63 items. However, 3 items had correlations that were notably greater in the previous study than in the current study. In addition, a comparison of the past and present survivor-proxy reliability analyses based on the Spearman correlations of the full 11-point frequency ratings revealed 3 items with notably divergent correlations: two of the items had been previously found to have very weak positive relationships and were presently found to have very weak negative relationships, and the third item was currently found to have a correlation much greater than that reported by Price (2007).

An explanation for all of these differences is not obvious. It is possible that the smaller differences (i.e., those only slightly greater than .10, including the two that were previously weakly positive but currently weakly negative) are due to the smaller sample size of the current study. The larger differences, on the other hand, may be indicative of the removal of "outliers" or perhaps the removal of some normative data which has allowed "outliers" to exert a stronger influence on the analyses. Despite these differences, the FOP was currently found to have good internal consistency and test-retest reliability, but questionable survivor-proxy agreement a conclusion similar to that arrived at by Price (2007).

Comparison to Existing Outcome Measures

The FOP was designed to be an inventory of possible problems following brain injury to provide clinicians or researchers with a profile of the subjective functioning of the survivor and his/her significant other after the injury. Because of this purpose and approach, the FOP is quite different from existing measures of brain injury outcome. First, the scope of the FOP is much broader than most existing measures with items inquiring about problems with gross motor functioning to items inquiring about satisfaction with and importance of the rater's spirituality and almost everything in between making it more appropriate for community-living survivors than many existing measures. Second, the FOP gathers information about not only problem frequency, like existing measures, but also about the impact of problems on the everyday functioning of the rater, the importance of the functional domains to the rater, and the satisfaction that the rater has in the domains. Finally, the FOP is not interested in deficits, but in functioning and the survivor's and significant other's ability to do the things that they need to do and what impact sequelae are having on this ability.

Despite these important differences, when compared to the psychometric analyses conducted with other measures of brain injury outcome, the FOP fares well. The FOP has been found to have similar internal consistency to other broad measures of outcome such as the BICRO-39 (Powell et al., 1998), the EBIQ (Teasdale et al., 1997) and the MPAI-4 (Malec et al., 2003). Similarly, the FOP subscales have been found to have comparable test-retest reliabilities to the subscales of existing outcome measures. However as none of the other measures have test-retest reliabilities reported for the items, there is no available data to compare these results to. The survivor-proxy agreement for the FOP has a much wider range than most of the existing outcome measures. Both weaker and stronger relationships than have been reported for other measures were presently found. However, inconsistencies in survivor-proxy agreement are quite common in outcome measures (see, for example, Cusick et al., 2000; Dawson, Markowitz, & Stuss, 2005; Tepper, Beatty, & DeJong, 1996). Thus, the FOP compares well to existing measures on these traditional psychometric characteristics.

Only a small number of other outcome measures have had sensitivity to change investigated: the DRS (van Baalen et al., 2006), FIM+FAM (van Baalen et al., 2006), EBIQ (Svendsen et al., 2004), and the BICRO-39 (Powell et al., 1998). Importantly, of these four instruments, two (the DRS and FIM+FAM) had their sensitivity to change investigated with a between-test interval of 1 year (van Baalen et al., 2006). As such, it is not surprising that they were found to be sensitive to change over time. Also, the DRS and FIM+FAM have been reported to be appropriate for the investigation of outcome only within the first 6 months after injury, after which they will not be sensitive to change, especially within community living survivors (Richardson, 2000). Therefore, the FOP has been shown to not only be a comparatively stable measure of functioning over a short interval (i.e., good test-retest reliability), but is also one of the few measures with empirical evidence that it is able to reflect improvements that may occur over the course of recovery even in the post-acute stage.

A Brief Discussion of Problematic Findings

Despite relatively strong results, two issues became clear from the current analyses. The first were the surprisingly low (and some negative) correlations between the items and the subscale and composite outcome scores calculated as part of the internal consistency investigation. Because some may argue that these findings are grounds for the removal of these questionable items (based on standard psychometric theory), the results spawned the need for the post hoc investigation of the Vancouver and Gorge samples separately. As mentioned, this follow-up analysis was based on the assumption that brain injury outcome is heterogeneous and it was therefore thought that investigating these groups separately may illustrate this point and reinforce the need to retain all of the items. The findings from this analysis imply that for different groups, different domains will be more or less related to overall functioning. Also, it is important to restate one of the main purposes of the FOP at this point: the FOP was designed to be an *inventory* of functioning. As such, it is not expected that every survivor will experience functional problems in every domain – an expectation that was confirmed. However, it is very important to have the domains represented so that the information can be obtained when the domain is problematic, no matter how infrequently. For example, the item inquiring about problems with drugs and/or alcohol was found to be rarely endorsed, but removal of the item would result in overlooking the problem and its impact on the individual's functioning.

Together with the additional evidence put forth for the validity of the FOP based on test content – that at least one survivor and one significant other reported functional problems in every content area covered by an item – removal of these items was not seen as the best option. Instead, a focused investigation is suggested to determine the problem with these items.

The “Activities” items, as a group, were found to be consistently problematic having poor correlations with the remainder of the items, the poorest test-retest reliability (when calculated using the full data and the “problem only” data), and very low survivor-proxy agreement for the 3 frequency items in the subscale. Interestingly, the item-subscale score correlations for the full sample revealed that the items inquiring about satisfaction with activities (i.e., Item #54, Leisure Satisfaction and Item #56, Work Satisfaction) had some of the lowest correlations. This likely indicates that these items are measuring something different from the rest of the items. However, the item-composite score correlations for the full sample revealed that the items inquiring about time spent engaging in different activities (i.e., Item #51, Self-care Activities; Item #52, Leisure Solitary; Item #53, Leisure Active/Social; and Item #55, Work Quantity) had the lowest correlations. This seems to indicate that these items are measuring something different from the rest of the FOP. Conversely, these same items were found to have moderate correlations for the Vancouver sample alone. Obviously, these items should be investigated further. It is important to determine if: 1) there are population factors that affect these scores (e.g., a difference between survivors of stroke and survivors of TBI); 2) the scores were affected (and to what degree) by the sample selection (e.g., a large proportion, 74%, engaged in week day rehabilitation thus limiting their availability to engage in other, work-related activities); 3) the scoring of these items needs to be re-thought (e.g., how the activity credit scores are converted into activity outcome scores); 4) the items are actually inappropriate (i.e., not able to gather valid information); or 5) there is actually an inverse relationship between level of independent activity and overall everyday functioning. In order to investigate these possibilities, both a larger sample should be obtained and validated measures that were designed to measure activity-related functioning (e.g., the CIQ) should be administered. A larger sample would allow for the investigation of the possible effects of the characteristics of the current sample on the scores. Administering a previously validated activities measure with the FOP would allow comparisons to be made between the FOP Activity subscale and the scores of the other measure. Such a comparison could help determine if the FOP provides an accurate measure of activity-related functioning and provide evidence for the validity of this subscale. Unfortunately, this could not be done for the current study, but it is suggested that this approach be taken in the near future.

The second issue is the less-than hoped for survivor-proxy agreement, which was found in both the current and the previous study. Both studies found that a large number of the items showed weak relationships between the ratings provided by the survivors and those provided by their significant others, despite using only the frequency data which ensures both raters are rating the same thing (i.e., how frequently the survivor has experienced the problem). While it is unclear why the survivor-proxy agreement is so poor, two possibilities were addressed using the additional analyses. One possible reason for the discrepancy is that there are “too many” options; most existing outcome measures use scales with 5 or less options (for example, the CIQ, the MPAI-4, the EBIQ) while the FOP asks participants to rate problem frequency on an 11-point scale. It was thought that re-coding the frequency scores to a 3-point scale (i.e., “never a problem”, “sometimes a problem”, “always a problem”) may improve the agreement. However, the results from this analysis were not consistently better. While the agreement for some of the items improved when the scale was restricted (e.g., for item #58, ρ = -.028 while V = .392) the agreement for other items worsened (e.g., for item #5, ρ = .800 while V = .561). Re-coding the frequency scores in this manner did little to improve the survivor-proxy agreement as the agreement was, for the most part, very similar to that found for the full 11-point scale. An obvious limitation is that the participants' responses to the original 11-point scale were merely condensed to produce the 3-point scale. It is unclear how asking the survivors and significant others to provide ratings on a 3-point scale initially might impact agreement. However, the results do seem to indicate that, for the FOP, the number of points on the rating scale has little to do with agreement between survivors and their significant others.

Another possible reason for the less-than-hoped-for agreement is that inaccurate ratings of problem frequency are provided by either the survivors and/or the significant others (based on, for example, poor awareness or a lack of information). Therefore, it was thought that the pairs may show better agreement if providing a more simplistic rating, like whether an area is problematic or not. When the FOP frequency ratings were dichotomized as "agree" (i.e., both survivor and significant other rate the area as a problem or not) or "disagree" (i.e., one member of the pair feels that the area is a problem while the other does not), the percent agreement was very good for a majority of the items. These findings indicate that, although agreement on the actual frequency of many of the problems inquired about in the FOP is poor, survivors and significant others do agree to a very high degree on whether an area is a problem or not.

Whose ratings (survivor or significant other) are more accurate is unclear, and trying to determine this is beyond the scope of the current study. It is possible, however, that asking significant others to rate the frequency of some of the more subjective problems may be problematic (Cusick et al., 2000). Consistent with this, the five items with the best Spearman correlations were: 1) Pain medications taken, 2) Self-care, 3) Problem behaviour, 4) Physical Intimacy problems, and 5) Gross motor, which are all relatively objective or with which the significant other may have had firsthand experience. The five items with the poorest correlations, on the other hand, were: 1) Safety judged by others, 2) Family relationships, 3) Behavioural organization, 4) Retrograde memory, and 5) Work demands, which are generally more subjective problems. Similar results have been found in other studies; raters generally show good agreement on reports of physical deficits, but poorer agreement on reports of cognitive, emotional and behavioural problems (Donaghy & Wass, 1998; Sbordone, Seyranian, & Ruff, 1998; Tepper et al., 1996).

Despite the poor agreement between the survivors and significant others on the frequency of problems, it is this author's opinion that it is still important to gain information from the significant other. The judgement of the significant other regarding the survivor's functioning, and his/her own experience of the impact that the injury has had on his/her life is important for three reasons specific to the significant other. First, obtaining the significant other's thoughts on the major areas of concern may help elicit a dialogue between the survivor and his/her significant other and thus help them address these issues. Second, it may also be beneficial to the significant other to have input into the course of rehabilitation; it is likely that having their opinions heard may help alleviate some of the stress and depression experienced following their loved one's injury. Third, it may make the significant other aware of areas of concern within his/her own life (as much of the focus of the significant other and other individuals will generally be on the survivor) and improve his/her own psychological functioning, as it has been consistently found that many significant others experience a range of mental health issues following their loved one's brain injury (Machamer, Temkin, & Dikmen, 2002; Marsh, Kersel, Havill, & Sleigh, 2002).

Limitations

There are a number of limitations with both the current study and aspects of the FOP that must be addressed. The first two limitations are related to the common practices employed when administering and scoring the FOP. First, grouping together the different question types (i.e., impact, satisfaction and activities) will be less troublesome in a clinical setting than it is in a research setting. When used clinically, it is likely that the individual items (i.e., the profile the FOP provides) will be of primary concern, thus the grouping of the items has little importance. However, as the current purpose was to investigate the psychometric characteristics of aspects of the FOP, this practice may have led to less-than optimal results, especially in the investigation of the internal consistency. While this concern was not addressed in the current study, it is suggested that future studies investigate how the multidimensional nature of the information provided by the FOP affects its reliability and also to determine if these types of information can or should be used more independently.

The second common practice, automatically "filling in" areas reported as not a problem with a score indicating optimal functioning, makes perfect clinical sense; if an individual reports no problems in a domain, he/she is likely functioning optimally in the domain in question. However, as found, this practice does influence the scores and the psychometric characteristics of the FOP outcome scores. While this practice was taken into consideration in a number of analyses, if this had not been done an inaccurate picture of the psychometric characteristics of the FOP outcome scores would have been provided. Thus, while these are not necessarily limitations of the FOP in practice, they are likely limitations to its ability to be accurately investigated using classical psychometric techniques.

The third major limitation is also related to the FOP and the scores used in the analyses - specifically the composite outcome score. It is important to acknowledge that by averaging all of the subscales, the FOP essentially compresses multiple dimensions into a singular, unidimensional score, representing “outcome”. Some of the most popular measures to date, the GOS and FIM + FAM, provide nothing more than an overall level of outcome. Including the option to compute a composite outcome score allows the FOP to be further validated and more easily compared to existing outcome measures. The composite outcome score also lends itself nicely to be utilized for more general research purposes (e.g., broad program evaluation) where the individuals most interested in the results would want to simply know if the survivors “got better” and not the specifics of what areas the most recovery was seen in.

However, it cannot be ignored that the FOP composite score has limitations. The most obvious limitation is that the composite score provides very limited information and thus comparisons within individuals, between individuals or even between measures (when attempting to determine validity, for example) can be misleading. For example, two survivors who have problems in very different areas, for example one who is functioning well in the community and one who is not, could have very similar composite outcome scores. However, their daily functioning and its implications for their quality of life and well-being are likely very different. A similar limitation to the use of the composite score is that, based on the way it is currently calculated, it necessarily includes all areas of functioning and all areas are given an equal weight. This could “hide” results if, for example, utilized to evaluate a very specific program. The composite score may not change significantly, but a specific subscale (e.g., cognitive functioning) may show significant improvement. In other words, the composite score could include items not relevant to the question being asked or group under investigation. The use of the composite outcome score alone also undermines many of the strengths of the FOP. For example, the FOP is currently the most comprehensive measure of outcome following brain injury however this comprehensiveness of little importance if the composite outcome score is all that is looked at. Also, one of the unique aspects of the FOP is that it provides a profile of functioning. This profile is overlooked if the composite outcome score is used.

For these reasons, I believe that the subscale scores provide the most valuable information that can be obtained from the FOP, and recommend that in most instances these scores be utilized. However, one of the primary benefits of the FOP is its flexibility. It allows for the use of all of the score types that it can provide and allows individual users, based on the type of question they are trying to answer, to decide which level of score is most appropriate for their purposes.

The fourth limitation is that, while the full sample is large in comparison to most studies of brain injury survivors, because of the purpose of the present study (i.e., psychometric validation of a new measure) the sample size is small, especially for the analyses that do not use the full sample. Because of this, the current results cannot be generalized to all brain injury survivors and significant others.

Fifth, this study relied on the survivors’ self-report of major injury characteristics (e.g., length of PTA) and this information was then used to investigate the validity of the FOP.

While the use of actual medical files would be ideal in such cases to ensure that accurate data were employed, this was not possible for the current study. It is very possible that different results for the analysis investigating the effect of participant and injury characteristics on FOP scores would be obtained if this official data were available.

The sixth limitation is associated with participant selection; a convenience sample was used as only those who were engaged in rehabilitation at Gorge Road Hospital or who replied to the newspaper advertisements in Vancouver could be included in the study. The similarity of these participants to the general population of brain injury survivors in British Columbia, let alone all of Canada, is questionable. However, the sample did provide a good breadth of injury etiology, severity, and age at the time of injury which is thought to be beneficial given the purpose of the current study.

A seventh limitation, also related to the sample, is that the long-interval group was composed of survivors from the GRH sample only while the short-interval group was composed of survivors from the Vancouver sample only. It is unclear what influence this may have had on the results. It is assumed that the change over time found for the long-interval group was based more on the fact that this sample was involved in rehabilitation between FOP administrations and not due to other confounding variables related to location (e.g., the fact that GRH is located in Victoria, a much smaller city than Vancouver). However it is advised that, in future studies, this be controlled for, preferably by including some participants from each location in both the long- and short-interval groups.

The final limitation is related to the methods. It is important to note that the analysis of change over time is an indication of *statistical* change only, not of *clinical* change. In other words, while there is a statistical difference between the short interval group and the long interval group with the long interval group showing greater change, it is unclear if the change is meaningful; did those in the long interval group feel as though their functioning had improved more than those in the short interval group? Did the significant others or clinicians involved with those in the long interval group notice a change in the survivors' functioning? Unfortunately, these questions cannot be answered with the data or statistical results obtained, but are worth future investigation.

Future Directions

There are many studies that should be completed in order to further validate the FOP for use with brain injury survivors and significant others. In general a study comprising a much larger, more diverse sample would provide results that could be deemed more generalizable. Specifically, a more in-depth investigation of the Activity questions should be completed in order to determine their appropriateness as these items were consistently found to be psychometrically problematic. Also, further investigation of the FOP SO needs to be done: the reliability (especially test-retest) and sensitivity to change are yet to be determined and are required if, as with the FOP, the FOP SO is to be used in clinical and experimental settings. Ideally, the FOP and FOP SO would be utilized together in the development of individual rehabilitation programs. Many experts have expressed that subjective information from the survivor and significant others should be included in this process (see, for example Prigatano, 1997; B. A. Wilson, 2002). Currently, the FOP is the only outcome measure that is able to systematically provide clinicians with this in-depth and subjective information. Therefore, showing that they are both reliable and valid measures of outcome and functioning is a necessary step.

Because of the unique qualities of the FOP, there are a wide variety of experimental investigations to which it can be applied that would be of great benefit to survivors, their families, and clinicians. For example, there is currently no empirical evidence that looks at how survivors perceive their functioning, and how this perception changes over time following injury. This information could be used to provide better education to survivors and their families regarding what they may experience as recovery progresses. Also, enhanced awareness of the recovery process could aid clinicians when dealing with survivors at different points in recovery. Because the FOP allows data to be collected about such a wide range of functional areas, it is the perfect tool to study the pattern(s) of recovery after brain injury. Recent literature has begun to note the burden that caregivers experience and the FOP SO would be ideal to measure the impact of the injury on caregivers over time in conjunction with their perception of problem frequency.

Conclusions

In 1999, Turner-Stokes outlined 5 criteria that apply to the selection of an outcome measure to be used to assess rehabilitation. These criteria are that the measure must be:

1. valid and reliable
2. relevant to the area of intervention
3. sensitive to change occurring over the time for which the treatment is given
4. feasible to use in the course of routine practice
5. the team should feel that they have ownership of the results

Based on the current findings, the FOP would fit these criteria for individuals in the post-acute stage of rehabilitation. 1) The results of the current study have shown the FOP to have satisfactory internal reliability and test-retest reliability. While the survivor-proxy agreement was questionable, this is not uncommon among measures of brain injury outcome. In addition, multiple lines of validity evidence have been presented indicating that the FOP is a valid measure of outcome following brain injury. 2) Together, the procedure taken to develop the FOP in conjunction with the results of the analyses, particularly those looking at content validity, are evidence that the FOP is relevant to the general population of community-living brain injury survivors. However, it is unclear whether the change that was found can be attributed to the intervention (i.e., the out-patient rehabilitation program) or to the length of time between interviews for the long interval group. While it is likely that the change is more strongly related to the intervention, especially given the average length of time since injury for the long interval group, based on the design of the current study this cannot be concluded definitely. 3) The specific investigation of sensitivity to change over time has shown that the FOP is able to detect change when it occurs, even over relatively short intervals when the survivors are engaged in a rehabilitation program, even if these survivors are many years post injury. 4) Based on knowledge of the measure and its administration procedures, it is this author's opinion that the FOP is feasible to use in the course of routine practice: administration of the whole measure takes only an hour yet it is flexible enough to allow only specific subscales to be administered if this is more appropriate. 5) Finally, because of the face value of the information provided, all members of the rehabilitation team including the survivor and his/her significant other will be able to understand and utilize the information obtained from the FOP.

Overall, this study has shown the FOP to be a valuable new measure of outcome following brain injury with satisfactory psychometric properties. The FOP is unique compared to existing outcome measures providing new information that can benefit the survivor, his/her significant other, clinicians and researchers.

References

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (1999). *Standards for educational and psychological testing.* Washington, DC: American Educational Research Association.

Anderson, M. I., Parmenter, T. R., & Mok, M. (2002). The relationship between neurobehavioural problems of severe traumatic brain injury (TBI), family functioning and the psychological well-being of the spouse/caregiver: Path model analysis. *Brain Injury, 16*(9), 743-757.

Angus, D. E., Cloutier, E., Albert, T., Chenard, D., Shariatmadar, A., Pickett, W., et al. (1998). *The economic burden of unintentional injury in Canada*: SMARTRISK.

Baguley, l. J., Cooper, J., & Felmingham, K. (2006). Aggressive behavior following traumatic brain injury: How common is common? *Journal of Head Trauma Rehabilitation, 21*(1), 45-56.

Ben-Yishay, Y. (2000). Postacute Neuropsychological Rehabilitation: A Holistic Prespective. In A.-L. Chrisetnsen & B. P. Uzzell (Eds.), *International Handbook of Neuropsychological Rehabilitation* (pp. 127-136). New York: Kluwer Academic/Plenum Publishers.

Bernstein, D. M. (1999). Recovery from mild head injury. *Brain Injury, 13*(3), 151-172.

Bezruczko, N. (2005). Rasch model essentials. In N. Bezruczko (Ed.), *Rasch Measurment in Health Sciences* (pp. 35-71). Minnesota: JAM Press.

Binder, S., Corrigan, J. D., & Langlois, J. A. (2005). The public health approach to traumatic brain injury: An overview of CDC's research and programs. *Journal of Head Trauma Rehabilitation, 20*(3), 189-195.

Canadian Institute for Health Information (2006, August). *Head Injuries in Canada: A Decade of Change (1994-1995 to 2003-2004).* Retrieved 6 Nov, 2006, from http://secure.cihi.ca/cihiweb/en/downloads/ analysis_ntr_2006_e.pdf

Cattelani, R., Tanzi, F., Lombardo, F., & Mazzucchi, A. (2002). Competitive re-employment after severe traumatic brain injury: Clinical, cognitive and behavioural predictive variables. *Brain Injury, 16*(1), 51-64.

Cicerone, K. D., Mott, T., Azulay, J., & Friel, J. C. (2004). Community integration and satisfaction with functioning after intensive cognitive rehabilitation for traumatic brain injury. *Archives of Physical Medicine and Rehabilitation, 85*, 943-950.

Coetzer, R., & Rushe, R. (2005). Post-acute rehabilitation following traumatic brain injury: Are both early and later improved outcomes possible? *International Journal of Rehabilitation Research, 28*(4), 361-363.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillside, NJ: Lawrence Erlbaum Associates.

Cusick, C. P., Gerhart, K. A., & Mellick, D. C. (2000). Participant-proxy reliability in traumatic brain injury outcome research. *Journal of Head Trauma Rehabilitation, 15*(1), 739-749.

Dawson, D. R., Markowitz, M., & Stuss, D. T. (2005). Community integration status 4 years after traumatic brain injury: Participant-proxy agreement. *Journal of Head Trauma Rehabilitation, 20*(5), 426-435.

Deloche, G., Dellatolas, G., & Christensen, A.-L. (2000). The European Brain Injury Questionnaire: Patients' and families' subjective evaluation of brain-injured patients' current and prior to injury difficulties In A.-L. Chrisetnsen & B. P. Uzzell (Eds.), *International Handbook of Neuropsychological Rehabilitation* (pp. 81-92). New York: Kluwer Academic/Plenum Publishers.

Dikmen, S. S., Machamer, J. E., Powell, J. M., & Temkin, N. R. (2003). Outcome 3 to 5 years after moderate to severe traumatic brain injury. *Archives of Physical Medicine and Rehabilitation, 84*, 1449-1457.

Diller, L., & Ben-Yishay, Y. (2003). The Clinical Utility and Cost-Effectiveness of Comprehensive (Holistic) Brain Injury Day-Treatment Programs. In G. P. Prigatano & N. H. Pliskin (Eds.), *Clinical Neuropsychology and Cost Outcome Research: A Beginning* (pp. 293-328). New York: Psychology Press, Inc.

Donaghy, S., & Wass, P. J. (1998). Interrater reliability of the Functional Assessment Measure in a brain injury rehabilitation program. *Archives of Physical Medicine and Rehabilitation, 79*, 1231-1236.

Donnelly, J. P., Donnelly, K., & Grohman, K. K. (2005). A multi-perspective concept mapping study of problems associated with traumatic brain injury. *Brain Injury, 19*(13), 1077-1085.

Fork, M., Bartels, C., Ebert, A. D., Grubich, C., Synowitz, H., & Wallesch, C.-W. (2005). Neuropsychological sequelae of diffuse traumatic brain injury. *Brain Injury, 19*(2), 101-108.

Garson, G. D. (2008). *Reliability Analysis.* Retrieved January 5, 2009, from North Carolina State University, Quantitative Research in Public Administration course Web site: http://faculty.chass.ncsu.edu/garson/PA765/pa765syl.htm

Ghajar, J. (2000). Traumatic brain injury. *Lancet, 356*, 923-929.

Gurka, J. A., Felmingham, K. L., Baguley, I. J., Schotte, D. E., Crooks, J., & Marosszeky, J. E. (1999). Utility of the Functional Assessment Measure after discharge from inpatient rehabilitation. *Journal of Head Trauma Rehabilitation, 14*(3), 247-256.

Hall, K. M., Bushnik, T., Lakisic-Kazazic, B., Wright, J., & Cantagallo, A. (2001). Assessing traumatic brain injury outcome measures for long-term follow-up of community-based individuals. *Archives of Physical Medicine and Rehabilitation, 82*, 367-374.

Harris Interactive Inc. (2000). *Public Perceptions of Brain and Head Injuries*. (Study No. 11681) New York.

Hawley, C. A., Taylor, R., Hellawell, D. J., & Pentland, B. (1999). Use of the functional assessment measure (FIM+FAM) in head injury rehabilitation: A psychometric analysis. *Journal of Neurology, Neurosurgery, and Psychiatry, 67*, 749-754.

Hellawell, D., & Pentland, B. (2001). Relatives' reports of long term problems following traumatic brain injury or subarachnoid haemorrhage. *Disability and rehabilitation, 23*(7), 300-305.

Henderson, R. K. (2001). Understanding internal consistency reliability estimates: A conceptual primer on coefficient alpha. *Measurement and Evaluation in Counselling and Development, 34*(3), 177-189.

Hoofien, D., Gilboa, A., Vakil, E., & Donovick, P. J. (2001). Traumatic brain injury (TBI) 10-20 years later: A comprehensive outcome study of psychiatric symptomatology, cognitive abilities and psychosocial functioning. *Brain Injury, 15*(3), 189-209.

Huang, M. E., Wartella, J., Kreutzer, J., Broaddus, W., & Lyckholm, L. (2001). Functional outcomes and quality of life in patients with brain tumours: A review of the literature. *Brain Injury, 15*(10), 843-856.

Jennett, B., & Bond, M. (1975). Assessment of outcome after severe brain damage: A practical scale. *Lancet, 305*(7905), 471-534.

Joschko, M. & Skelton, R. (2003). *The Functional Outcome Profile*. Unpublished. University of Victoria.

Keith, R. A., Granger, C. V., Hamilton, B. B., & Sherwin, F. S. (1987). The functional independence measure: A new tool for rehabilitation. *Advances in Clinical Rehabilitation, 1*, 6-18.

Kolb, B. (1995). *Brain Plasticity and Behavior*. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Langlois, J. A., Marr, A., Mitchko, J., & Johnson, R. L. (2005). Tracking the silent epidemic and educating the public: CDC's traumatic brain injury-associated activities under the TBI Act of 1996 and the Children's Health Act of 2000. *Journal of Head Trauma Rehabilitation, 20*(3), 196-204.

Linn, R. T., Blair, R. S., Granger, C. V., Harper, D. W., O'Hara, P. A., & Maciura, E. (1999). Does the Functional Assessment Measure (FAM) Extend the Functional Independence Measure (FIM) Instrument? A rasch analysis of stroke inpatients. *Journal of Outcome Measurement, 3*(4), 339-359.

Machamer, J., Temkin, N., & Dikmen, S. S. (2002). Significant other burden and factors related to it in traumatic brain injury. *Journal of Clinical and Experimental Neuropsychology, 24*(4), 420-433.

Malec, J. F. (2004). Comparability of Mayo-Portland Adaptability Inventory ratings by staff, significant others and people with acquired brain injury. *Brain Injury, 18*(6), 563-575.

Malec, J. F., Kragness, M., Evans, R. W., Finlay, K. L., Kent, A., & Lezak, M. D. (2003). Further psychometric evaluation and revision of the Mayo-Portland Adaptability Inventory in a national sample. *Journal of Head Trauma Rehabilitation, 18*(6), 479-492.

Malec, J. F., Moessner, A. M., Kragness, M., & Lezak, M. D. (2000). Refining a measure of brain injury sequelae to predict postacute rehabilitation outcome: Rating scale analysis of the Mayo-Portland Adaptability Inventory. *Journal of Head Trauma Rehabilitation, 15*(1), 670-682.

Malec, J. F., & Thompson, J. M. (1994). Relationship of the Mayo-Portland Adaptability Inventory to functional outcome and cognitive performance measures. *Journal of Head Trauma Rehabilitation, 9*(4), 1-15.

Marsh, N. V., Kersel, D. A., Havill, J. H., & Sleigh, J. W. (2002). Caregiver burden during the year following severe traumatic brain injury. *Journal of Clinical and Experimental Neuropsychology, 24*(4), 434-447.

Mateer, C. A. (1997). Rehabilitation of individuals with frontal lobe impairment. In J. Leon-Carrion (Ed.), *Neuropsychological rehabilitation: Fundamentals, innovations and directions* (pp. 285-300). Delray Beach, FL: GR/St Lucie Press.

McCarthy, M. L., Dikmen, S. S., Langlois, J. A., Selassie, A. W., Gu, J. K., & Horner, M. D. (2006). Self-reported psychosocial health among adults with traumatic brain injury. *Archives of Physical Medicine and Rehabilitation, 87*(7), 953-961.

McPherson, K. M., Pentland, B., & McNaughton, H. K. (2000). Brain injury - The perceived health of careres. *Disability and Rehabilitation, 22*(15), 683-689.

Mailhan, L., Azouvi, P., & Dazord, A. (2005). Life satisfaction and disability after severe traumatic brain injury. *Brain Injury, 19*(4), 227-238.

National Institutes of Health Consensus Development Panel on Rehabilitation of Persons with Traumatic Brain Injury (1999). Rehabilitation of Persons with Traumatic Brain Injury. *JAMA, 282*, 974-983.

Osterlind, S. J. (2006). *Modern Measurement: Theory, Principles and Applications of Mental Appraisal.* Upper Saddle River, NJ: Prentice Hall.

Pietrapiana, P., Bronzino, M. P., Perino, C., & Rago, R. (1997). Acquired brain injury physical rehabilitation. In J. Leon-Carrion (Ed.), *Neuropsychological Rehabilitation: Fundamentals, Innovations and Directions* (pp. 227-252). Delray Beach, Florida: GR/St. Lucie Press.

Powell, J. H., Beckers, K., & Greenwood, R. J. (1998). Measuring progress and outcome in community rehabilitation after brain injury with a new asessment instrument - The BICRO-39 scales. *Archives of Physical Medicine and Rehabilitation, 79*, 1213-1225.

Price, J. R. (2007). *A Psychometric Analysis of the Functional Outcome Profile (FOP).* Unpublished doctoral dissertation, University of Victoria, Victoria, British Columbia, Canada.

Prigatano, G. P. (1997). The problem of impaired self-awareness in neuropsychological rehabilitation. In J. Leon-Carrion (Ed.), *Neuropsychological Rehabilitation: Fundamentals, Innovations and Directions* (pp. 301-312). Delray Beach, Florida: GR/St. Lucie Press.

Prigatano, G. P., Fordyce, D. J., Zeiner, H. K., Roueche, J. R., Pepping, M., & Case-Wood, B. (1984). Neuropsychological rehabilitation after closed head injury in young adults. *Journal of Neurology, Neurosurgery, and Psychiatry, 47*, 505-513.

Rappaport, M., Hall, K. M., Hopkins, K., Belleza, T., & Cope, N. (1982). Disability Rating Scale for severe head trauma: Coma to community. *Archives of Physical Medicine and Rehabilitation, 63*, 118-123.

Reynolds, C. R., Livingston, R. B., & Willson, V. (2006). Validity for teachers. In *Measurement and Assessment in Education* (pp. 117-140). Boston: Pearson Education, Inc.

Rice-Oxley, M., & Turner-Stokes, L. (1999). Effectiveness of brain injury rehabilitation. *Clinical Rehabilitation, 13*(suppl 1), 7-24.

Richardson, J. T. E. (2000). *Clinical and Neuropsychological Aspects of Closed Head Injury: Second Edition*. Philadelphia: Taylor & Francis Group.

Rothweiler, B., Temkin, N., & Dikman, S. S. (1998). Aging effect on psychosocial outcome in traumatic brain injury. *Archives of Physical Medicine and Rehabilitation, 79*, 881-887.

Sander, A. M., Fuchs, K. L., High, W. M., Hall, K. M., Kreutzer, J. S., & Rosenthal, M. (1999). The Community Integration Questionnaire revisited: An assessment of factor structure and validity. *Archives of Physical Medicine and Rehabilitation, 80*, 1303-1308.

Sander, A. M., Seel, R. T., Kreutzer, J. S., Hall, K. M., High, W. M., & Rosenthal, M. (1997). Agreement between persons with traumatic brain injury and their relatives regarding psychosocial outcome using the Community Integration Questionnaire. *Archives of Physical Medicine and Rehabilitation, 78*(4), 353-357.

Sarajuuri, J. M., Kaipio, M.-L., Koskinen, S. K., Niemela, M. R., Servo, A. R., & Vilkki, J. S. (2005). Outcome of a comprehensive neurorehabilitation program for patients with traumatic brain injury. *Archives of Physical Medicine and Rehabilitation, 86*, 2296-2302.

Sbordone, R. J., Seyranian, G. D., & Ruff, R. M. (1998). Are the subjective complaints of traumatically brain injured patients reliable? *Brain Injury, 12*(6), 505-515.

Schmidt, M. F., Garvin, L. J., Heinemann, A. W., & Kelly, J. P. (1995). Gender- and age-related role changes following brain injury. *Journal of Head Trauma Rehabilitation, 10*(4), 14-27.

Skelton, R. W., Joschko, M., Larke, S., Kodalan, K., Mariscak, A., & Price, J. (2005). *The Functional Outcome Profile for evaluating outcome after acquired brain injury: Sensitivity to (impact on) suvivors, family and change over time*. Unpublished manuscript, University of Victoria.

Sopena, S., Dewar, B.-K., Nannery, R., Teasdale, T. W., & Wilson, B. A. (2007). The European Brain Injury Questionnaire (EBIQ) as a reliable outcome measure for use with people with brain injury. *Brain Injury, 21*(10), 1063-1068.

Stambrook, M., Moore, A. D., Lubusko, A. A., Peters, L. C., & Blumenschein, S. (1993). Alternatives to the Glasgow Coma Scale as a quality of life predictor following traumatic brain injury. *Archives of Clinical Neuropsychology, 8*(2), 95-103.

Stein, D. G. (2000). Brain injury and theories of recovery. In A.-L. Chrisetnsen & B. P. Uzzell (Eds.), *Internaional Handbook of Neuropsychological Rehabilitation* (pp. 9-32). New York: Kluwer Academic.

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

Svendsen, H. A., & Teasdale, T. W. (2006). The influence of neuropsychological rehabilitation on symptomatology and quality of life following brain injury: A controlled long-term follow-up. *Brain Injury, 20*(12), 1295-1306.

Svendsen, H. A., Teasdale, T. W., & Pinner, M. (2004). Subjective experience in patients with brain injury and their close relatives before and after a rehabilitation programme. *Neuropsychological Rehabilitation, 14*(5), 495-515.

Teasdale, T. W., Chrisetnsen, A.-L., Gerard, K. W., Deloche, L. B., Stachowiak, F., Castro-Cldas, A., et al. (1997). Subjective experience in brain-injured patients and their close relatives: A European Brain Injury Questionnaire study. *Brain Injury, 11*(8), 543-563.

Teasdale, T. W., & Engberg, A. W. (2005). Subjective well-being and quality of life following traumatic brai injury in adults: A long-term population-based follow-up. *Brain Injury, 19*(12), 1041-1048.

Tepper, S., Beatty, P., & DeJong, G. (1996). Outcomes in traumatic brain injury: Self-report versus report of significant others. *Brain Injury, 10*(8), 575-581.

Tomberg, T., Toomela, A., Pulver, A., & Tikk, A. (2005). Coping strategies, social support, life orientation and health-related quality of life following traumatic brain injury. *Brain Injury, 19*(4), 1181-1190.

Turner-Stokes, L. (1999). Outcome measurement in brain injury rehabilitation – towards a common language. *Clinical Rehabilitation, 13*, 273-275.

van Baalen, B., Odding, E., van Woensel, M., van Kessel, M., Roebroeck, M., & Stam, H. (2006). Reliability and sensitivity to change of measurement instruments used in a traumatic brain injury population. *Clinical Rehabilitation, 20*, 686-700.

van Baalen, B., Odding, E., Maas, A. I. R., Ribbers, G. M., Bergen, M. P., & Stam, H. J. (2003). Traumatic brain injury: Classification of initial severity and determination of functional outcome. *Disability and Rehabilitation, 25*(1), 9-18.

Vermeersch, D. A., Lambert, M. J., & Burlingame, G. M. (2000). Outcome questionnaire: Item sensitivity to change. *Journal of Personality Assessment, 74*(2), 242-261.

Wehman, P., Targett, P., West, M., & Kregel, J. (2005). Productive work and employment for persons with traumatic brain injury: What have we learned after 20 years? *Journal of Head Trauma Rehabilitation, 20*(2), 115-127.

Willer, B., Rosenthal, M., Kreutzer, J. S., Gordon, W. A., & Rempel, R. (1993). Assessment of community integration following rehabilitation for traumatic brain injury. *Journal of Head Trauma Rehabilitation, 8*, 75-87.

Wilson, B. A. (2002). Towards a comprehensive model of cognitive rehabilitation. *Neuropsychological Rehabilitation, 12*(2), 91-110.

Wilson, J. T. L., Pettigrew, L. E. L., & Teasdale, G. M. (1998). Structrued interviews for the Glasgow Outcome Scale and the Extended Glasgow Outcome Scale: Guidelines for their use. *Journal of Neurotrauma, 15*(8), 573-585.

Wood, R. L., Liossi, C., & Wood, L. (2005). The impact of head injury neurobehavioural sequelae on personal relationships: Preliminary findings. *Brain Injury, 19*(10), 845-851.

World Health Organization (2001). *International Classification of Functioning, Disability and Health.* Geneva: Author.

Wright, J. (2000). The Disability Rating Scale. In *The Center for Outcome Measurement in Brain Injury.* Retrieved May 26, 2009, from http://www.tbims.org/combi/drs

Appendix A

Detailed description of existing outcome measures

Glasgow Outcome Scale (GOS) and Extended Glasgow Outcome Scale (GOSE)

Description: The GOS (Jennett & Bond, 1975) is a brief descriptive outcome scale that is used in the investigation of early gross outcome following brain injury (Hall, Bushnik, Lakisic-Kazazic, Wright & Cantagallo, 2001), and is widely used due to its ability to provide a quick, clinically relevant measure of a brain injury survivor's overall functioning. The GOS provides 5 outcome categories on which to rate the brain injury survivor: dead, vegetative, severely disabled, moderately disabled, and good recovery (Hall et al., 2001). Thus, the measure focuses on how the injury has affected the overall functioning of the individual as opposed to specific deficits and symptoms caused by the injury (Wilson, Pettigrew & Teasdale, 1998).

An extended version of the GOS - the Extended Glasgow Outcome Scale (GOSE) (Wilson, Pettigrew & Teasdale, 1998) - was developed after determining that the severe and moderate disability and good recovery categories were insensitive to the range of recovery possible within the categories (Wilson, Pettigrew & Teasdale, 1998). Therefore, each of these categories were further divided into "upper" and "lower" levels of outcome (Richardson, 2000) resulting in 8 categories.

Scores for both the GOS and the GOSE are determined by a trained rater, and a structured interview has been developed for the GOSE to ensure accurate ratings across individuals (Wilson, Pettigrew & Teasdale, 1998). The ratings are based on observations of, and interviews with, the brain injury survivor. Possible scores range from 0 (dead) to 5 (good recovery) on the GOS, and from 0 (dead) to 8 (upper good recovery) on the GOSE.

Psychometric Validation: Inconsistent inter-rater reliability for the GOS and the GOSE is found within the literature. Wilson et al. (1998) investigated the agreement between two raters for the outcome for 50 patients and found that the overall agreement was 92% for the GOS and 78% for the GOSE, with weighted kappa statistics of 0.89 for the GOS and 0.85 for the GOSE indicating excellent inter-rater agreement. However, another study done by van Baalen and colleagues (2006), found the ratings for 25 patients showed 73% agreement for GOS and only 50% agreement for the GOSE and squared weighted kappa statistics of only 0.38 for the GOS, and 0.56 for the GOSE. A much larger study (N = 861) found that prior to training, an overall agreement between ratings assigned and those deemed appropriate by an expert rater was only 59%, with the overall difference between these ratings being highly significant (*Wilcoxon* $Z = -3.38$; $p = 0.001$, two-tailed) (Wilson, Slieker, Legrand, Murray, Stocchetti & Maas, 2007). However, following training, a marked decline in the number of discrepancies was found (Wilson et al., 2007). The reason for the discrepancies in inter-rater agreement between studies is unclear; it is possible that level of training in the administration and scoring the GOS and GOSE plays a role.

The GOS has been, however, shown to have good concurrent validity with standardized neuropsychological tests. Satz et al. (1998) found that the GOS classifications were significantly related to scores on the Grooved Pegboard, Colour Trails 1 and 2, Symbol Digit Test, Word List memory. GOS scores have also been shown to be significantly related to the Patient Competency Rating Scale and the Employability Rating Scale in a sample of 100 moderate to severe TBI survivors (Satz et al., 1998).

While the GOS and GOSE are still commonly used in the acute stage of recovery (such as hospital discharge), research has shown that very few brain injury survivors exhibit a change of category on the GOS more than six months following their accident (see Richardson, 2000 for a review). These results imply that either recovery is complete by 6 months post-injury – which, as previously mentioned, the literature has consistently shown to not be the case – or that the GOS is insensitive to the changes that do occur after the first 6 months of recovery.

Evaluation: The GOS has been shown to suffer from ceiling effects (Hall, Bushnik, Lakisic-Kazazic, Wright, & Cantagallo, 2001). Also, as previously mentioned, the GOS appears to be insensitive to recovery after approximately 6-months post-injury (Richardson, 2000). Therefore, the GOS appears to lack the sensitivity required to investigate long-term outcome, especially in community-living individuals.

The degree of variability found for analyses of inter-rater reliability is also of concern. Wilson et al. (2007) found that raters often erred in the direction of attributing less disability to the brain injury survivor than was deemed necessary by an expert. This may be due to rater bias in the case of individuals who survive severe brain injuries; a “good outcome” relative to what is expected of someone with that degree of injury does not meet the actual requirements for a rating of “good outcome” on the GOSE. This may account for some of the variability found in inter-rater reliability.

Furthermore, the GOS and GOSE are lacking a full psychometric analysis. There has been no investigation of the internal reliability of the measures, and they have also not been submitted to modern psychometric analyses such as Rasch analysis.

Overall, the GOS and GOSE both sacrifice detail in order to maintain their simplicity, which, depending on the circumstances under which outcome is being measured, can be regarded as either a benefit or a limitation. However, the GOS and GOSE do not allow for interpretation of specific areas that an individual shows improvement in or those which he or she still finds problematic, nor do they consider the perspective of the significant other when determining ratings. Therefore, the GOS and GOSE are limited in use beyond the acute rehabilitation stage or as a gross classification tool.

Citations: The Web of Science data base revealed 2766 citations of the GOS seminal article by Jennett and Bond (1975). The seminal article for the GOSE – Wilson, Pettigrew and Teasdale (1998) – has also been cited an additional 137 times according to Web of Science. Both of these articles are still, as of the last 5 years, being cited at a rather consistent rate.

Disability Rating Scale (DRS)

Description: The DRS (Rappaport, Hall, Hopkins, Belleza & Cope, 1982) is used to describe a brain injury survivor's level of disability based on his/her ability to complete activities of daily living and the degree of assistance that is required. As a brief descriptive outcome scale, it has essentially replaced the GOS (Wright, 2000). The scale consists of 8 items that are within 4 categories: 1) arousal, awareness, and responsivity; 2) cognitive ability for self-care activities; 3) dependence on others; and 4) psychosocial adaptability. A higher total score is indicative of more severe levels of disability. Rappaport and colleagues (1982) emphasize that it is intended to chart the progress of the brain injury survivor "through the midzone of the recovery spectrum, between early arousal from coma and early sentient functioning". Administration of the DRS involves a trained rater determining the survivor's score based on observation, and interviews with those who care for the individual if necessary.

Psychometric Validation: The DRS has been found to have variable internal reliability depending on the subscale. Rappaport and colleagues (1982) found that the correlations between scale items ranged from 0.23 to 0.98, with the DRS total score have correlations ranging from 0.54 to 0.96 with each item (all significant $p < 0.01$). However, this investigation of internal reliability used Pearson correlations as opposed to the alpha statistic which provides a more accurate estimate of internal reliability as mentioned above.

The inter-rater reliability of the DRS is inconsistent throughout the literature. Rappaport et al. (1982) found, for 88 brain injury survivors each rated by 3 raters, inter-rater reliability across pairs of raters ranged from 0.97 and 0.98. However, van Baalen et al. (2006) found the DRS to have a squared weighted kappa of 0.62 at hospital discharge, and a kappa of only 0.003 one year following discharge. The reason for these inconsistencies is not clear. One possibility is sample size; van Baalen and colleagues (2006) had a much smaller sample than Rappaport et al. (1982), with 25 participants at hospital discharge and 14 at the 1 year follow-up. Another possibility is the use of different statistics; Rappaport et al. (1982) used traditional Pearson correlations, while van Baalen et al. (2006) employed squared weighted kappa statistics which could account for some of the discrepancy found.

On the other hand, the DRS has been found to show good concurrent validity. Neese, Caroselli, Klaas, High, Becker & Scheibel (2000) found the scores on the DRS to be significantly associated with performance in multiple domains measured by neuropsychological tests such that better functioning as measured by the DRS is significantly correlated with better performance on neuropsychological testing. Also, Pearson correlation coefficients for the DRS and the CHART were found to be -0.53 (Zhang, Abreu, Gonzales, Seale, Masel & Ottenbacher, 2002).

Studies have also shown the DRS to be sensitive to change following brain injury. The DRS demonstrated sensitivity to change when scores obtained at hospital discharge were compared to scores from one year post-injury (van Baalen et al., 2006). The DRS has also been reported to be more sensitive than the GOS to changes occurring over this time period (Rappaport et al., 1982).

Evaluation: While still a measure of gross functioning, the DRS may be more sensitive to change within the first year post-injury than the GOS (Rappaport et al., 1982). However, studies investigating the sensitivity of the GOS compared scores obtained approximately a year apart; it is unclear if it is sensitive to changes occurring over shorter intervals of recovery. Also, the psychometric analyses of the DRS are for the most part inconsistent and also incomplete; the DRS is lacking any easily identifiable investigation of test-retest reliability.

Beyond these problems, one of the major limitations of the DRS is that it is solely an expert-rated measure and does not include the perspective of the significant other, nor input from the survivor. Therefore, like the GOS and GOSE, the DRS has great usefulness as a measure of gross outcome, especially in the acute rehabilitation phase after brain injury. However, its utility is limited as a measure of long-term outcome in community-based individuals.

Citations: Web of Science produced 344 citations of the DRS seminal article by Rappaport, Hall, Hopkins, Belleza and Cope (1982). While the recent yearly citations of this article have been somewhat erratic, over the past 5 years Rappaport et al. (1982) has been cited at least 20 times per year, with 13 citations in 2007 up to the 18th of August.

Functional Independence Measure (FIM™) and Functional Assessment Measure (FAM)

Description: The FIM™ (Keith, Granger, Hamilton & Sherwin, 1987) was developed as a standardized measure of functioning to be used with individuals during inpatient care (Gurka, Felmingham, Baguley, Schotte, Crooks & Marosszeky, 1999). The FIM™ focuses primarily on the areas of motor and self-care skills as they are involved in activities of daily living. The measure includes 18 items which are rated on a seven-point scale were 1 reflects complete dependence and 7 complete independence in completion of the task in question (Gurka et al., 1999), resulting in a range of scores from 18 to 126.

However, the FIM™ is not brain-injury-specific as it can be used to assess any rehabilitation patient. As a result, the FIM™ places little emphasis on cognitive functioning. To address this issue, the FAM was developed and includes an additional 12 items that expand on the TBI-specific functional areas which are under-emphasized in the FIM™ alone (Hall et al., 2001). The FAM was introduced to accompany the FIM™ and, in the context of brain injury research and rehabilitation, they are generally administered as one measure known as the FIM+FAM. Together the FIM+FAM is comprised of 30 items rated on the same seven-point scale described above. Scores on the whole FIM+FAM can range from 30 to 210.

The FIM+FAM is administered and scored by a trained rater based on observation of, and interviews with, the brain injury survivor.

Psychometric Validation: The FIM+FAM has been shown to have excellent internal reliability. Factor analysis of the FIM+FAM done by Hawley, Taylor, Hellawell and Pentland (1999) looking at 965 patients with 2268 administrations found two principle components: the physical component with 16 items and the cognitive component with 14 items. Based on these two components, Cronbach's α was 0.99 for the physical scale, 0.98 for the cognitive scale, and 0.99 for the whole 30 item scale (Hawley et al., 1999).

The FIM+FAM also demonstrates good inter-rater reliability. In a sample of 30 inpatients with mixed diagnoses, the inter-rater reliability for the items was found to range from 0.35 to 0.95, with only seven items showing less than 70% total (McPherson, Pentland, Cudmore & Prescott, 1996). Similar results from van Baalen et al. (2006) found the FIM™ to have a squared weighted kappa statistic of 0.80, and the FAM a kappa of 0.69 at hospital discharge. At 1-year post-discharge, the kappa for the FIM™ was very similar to that obtained earlier (0.75), however for the FAM the kappa increased to 0.95 indicating almost perfect agreement between raters (van Baalen et al., 2006).

Although the FIM™ is usually administered by a trained rater and scores are based on observation of the brain injury survivor, the participant-proxy agreement has been investigated. Cusick, Gerhart and Mellick (2000) interviewed both survivors and significant others regarding the functioning of the brain injury survivor, and completed the FIM™ based on these interviews. It was found that the FIM™ total showed strong participant-proxy agreement (ICC = 0.70). However, it was also found that the mean difference between survivors and proxies was significant (Cusick et al., 2000). Thus, while the scores based on interviews with the survivors and proxies may have been in the same direction (i.e. interviews with survivors and proxies would both lead to scores representing either a higher degree of dependence or a lower degree of dependence) the scores themselves were significantly different from each other.

The FIM+FAM has been shown to have concurrent validity with neuropsychological tests, as well as other outcome measures. McPherson, Berry and Pentland (1997) found scores on the FIM+FAM to be correlated with outcomes on Trail-making B, Story Recall, Rey Complex Figure, as well as the MMSE Orientation and the MMSE Total. Gurka and colleagues (1999) conducted separate analyses on subscales of the FIM™ and FAM – the FIM™ motor, FAM motor, FIM™ cognitive and FAM cognitive – and found that all four subscales were significant predictors of Community Integration Questionnaire (CIQ) scores at 6-months post injury. However, only the FAM motor score was a significant predictor of CIQ score at 24-months post injury (Gurka, Felmingham, Baguley, Schotte, Crooks & Marosszeky, 1999).

Rasch analysis has been completed on the FIM+FAM by both Hawley et al., (1999) and Linn, Blair, Granger, Harper, O'Hara & Maciura (1999). Both groups found that the FIM™ and FAM have numerous items that do not fit with the model, and thus do not measure the underlying trait well. Linn and colleagues (1999) also found that the FAM items do not extend the difficulty of the measure beyond that of the FIM™ items, and do not ameliorate the ceiling effects found when using the FIM alone, as was the developmental purpose.

Both the FIM™ and the FAM separately have been shown to be sensitive to change over the first year of recovery in a TBI population (van Baalen et al., 2006) when tested at hospital discharge and again at 1 year post-injury. However, no easily identifiable studies were found that investigated the sensitivity of these measures at smaller intervals of time post-injury.

Evaluation: The FIM+FAM is one of the few measures that has had a complete formal psychometric analysis done, including Rasch analysis. These analyses show the FIM+FAM to be a highly reliable instrument for use in the brain injury population. However, it is important to point out that the expert training for the administration of the FIM+FAM is particularly difficult.

For a gross outcome measure, the FIM+FAM does appear to have good comprehensiveness, with items that inquire about self-care, mobility, communication, psychosocial adjustment and cognitive functioning. Yet, despite the relative breadth of concepts included, the FIM+FAM has been consistently shown to suffer from ceiling effects when used to assess brain injury survivors who have returned to the community (Gurka et al., 1999; Hall et al., 2001; Linn et al., 1999). This finding implies that, while effective for use with an inpatient population and during acute rehabilitation, the FIM+FAM are not appropriate for the assessment of higher functioning individuals, especially those who have returned to the community and are participating in post-acute rehabilitation.

Citations: A search of the Web of Science, PSYCHINFO and PubMed databases for the

seminal article for the FIM – Keith, Granger, Hamilton and Sherwin (1987) – unfortunately yielded no results as the article could not be found. However, a Google Scholar search discovered 304 citations of Keith et al. (1987). A Web of Science search for McPherson, Pentland, Cudmore and Prescott (1996) – the original article for the FAM – uncovered 29 citations of this article.

Community Integration Questionnaire (CIQ)

Description: The CIQ (Willer, Rosenthal, Kreutzer, Gordon & Rempel, 1993) was developed to measure community integration following brain injury, and to assess the level to which community integration is less-than-normal for age-, sex-, and culture-appropriate roles (Sander, Fuchs, High, Hall, Kreutzer & Rosenthal, 1999). The measure consists of 15 items comprising three subscales: Home Competency, Social Integration, and Productive Activity. Each scale contains a different number of items and a different range of scores, with a Total CIQ score that can range from 0 to 29, with higher scores representing greater community integration and independence. Two versions of the CIQ are available, one that is completed by the brain injury survivor and one can be completed by a significant other or other close friend or family member. The CIQ can be self administered, or administered like an interview, with a trained staff member reading the questions; it can also be completed via telephone if required (Sander, et al., 1997).

Psychometric Validation: Factor analysis on the CIQ has shown it to have good internal validity. A large study including data from 312 brain injury survivors done by Sander et al. (1999) reaffirmed the validity of the three original factors proposed by Willer et al. (1993) – Home Competency, Social Integration, and Productive Activity – with only a few item exceptions.

The CIQ has also been shown to have good test-retest reliability. Cusick and colleagues (2000) had 204 brain injury survivors at least 6-months post injury complete the CIQ on two occasions approximately two-weeks apart. The CIQ total had an inter-class correlation (ICC) of 0.86, and the subscales had ICC's ranging from 0.66 – 0.88 (Cusick, Gerhart & Mellick, 2000).

The CIQ has also been demonstrated to have good survivor-proxy agreement. Analysis of the individual items at 1-year post injury (N = 122) revealed kappa values ranging from 0.42 to 0.94 (Sander et al., 1997) and the ICC for the CIQ total between survivors and proxies was found to be 0.78 (Cusick et al., 2000).

The CIQ generally shows moderate concurrent validity when compared to integration-focused subscales of other outcome measures. Sander and colleagues (1999) investigated the concurrent validity of the CIQ compared to appropriate subscales of both the FIM+FAM (FIM Social Interaction, FAM Community Access, and FAM Employability) and the DRS (DRS Level of Functioning and Employability). Each CIQ scale score showed significant, yet weak-to-moderate, relationships to the FIM+FAM and DRS items, with Spearman ρ correlation coefficients that ranged from 0.24 to 0.60. The CIQ Total Score had higher correlations with each FIM+FAM and DRS subscale than each individual CIQ scale, with a range of 0.34 – 0.60 (most in the moderate range of 0.40 – 0.60). The CIQ has also been found to be significantly correlated with the CHART, with a Pearson correlation coefficient of 0.68 (Zhang, Abreu, Gonzales, Seale, Masel & Ottenbacher, 2002).

The CIQ appears to have a good range of item difficulty and thus good sensitivity to different levels of outcome. In a large sample size (N = 312) it was found that the CIQ does not show substantial floor or ceiling effects at 1-year post injury, indicating good sensitivity to the impairments experienced following brain injury (Sander et al., 1999). However, the CIQ's sensitivity to change has not been investigated.

Evaluation: The CIQ has many advantages to its use in assessment of outcome following brain injury. It is one of the only measures to focus specifically on community integration, an area that is commonly affected by the deficits following brain injury. Also, it is one of the first measures to have a self-administered survivor and significant other form, therefore taking into account not only the survivor's prospective on his/her re-integration, but also that of his/her significant other – these are the two people who will have had the most experience with the survivor's attempts at re-integrating into the community.

Since the CIQ does focus solely on community integration, it is limited in scope and there are many other aspects of recovery following brain injury that are not investigated using this measure alone. For this reason, the CIQ is not independently sufficient to investigate the range of aspects involved in functional outcome.

Citations: A Web of Science search for Willer, Ottenbacher and Coad (1994) – a follow-up to the original CIQ article – found 109 citations. Unfortunately, the seminal article was not found in Web of Science. A Google Scholar search for the seminal article – Willer, Rosenthal, Kreutzer, Gordon and Rempel (1993) – yielded 150 citations of this article.

Mayo-Portland Adaptability Inventory (MPAI-4)

Description: The Portland Adaptability Inventory was originally developed by Lezak in 1987 in an attempt to document and measure the behavioural and social problems that commonly arise following brain injury (Malec et al., 2003). In 1994, Malec and Thompson built on the original measure and developed the first version of the Mayo-Portland Adaptability Inventory which included all of the original items from the PAI, and additional items for rating pain and specific areas of cognitive impairment (Malec, Moessner, Kragness & Lezak, 2000). Malec and Thompson also adjusted the rating scales used, so that the MPAI is based on the patient's current functioning without reference to pre-injury abilities, and also standardized the scale across items so that a rating of 0 indicates no impairment and a rating of 3 indicates complete or near complete loss of function in the area in question (Malec & Thompson, 1994). The most current version, the MPAI-4, is comprised of 35 items (6 of which inquire about pre-existing and associated conditions, but do not contribute to the total score) grouped into 3 subscales (Abilities, Adjustment and Participation subscales). The measure can be completed by a staff member who is familiar with the brain injury survivor, the survivor him/herself or a significant other.

Psychometric Validation: Analyses done on the MPAI-4 found good internal reliability. The subscales show good internal consistency with alpha coefficients ranging from 0.76 to 0.83 (Malec et al., 2003). In addition, the MPAI-4 has been shown to have similar levels of Person and Item Reliability and Separation when completed by ABI survivors (N = 111) and their significant others (N = 100) (Malec, 2004).

The MPAI-4 was also found to have good inter-rater reliability and survivor-proxy agreement. Malec (2004) investigated the agreement between ratings made by survivors, significant others, and staff. A substantial amount of agreement on individual items was found among the three raters, with 42% of pairs of ratings having exact agreement (Malec, 2004).

The original version of the MPAI has been shown to have good concurrent validity with other outcome measures as well as cognitive tests (Malec & Thompson, 1994). The MPAI scores have a strong Spearman correlation with the DRS ($\rho = 0.81$), and a moderate correlation in the expected direction with the Rivermead Behavioural Memory Test ($\rho = -0.47$) (Malec and Thompson, 1994). The MPAI total score has also been found to be moderately correlated with a number of cognitive neuropsychological measures, however when only the cognitive items of the MPAI are compared to these measures, the correlations are generally greater that 0.50 (Malec & Thompson, 1994).

Rasch analysis done on an earlier, 30-item version of the MPAI yielded satisfactory results. In a sample of 305 brain injury survivors attending outpatient rehabilitation, the person separation was found to be 1.90 with a reliability of 0.78 and the item separation was 9.54 with a reliability of 0.99 (Malec et al., 2000). Analysis done on the MPAI-4 found that acceptable person and item reliability and separation were found when the measure was completed by survivors, significant others, and rehabilitation staff: results show full scale person reliability between 0.86 and 0.92, item reliability between 0.94 and 0.98, person separation between 2.94 and 3.37, and item separation between 3.84 and 6.81 (Malec, 2004). Furthermore, infit and outfit statistics for all three raters were found to be at acceptable levels for almost all items (Malec, 2004).

The MPAI has also been shown to be sensitive to different levels of impairment: individuals who were found to be more disabled based on their score on the Rancho Los Amigos Levels of Cognitive Functioning Scale (Rancho) showed lower median MPAI scores than those who were deemed less disabled via the Rancho (Malec & Thompson, 1994).

Evaluation: While the MPAI has not had a complete traditional psychometric analysis, thorough Rasch analysis on the final versions have shown it to be a reliable measure in assessing outcome following brain injury.

Two of the major strengths of the MPAI are that it has survivor, significant other, and clinician versions available, and it is more comprehensive than the other available outcome measures. However, little psychometric validation has been done with the survivor and significant other versions of the MPAI – most analysis has involved evaluations based on a team consensus of individuals working with or observing the survivor – and less investigation has been done with the MPAI-4 so it is unclear if the findings are applicable to the current version of the measure. Furthermore, although the MPAI covers many of the common problematic areas experienced after brain injury, the measure focuses exclusively on the frequency of the problem, and does not inquire about the impact that the problem has on the lives of the survivor and the significant other.

Citations: The original MPAI article by Malec and Thompson (1994) could not be found in most databases including Web of Science, PSYCHINFO and PubMed. However, a Google Scholar search of this article found it to be sited 34 times in their database.

European Brain Injury Questionnaire (EBIQ)

Description: The EBIQ (Teasdale et al., 1997) was developed specifically for use with brain injury survivors and their significant others to address a wide range of difficulties common to these populations. It has two versions - one for the survivor and one for the significant other. Both versions have 63 questions that address the problems experienced by the survivor as a result of the brain injury, and 3 questions that focus on the impact of the brain injury on the significant other. The 63 survivor-focused questions are divided into 9 scales: somatic, cognitive, motivation, impulsivity, depression, isolation, physical, communication, and core symptoms. All 66 questions are answered by both the survivor and their significant other from their personal perspective with the use of a 3-point scale regarding the frequency of the problem within the last month where a score of 1 represents “not at all”, 2 represents “a little”, and 3 represents “a lot” (Teasdale et al., 1997). Both versions are administered as a paper-and-pencil measure and are completed by the survivor and the significant other independently. (A .pdf of the survivor and significant other versions of the EBIQ is available at http://www.psy.ku.dk/teasdale/EBIQ.pdf).

Psychometric Validation: As a newer measure, limited psychometric analysis has been done on the EBIQ. However, internal reliability for the nine scales, measured using Cronbach’s alpha, were found to range 0.47 to 0.90 for survivor self-reports, and 0.54 to 0.92 for the relative-reports (Teasdale et al., 1997).

An investigation of the test-retest reliability of the EBIQ showed it to be satisfactory over a relatively short period of time (Sopena, Dewar, Nannery, Teasdale, & Wilson, 2007). The Pearson correlations for the time 1 and time 2 subscale scores ranged from 0.61 to 0.88 for a group of brain injury survivors. Similar test-retest correlations were found for relatives of brain injury survivors (Sopena et al., 2007). It has been found, however, that the cognitive subscale may be less reliable as a significant change was found from time 1 to time 2 (Sopena et al., 2007).

The EBIQ has also been shown to have moderate to good survivor-proxy agreement. Pairwise correlation coefficients ranging from .44 to.58 were found for the ratings provided by survivors (n=489), close relatives (n=465) and clinicians (n=230) (Deloche, Dellatolas & Christensen, 2000).

Furthermore, the EBIQ has been shown to be sensitive to the effects of brain injury, with a significant difference between brain injury survivors and a neurologically intact control group on 45 of 63 questions,[i] 40 of which were in the "positive" direction indicating that the brain injury survivors reported a greater degree of problems than the control group (Teasdale et al., 1997). Also, when tested both pre- and post-rehabilitation, it was found that EBIQ scores based on ratings from survivors and significant others were lower post-rehabilitation than they had been pre-rehabilitation (Svendsen, 2004).

Evaluation: The EBIQ has been put through the least psychometric analyses of the outcome measures considered here, and those which have been done show rather inconsistent results. Therefore, while the EBIQ is the most comprehensive of the outcome measures created to date and includes items that inquire about the impact of the brain injury on the life of the significant other, its utility as an outcome measure is unclear. With further investigation, the current author thinks that the EBIQ could be one of the more useful long-term outcome measures currently available.

Citations: Despite the relative recency of the development of the EBIQ, a Web of Science search of the seminal article by Teasdale et al. (1997) found that it had been cited in the database 36 times. By August 2007 the article had already been cited 5 times that year – as many times as it was cited in all of 2006. This may indicate increasing use of this measure in neuropsychological research.

Brain Injury Community Rehabilitation Outcome Scale (BICRO-39)

Description: The BICRO-39 (Powell, Beckers & Greenwood, 1998) was developed to assess the progress and outcome of community re-integration following brain injury. Thus, the measure can be used to assess the individual consequences of the brain injury (evaluate the survivor's level of functioning on each item before[ii] and after the injury) as well as changes that may occur over time by comparing scores from multiple administrations. The BICRO-39 can be completed by survivors and/or their significant others, and three different forms of the questionnaire have been developed: the patient pre-injury, which requires the survivor to retrospectively rate his or her functioning before the injury; the patient post-injury, which requires the survivor to rate his or her current functioning; and the carer post-injury, where the survivor's carer or significant other rates the survivor's current functioning. The BICRO-39 is made up of 39 questions that are scored on a 6-point scale (0 – 5) with higher scores indicating poorer community re-integration. Along with an overall score, 7 factor scores can also be derived (Personal Care, Psychological, Socializing, Self-Organization, Mobility, Family Contact, and Productive Employment) (Powell et al., 1998).

Psychometric Validation: Powell et al. (1998) investigated many aspects of the psychometric validity of the BICRO-39. The internal validity was found to vary greatly depending on the subscale measured. For example, the Personal Care, Mobility, Self-Organization, and Psychological subscales all had very strong alphas ranging from 0.88 to 0.95. Yet, the Socializing, Parent/Sibling Contact and Partner/Child Contact had moderate alphas (0.55 to 0.70) and the alpha for Productive Employment was weak (0.30). The authors propose that the low alpha coefficients in the latter subscales are likely due to one or more of the items included in the scale being inapplicable to the individual (e.g. he/she may have a partner but no children, or be fully occupied with one or two activities thus reducing the time available for others) which could result in item scores within the subscale being very different (Powell et al., 1998).

For most of the factors, the BICRO-39 was found to have good test-retest reliability. When completed at least 1 day and no more than 28-days apart, Spearman correlations for the nine subscales were found to be between 0.53 and 0.89 for the patient pre-injury form, between 0.67 and 0.92 on the patient post-injury form, and between 0.59 and 0.98 for the carer post-injury form (Powell et al., 1998).

The BICRO-39 was also found to have good survivor-proxy agreement. Spearman correlations for survivors and carers for the post-injury form ranged from 0.62 to 0.89, and the only factor-mean that was found to be significantly different between the two raters was for the Psychological subscale (Powell et al., 1998).

Good concurrent validity for the BICRO-39 was found when compared to scores on other outcome measures. For example, significant correlations were found between specific BICRO-39 factors and corresponding subscales of the FIM + FAM with a range from 0.49 to 0.76. The Psychological subscale of the BICRO-39 was also found to show a significant correlation with the Hospital Anxiety and Depression Scale (HADS)-Depression scale (r = 0.68) and the HADS-Anxiety scale (r = 0.81). Subscales of the BICRO-39 also show significant correlations to scales of the Community Integration Questionnaire (CIQ) ranging from -0.54 to -0.77 (Powell et al., 1998).

Finally, the BICRO-39 was shown to be sensitive to change. A group of 65 patients tested at intake and again after discharge from a rehabilitation program, with a mean interval between assessments of 46.6 +/- 46.1 weeks showed significant improvements on the Personal

Care, Mobility, and Psychological subscales with the Self-Organization subscale showing a trend towards improvement. When only those individuals who were deemed “effected” on a subscale at rehabilitation intake were investigated for change on those subscales at discharge, all subscales showed significant increases in scores except Parent/Sibling Contact (Powell et al., 1998).

Evaluation: The BICRO-39, as a fairly new outcome measure, is limited by a lack of psychometric data. Also, there has been no modern psychometric analysis, such as Rasch analysis, used to investigate the measure. In the analyses that have been done, the internal consistency scores for four of the subscale scales are fairly low, which suggests that grouping these items together on a common scale may be inappropriate. Further limitations of the BICRO-39 include the questionable nature of the pre-injury form based on previous research findings, and an almost complete focus on problem frequency. Aside from these limitations, the BICRO-39 may be a promising instrument for measuring brain injury outcome with moderate comprehensiveness.

Citations: A Web of Science database search for Powell, Beckers and Greenwood (1998) found that this article had been cited 27 times. However, the citation rate over the past 5 years is quite erratic.

Appendix B

The Functional Outcome Profile (FOP)

{Change point for Ver 3.04a: entire box} Subject ID:	_________
Date (MM/YY):	_____________
Interviewed by (Initials/Discipline):	______________
Most Responsible Therapist (Initials/Discipline):	_____________
Location (site, e.g. VGH):	_____________________________

Functional Outcome Profile

Interview

Client Version[1]

Version 3.04a

Version date: June 2004[2]

Michael Joschko & Ronald Skelton[3]

University of Victoria

---

[1] This version is the first designed for use in multiple clinical sites. Version "a" incorporates minor edits that have shown up as needed during clinical use. Last modified: June 4, 2004

[2] Printed 04/06/2009

[3] Major contributor to development: Susan Larke

FUNCTIONAL OUTCOME SCALES – LIST OF CONSTRUCTS

Physical, ADL

1. Senses
2. Gross Motor
3. Fine Motor
4. Self-Care
5. Eating & Meal Preparation
6. Domestic Skills
7. Fatigue
8. Sleeping

Health

9. Medical Conditions
10. Pain
11. Health Satisfaction
12. Pain Medications
13. Psychotropic Meds
14. Med Side Effects
15. Problems from Meds/Drugs

Cognitive

16. Attention, Sustained
17. Attention, Divided
18. Language Expression
19. Language Comprehension
20. Information Processing Speed
21. Memory, Anterograde
22. Memory, Retrograde
23. Procedural Learning
24. Disorientation, Time
25. Way finding

Executive

26. Judgement
27. Organization, Cognitive
28. Confusability
29. Impulsivity
30. Organization, Behavioural
31. Initiation
32. Time and Punctuality
33. Money Management
34. Self-Insight
35. Coping

Emotional, Behavioural

36. Emotional Reactivity
37. Problem Behaviour
38. Intrusions: Emotional and Cognitive
39. Mood
40. Self-Esteem
41. Spirituality

Social

42. Relationships, Family
43. Relationships, Satisfaction
44. Physical Intimacy Problems
45. Physical Intimacy Satisfaction
46. Social Life
47. Social Relationships
48. Community/Social Interactions
49. Community Barriers
50. Community Involvement

Activities

51. Self-Care Activities
52. Leisure: Solitary
53. Leisure: Active/Social
54. Leisure Satisfaction
55. Work, Quantity
56. Work Satisfaction
57. Work Demands
58. Safety concerns, by Others
59. Safety Concerns, by Self

Overall

60. Economic Security
61. Other Problems
62. Life Satisfaction
63. Good Things

FOP Instructions (To Be Read to Interviewee)

FOP Instructions (To Be Read to Interviewee)

During this part of the interview, I will be asking you questions about whether you are having problems in various areas in your life. Your answers will help me to get to know you, to figure out which parts of your life are going well, and which are areas in which you are still recovering. If we have any interviews in the future, your answers can be used to measure how things are changing over time.

Examples of some of the things we will be talking about include your memory, how you spend your time every day, and how things are going for you and your family. Sometimes you will be asked how often certain things happen, and also how much those things have affected you. Please try to give only your opinion when you are answering the questions, and please try to be as accurate as possible. Remember, there are no right or wrong answers. Some of the questions may have more than one part, so it's better to wait until I've finished the whole question before giving your answer.

SHOW FREQUENCY SCALE[iii]

We will use this scale to help you describe how often things happen or how often they cause you problems in your everyday life. ***A problem is something that causes you some difficulty, discomfort, embarrassment, or prevents you from doing something you want to do***. We will ask you to put your finger on the scale at the place you think best answers the question. Your choices range from **"NEVER", to "Once a month", "Once a week", "Once a day", "Once an hour",** and **"Constantly"** *(point and read out the options)*. So, if something happens once an hour you would put your finger here *(demonstrate)***,** or if something happens every couple of days, a few times a week, you would put your finger higher than **"Once a week"**, but lower than **"Once a day"**. Constantly means that it happens all the time, like breathing.

Do you have any questions before we go any further?

SHOW IMPACT SCALE

We will use this scale to help you describe how much certain things have affected or impacted your everyday life. Your choices range from **"No Impact"** to **"Couldn't be more!"** *(point and read out the option)*. **"No Impact"** means you would barely notice it, and it wouldn't bother you. **"Couldn't be more"** means that the thing could not have affected your life more than it has. This is not supposed to be how it affects you emotionally, but how if affects what you do.

SHOW SATISFACTION SCALE

Sometimes we will ask how satisfied you are with something in your life. This time, you will use your finger to show how happy or satisfied you are with something. Your choices range from "**Couldn't be less satisfied"** to **"Couldn't be more!"** *(point and read out the option)*. **"Couldn't be less satisfied"** means you are not at all happy or satisfied with that thing. **"Couldn't be more"** means that you could not be happier or more satisfied about that thing.

SHOW IMPORTANCE SCALE

Finally, we will sometimes ask you how important something is to you. Once again, you will use your finger to tell me whether something is of **"No Importance"** to you, or that it **"Couldn't be more!"** important.

I'll start the interview with a few sample questions to make sure that you understand how to use the scales, but before I do that, do you have any questions you'd like me to answer right now?

START HERE

Here are the sample questions:

A. {Frequency} How often do you eat? ***(It doesn't matter whether response is just for meals or includes snacks too).***

SHOW FREQUENCY SCALE **Point on the line to show that.** Frequency

Ensure the respondent provides a rating by pointing to the appropriate place on the line.

A response of "Three times a day" should be demonstrated on the scale if necessary.

B. {Impact} Now, I'd like you to imagine you no longer have a TV.

How much would this affect your daily life?

Impact

SHOW IMPACT SCALE **Point on the line to show that.**

Ensure the respondent provides a rating. Clarify if necessary (i.e. blatant error only).

C. {Satisfaction, Importance} Think of something you are reasonably happy with right now. Anything you like, even the clothes you are wearing or the weather today.

Satisfaction

What is it? ___________ Describe how happy or satisfied you are with it.

SHOW SATISFACTION SCALE **Point on the line to show that.**

Clarify if necessary.

How important is ____________ to you?

SHOW IMPORTANCE SCALE **Point on the line to show that.**

Importance

Clarify if necessary.

Do you have any questions about the rating scales?

Now, we are going to move on to the rest of the questions. If you do not understand a question, or a word in the question, please ask me to explain it to you. Also, these questions were designed to describe many different types of people in many different situations. I do not expect all of the questions to apply to you, but please answer every question as best as you can. Occasionally I will be taking note of what you say so that I can remember it better.

QUESTIONS

I'd like to start by asking you about your physical abilities, like your five senses, and your ability to use your arms, legs, and hands.

1. {Senses} **Do you have difficulty with <u>any of your senses</u> - like vision, hearing, smell, taste or touch?**	**Y N N/R**

Like what?

____

(Anything else?)

____

Taking all these into account, how often have these difficulties caused you problems in your everyday life?

Frequency

Impact

Over the past month, how much has this affected your everyday life?

Which difficulty would you say concerns you most?

____

2. {Gross Motor} **Do you have difficulty walking or <u>doing things with your arms or legs</u>? This could include everything from problems with paralysis, weakness, balance or tremors, to difficulty riding a bike, walking along a narrow surface like a curb, or catching a ball.**	**Y N N/R**

Over the past month, how often has your difficulty doing things with your arms or legs caused you problems in your everyday life?

Frequency

Impact

Over the past month, how much have these problems affected your everyday life?

3. {Fine Motor} **Do you have difficulty <u>making fine movements with your hands</u>? For example, in tasks such as writing, sewing or fixing things.**	**Y N N/R**

Over the past month, how often has your difficulty making fine movements with your hands caused you problems in your everyday life?

Frequency

Impact

Over the past month, how much have these problems affected your everyday life?

Now I'm interested in finding out how independent you are. In this context I'll be asking about things like doing household chores, preparing meals, etc..

4. {Self-Care} **Do you have any difficulty with <u>self-care</u>? This includes everything from using the bathroom, brushing your teeth, bathing, and dressing yourself, to taking care of yourself when you are hurt, such as when you have a cut or a burn.**	**Y N** **N/R**

Over the past month, how often have you had difficulty with self-care?

Frequency

Impact

Over the past month, how much have these problems affected your everyday life?

5. {Eating and Meal Preparation} **Do you have difficulty <u>with eating or meal preparation</u>?** **This doesn't include problems with your appetite but does include difficulty eating, cooking, or taking care of leftovers.**	**Y N N/R**

Over the past month, how often have you had difficulty with eating or meal preparation?

Frequency

Impact

Over the past month, how much have these problems affected your everyday life?

6. {Domestic Skills} **Do you have difficulty <u>doing chores around the house</u>? (This includes doing the dishes, making your bed, cleaning the bathroom or doing simple maintenance chores like replacing a light bulb, or cleaning the fridge.)**	**Y N** **N/R**

So, over the past month, how often have you had difficulty doing household chores?

Frequency

Impact

Over the past month, how much have these problems affected your everyday life?

Now, I'd like to ask about your energy levels and sleeping patterns.

7. {Fatigue} **Do you put off activities because you are <u>physically or mentally tired</u>, or find you can't finish what you have started?**	**Y N N/R**

Over the past month, how often has your being tired caused problems for you?

Frequency
Impact

Over the past month, how much has this affected your everyday life?

8. {Sleeping} **Would you say that you have been having problems <u>sleeping</u>? This could include waking up too early, or sleeping too much, having nightmares, sleep walking, or tossing and turning a lot.**	**Y N N/R**

Record sleep problem

____

____

_

Frequency

Over the past month how often have these problems affected your everyday life?

So overall, how satisfied have you been with your sleeping over the past month?

Satisfaction

How much time do you spend napping? ___________Hrs per day/week *(Circle one)*

How much time do you spend sleeping or trying to sleep each night? _________Hrs per day

Total the number of hours sleeping, trying to sleep, and napping ________

Record hours on table on pg. 28

For this next part I will be asking you about your health, medical conditions and pain you may have been having.

9. {Medical Condition} **Do you have any <u>medical conditions</u> other than your brain injury that affect your everyday life? For example, anything you might see a doctor for, like seizures, asthma, allergies, or heart problems.**	**Y N N/R**

Medical conditions could include sequelae to brain injury, such as headaches.

Like what?

____

(Anything else?) ____

Taking into account all these difficulties with your medical conditions, how often have these difficulties caused you problems in your everyday life?

☐ Frequency

☐ Impact

Over the past month, how much has this affected your everyday life?

Which condition concerns you the most? ____

Did you have this condition before your injury? **Y** **N**

10. {Pain} **Over the past month, have you been experiencing <u>physical pain</u>?**	**Y** **N** **N/R**

Over the past month, how often has your pain caused you problems (in your everyday lif

☐ Frequency

☐ Impact

Over the past month, how much has this affected your everyday life?

{Health Satisfaction} Overall, how happy or <u>satisfied</u> with your <u>health</u> have you been over the past month?

☐ Satisfaction

The next few questions are about <u>medicines</u> that you might be taking. This could be medication you or your family get at a drug store, like something for a cold, and also medication prescribed by a doctor. I will be asking you about your use of a few different kinds of medications. Later on, I will ask you some general questions about your use of recreational drugs or alcohol.

Over the past month, did you take any medicine of any kind?	**Y N**

Data entry instruction: If “No” here, enter N, N, and

*If no Meds, skip to ***, after question 14*

11. {Pain Medications} **Do you take <u>medication</u> for headaches or other kinds of <u>pain</u>?**	**Y N N/R**

What kind?

____

Over the past month, how often have you taken medication for pain? ☐ Frequency

Over the past month, how important has your pain medication been to you? ☐ Importance

Rate taking all medication into account

12. {Psychotropic Meds} **Do you use any <u>drugs</u> prescribed by a doctor to help you relax, or to change your <u>mood</u> or behaviour?**	**Y N N/R**

What kind?

____

Over the past month, how often have you taken this (these) medication(s)? ☐ Frequency

Over the past month, how important has (have) this (these) medication(s) been to you? That is, how much do you think you need this medication? ☐ Importance

Rate taking all medication into account

13. {Med Side Effects} **Thinking about all the <u>medications</u> prescribed by your doctor (including any medications we haven't talked about), have you been <u>unhappy</u> with the <u>effects or side-effects</u> of any of them?**	**Y N N/R** **N/App**

Over the past month, how often have these effects caused you problems (in your everyday life)? ☐ Frequency

Impact

Over the past month, how much has this affected your everyday life?

****If no Meds, start again here.*

Now I'd like to ask you about your use of drugs and alcohol. I won't ask for any details about what you take but I will ask whether drug or alcohol use causes you any problems.

If interviewee is not taking medication (i.e., questuions12, 13, and 14 skipped), only ask drugs

Do you use **(recreational/street)** **drugs or alcohol,** **(stop or pause)** **or do you take more of a prescribed medication than your doctor told you to?** *If No or N/R, skip to question 16*	**Y N N/R**

Data entry instruction: answer is not recorded

Now I am going to ask you if drug or alcohol use is a problem for you. The sorts of things I am thinking about include missing work because of it. Or having family and friends complain about it, or suggest that you cut down. Or getting in trouble because of your use of alcohol, drugs or medication.

14. {Problems from Meds/Drugs} **So, over the past month, has your use of <u>drugs, alcohol</u> or over-use of <u>medications</u> caused you <u>problems</u> (in your everyday life)?**	**Y N N/R**

How often?

☐ Frequency

Over the past month, how much have these problems affected your everyday life?

☐ Impact

For this next part, I am going to ask you some questions about your ability to express yourself, understand others, pay attention, and concentrate in your day-to-day life.

15. {Attention, Sustained} **When you do things you enjoy, do you have difficulty <u>staying with it</u>, or getting done what you wanted?**	**Y N N/R**

Over the past month, how have often has this difficulty staying with things caused you problems (in your everyday life)?

☐ Frequency

☐ Impact

Over the past month, how much has this affected your everyday life?

16. {Attention, Divided} **Do you have difficulty <u>doing two things at once,</u> for example, talking and cooking at the same time?** **(For example, listening on the phone and taking down a message?)**	**Y N N/R**

Over the past month, how often has this difficulty caused you problems (in your everyday life)?

Frequency

Impact

Over the past month, how much has this difficulty doing two things at once affected your everyday life?

17. {Language Expression} **Do you have difficulty <u>expressing yourself by speaking or in writing</u>? This includes difficulty finding the right word, or not being able to put your thoughts into words, or not being able to get your point across.**	**Y N** **N/R**

Over the past month, how often has your difficulty expressing yourself caused you problems (in your everyday life)?

Frequency

Impact

Over the past month, how much have these problems affected your everyday life?

18. {Language Comprehension} **Do you have difficulty <u>understanding what most people are saying or what you are reading</u>? (For example, do you have difficulty understanding what you are asked to do, or following the news on TV, radio or in the paper?)**	**Y N** **N/R**

Over the past month, how often has your difficulty understanding what people are saying or what you read caused you problems (in your everyday life)?

☐ Frequency

☐ Impact

Over the past month, how much have these problems affected your everyday life?

19. {Information Processing Speed} **In terms of your <u>thinking</u>, do you have difficulty <u>keeping up with what's going on</u>? For example, do conversations or TV shows move too quickly for you to follow?**	**Y N N/R**

Over the past month, how often has this difficulty keeping up with what's going on caused you problems (in your everyday life)?

Frequency

Impact

Over the past month, how much has this [difficulty keeping up] affected your everyday life?

Now I am going to ask you some questions about your memory, and your ability to learn new things.

20. {Memory, Anterograde} **Do you have difficulty <u>remembering things</u>, like where you left things, or appointments you have, or the names of new people you meet?**	**Y N N/R**

Over the past month, how often has your difficulty remembering these kinds of things caused you problems (in your everyday life)?

Frequency

Impact

Over the past month, how much has this affected your everyday life?

21. {Memory, Retrograde} **Do you have difficulty <u>remembering old information</u>, like the names of people you know well, or things that happened in the past?**	**Y N N/R**

Over the past month, how often has your difficulty remembering [old information] caused you problems (in your everyday life)?

Frequency

Impact

Over the past month, how much has this affected your everyday life?

22. {Procedural Learning} **Do you have difficulty <u>learning how to do new things</u>, like a new task at work (at school) or how to use a new alarm clock (in a hotel), or how to use someone else's computer?**	**Y N N/R**

Frequency

Over the past month, how often has this [sort of thing] caused you problems (in your everyday life)?

Impact

Over the past month, how much has this affected your everyday life?

23. {Disorientation, Time} **Do you have difficulty keeping track of "when'" it is? For example, do you lose track of what day of the week it is, or whether it's morning or afternoon?**	**Y N N/R**

Over the past month, how often has this difficulty keeping track of time caused you problems (in your everyday life)?

Frequency
Impact

Over the past month, how much has this affected your everyday life?

24. {Wayfinding} **Do you have difficulty finding your way to where you are going? For example, to a shopping mall or another place you are familiar with?**	**Y N N/R**

Over the past month, how often has your difficulty with these kinds of things caused you problems (in your everyday life)?

Frequency
Impact

Over the past month, how much has this difficulty finding your way around affected your everyday life?

For this next section, I am going to ask you about your judgement and thinking.

25. {Judgement} **Do you sometimes <u>make mistakes in judgement</u>, such as making choices that turn out badly**? **For example, not showing common sense?**	**Y N N/R**

Over the past month, how often have your mistakes in judgement caused you problems (in your everyday life)?

Frequency

Impact

Over the past month, how much has this affected your everyday life?

26. **{Organization, Cognitive}** Do your <u>thoughts seem scattered</u> or disorganized? For example, do your thoughts jump around or do you have difficulty following an idea through to its completion?	**Y N N/R**

Over the past month, how often has your difficulty with your thoughts caused you problems (in your everyday life)?

Frequency

Impact

Over the past month, how much has this affected your everyday life?

27. {Confusability} **Do you get confused, upset, or thrown off track by unexpected things that happen? For example, if the furnace breaks down, or you can't find a parking space?**	**Y N N/R**

Over the past month, how often has this difficulty caused you problems (in your everyday life)?

Frequency

Impact

Over the past month, how much has this affected your everyday life?

28. {Impulsivity} **Do you do things impulsively, without thinking? For example, do you do the first thing that comes to mind without thinking about what might happen?**	**Y N N/R**

Over the past month, how often has your impulsiveness caused you problems (in your everyday life)?

Frequency

Impact

Over the past month, how much has this affected your everyday life?

In the next part of the interview, I will be asking about some of your day-to-day abilities, such as organizing your time, and managing your money.

29. {Organization, Behavioural} **Do you have difficulty <u>organizing your activities</u> for the day?** **For example, do you plan to do too much, or have difficulty getting done what you wanted to do?**	**Y N N/R**

Over the past month, how often has this difficulty caused you problems (in your everyday life)?

☐ Frequency

☐ Impact

Over the past month, how much has your difficulty with organization affected your everyday life?

30. {Initiation} **Do you have difficulty <u>starting things you need to do</u>, such as getting up out of bed and dressed for the day, cooking or doing other daily chores?**	**Y N N/R**

Over the past month, how often has this difficulty starting things caused you problems (in your everyday life)?

☐ Frequency

☐ Impact

Over the past month, how much has this [difficulty in starting things] affected your everyday life?

31. {Time and Punctuality} **Do you have difficulty keeping track of time and <u>being on time</u>? For example, do you have difficulty being on time for appointments?** ***(includes missing appointments entirely)***	**Y N N/R**

Over the past month, how often has your difficulty being on time caused you problems (in your everyday life)?

Frequency
Impact

Over the past month, how much have these problems affected your everyday life?

32. {Money Management} **Now I'm going to ask you whether you have difficulty managing money. This includes making choices spending money, being able to make purchases, dealing with the bank, paying bills, or setting priorities. Do you have difficulty <u>managing your money</u>?**	**Y N N/R**

Frequency

Over the past month, how often has your difficulty managing money caused you problems (in your everyday life)?

Impact

Over the past month, how much have these problems affected your everyday life?

33. {Self-Insight} **Do you think you ever over- or under-estimate your abilities?** **Y N N/R**

Data entry instruction: enter under / over in Y/N box

Note: Indicate response by circling appropriate descriptor (under / over).

Over the past month how often has this over/under estimation caused problems for you?

Frequency

Impact

How much impact has this had on your everyday life?

Now, I'd like to ask you about your emotional reactions and your ability to cope with stresses. Everybody has stress in their life. Stresses people experience usually change from day to day or from week to week.

34. {Coping} **Do you sometimes find that it is hard to cope with the things that are happening in your life?**	**Y N N/R**

Over the past month, how often has this difficulty coping caused you problems (in your everyday life)?	Frequency
Over the past month, how much did this affect your everyday life?	Impact

35. {Emotional Reactivity} **Do you sometimes have strong emotional reactions or difficulty controlling your emotions? For example, do you have difficulty with irritability, depression, intense anger, bad temper, or tearfulness?**	**Y N N/R**

Could you give me some examples?

____

(Anything else?)

____

Taking all of these things together how often have your emotional reactions caused you problems (in your everyday life)?

Frequency

(Or if only one emotion cited) How often has this caused you problems in your everyday life?

Impact

Over the past month, how much have these reactions affected your everyday life?

Now I want to ask about some behaviours that might cause problems in getting along with others, perhaps because they are odd or annoying. These behaviours can be done out of anger, impulsively, on purpose, or just because you can't help yourself.

36. {Problem Behaviour} **Do you ever <u>behave</u> in ways that cause you <u>problems</u> getting along with others? This includes things like belching, swearing, not cleaning up after yourself, physically hurting yourself or other people or doing things over and over again because you just can't help yourself.** **So, have you done anything like this over the past month?**	**Y N N/R**

Like what? ____

____

Accept any behaviour that would interfere with getting along with people, including unusual or repetitive habits

Over the past month, how often has this (have these) behaviour(s) caused you problems getting along with others? *Rate taking all behaviours into account*

Frequency

Impact

Over the past month, how much have these problems affected your everyday life?

Now, I'd like to ask you about your emotional experiences, mood, and how you feel about yourself.

37. {Intrusions: Emotional and Cognitive} **Do you sometimes have <u>emotional experiences or thoughts that are troubling to you</u>? For example, periods of anxiety, bad dreams, flashbacks, thoughts of suicide, unusual thoughts or fears, or thoughts that won't go away?**	**Y N N/R**

Could you give me some examples?

____

____

(Anything else?)

____

Over the past month, how often have you had these kinds of experiences?	Frequency
Rate taking all experiences into account	
Over the past month, how much has this affected your everyday life?	Impact

{Mood} Overall, how would you describe your <u>mood</u>?

____

Data entry instruction: enter description with Q. 30

If the client does not understand the term "Mood," read the following description: "Mood is a way of describing the general state of your feelings over time. Some examples of moods are: happy, sad, satisfied, irritable, relaxed, and anxious."

38. {Mood} Over the past month, has your <u>mood</u> caused problems for you?	**Y N N/R**

How often (has your mood caused you problems in your everyday life)? Frequency

So, how much your mood affected your everyday life? Impact

Over the past month, how happy or satisfied have you been with your mood?

Satisfaction

39. {Self-Esteem} **Now I'd like to ask you about your <u>self-esteem</u>. This would include your confidence in doing things, your feelings of self-worth, whether you accept yourself for who you are, and how satisfied you are with the amount of control you have over your own life.**

So, over the past month, how satisfied have you been with who you are? Satisfaction

Note: Try to keep person focussed on their inner being, not their physical form

I'd like to ask you about your spiritual well-being. For some people this relates to their religion and values, their belief in God, or their relationship to a god. For others, it relates to their sense of who they are, their acceptance of themselves, an inner peace, or their sense of the meaning and value of their life.

40. {Spirituality} **Would you say that your spiritual life or <u>spiritual well-being</u> is important to you?** ***(is a concern of yours?)***	Y N N/R

SAT

How important is your spiritual well-being to you?

Importance

Over the past month, how satisfied have you been with your spiritual life or spiritual well-being?

Satisfaction

In this next section I will be asking you about your relationships with people you love, including family, friends, and others who are close to you.

41. {Relationships, Family} **Over the past month, have you been having problems getting along with family members or people you are close to? (E.g. boy/girlfriend)**	**Y**	**N**	**N/R**

Frequency

Over the past month, how often have you been having problems getting along with people [you are close to]?

Impact

Over the past month, how much has this affected your everyday life?

42. {Relationships, Satisfaction} **So, taking into account all of your relationships with family and friends, including your close and loving ones, how satisfied have you been with these relationships over the past month?**

Satisfaction

Importance

How important are these relationships to you?

43. {Physical Intimacy} **Over the past month, have you been having difficulty with physical intimacy? This includes touching, giving and receiving hugs, and sexual relationships.**	**Y**	**N**	**N/R**

Frequency

Over the past month, how often has this been a problem for you?

Impact

Over the past month, how much has this [difficulty] affected your everyday life?

44. {Physical Intimacy Satisfaction} **Over the past month, how satisfied have you been with your <u>ability to be physically intimate</u>?**

satisfaction

Importance

How important to you is your ability to be physically intimate?

Now I would like to ask you about your social life, your community involvement, and your ability to get around town.

45. {Social Life} Now, talking about your <u>social life,</u> I mean this in the broadest sense - not just parties and social get-togethers, but whether you see friends as much as you'd like, or whether you are feeling lonely or isolated.

Satisfaction

So, over the past month, how happy or satisfied with your social life have you been?

How important is your social life to you?

☐ Importance

46. {Social Relationships} **Do you have difficulty <u>making or keeping friends</u>? For example, do you have difficulty keeping in touch with them or making plans to do things together?**	**Y N** **N/R**

☐ Frequency

Over the past month, how often has your difficulty making or keeping friends caused you problems (in your everyday life)?

☐ Impact

Over the past month, how much have these problems affected your everyday life?

47. {Community/ Social Interactions} **Do you have difficulty interacting with people in the community, like bus drivers, sales people, or cashiers? For example,**	**Y N N/R**

are you uncomfortable approaching such people, or do you have difficulty understanding them or making yourself understood? **(i.e., difficulty in everyday social interactions)**	

Over the past month, how often have you had difficulty dealing with people in the community?

Frequency

Impact

Over the past month, how much has this affected your everyday life?

48. {Community Barriers} **Are there any things in your home, work, school, or community that prevent you from doing what you want to do? For example, services you need that aren't available, or places you can't get to?**	**Y N N/R**

Like what?

____

(Anything else?)

____

Frequency

Over the past month, how often have [all these things] gotten in your way?

Over the past month, how much has this affected your everyday life? Impact

Now I'd like to ask about your current involvement with your community. There are different ways you can get involved, for example, by being a member of a club, sports team, choir, or church, or volunteering at a brain injury group, senior centre, or other organization. However, there are many people who aren't interested in such things.

50. {Community Involvement} **Does community involvement matter to you?**	**Y N N/R**

SAT

How important is community involvement to you? Important

How satisfied are you with your current level of involvement? Satisfaction

Now, I'd like to ask you about the different ways you spend your time. Up until now, I have been asking you about how things have been for the past month. Now, I'm going to ask you about the things you do in a "typical" day or week.

If the client is having difficulty imagining a "typical week", the following may be given as clarification: (In other words, think about a week in the past month where nothing really unusual happened such as getting ill or having house-guests.)

51. {Self-Care Activities} **Now I'd like to ask you about your personal <u>self-care activities</u>, the things that you do to look after yourself. This includes all the things you do in the bathroom and other things like dressing, eating and daily physio exercises, but not chores like cooking or laundry. So, in a typical day, how much time do you spend doing personal self-care activities?** ***(Reminder: Question 4 asked about self-care problems, not activity amounts.)***

Detach and use table on pg. 28 to record information. Collect self-care activities all at once or divide the time into Toileting, Eating, and Other, whichever is easier for the client to answer. Exercise can go here or under "Leisure". Note: Include Rehab/Therapy time with a professional under "Work".

After obtaining a list of activities, discuss the time spent on each in a typical week (Enter as hours per day) and the amount of help required (to nearest 25% is fine). Include travel hours and travel help.

One way people spend their time is doing leisure or recreational activities, like hobbies, working out, or watching TV. Some of this time is spent alone, some is spent with other people.

52. {Leisure: Solitary} How much <u>time</u> do you spend doing leisure activities <u>alone</u>? For example, how much time do you spend watching TV, playing video games, or being on the computer all by yourself *(That is, non-socially interactive activity, and not chatting online)*

_________ Hours day/ week

Note: online chatting and e-mail correspondence should be coded under "leisure" (i.e., social activity) unless it is work related. If the response is vague (e.g., "a lot of TV"; "many hours"), query to obtain a specific number of hours. Use the table below to convert hrs/week to hrs/day.

53. {Leisure: Active/Social} **Other than those things we just talked about, what are some of the different sports and <u>leisure activities</u> you do every week? For example, swimming, shopping, doing stuff with friends, talking on the phone or Internet, watching TV or playing video games with friends.** *(Record in Activity Column of table)*

After obtaining a list of activities, discuss the time spent on each in a typical week (Enter as "hours/week" or "Hrs/day & Days/wk") and the amount help required for each activity (to nearest 25% is fine). Include travel times and travel help.

54. {Leisure Satisfaction} **Now taking all of these things into account, how <u>satisfied</u> have you been with your <u>leisure activities</u> over the past month?**

☐ Satisfaction

☐ Importance

Over the past month, how important have your leisure activities been to you?

Many people also spend time doing different kinds of work. Sometimes it's paid work and sometimes it's volunteering, or taking care of children, going to school or a training program, or doing housework and other chores.

55. {Work, Quantity} **What are some of the different kinds of work you do in a typical week?**

(Record in Work Column of table)

After obtaining a list of activities, discuss the time spent on each in a typical week (Enter as "hours/week" or "Hrs/day & Days/wk") and the amount help required for each activity (to nearest 25% is fine). Include travel times and travel help.

56. {Work Satisfaction} **Taking all these things into account over the past month, how satisfied have you been with the amount and types of work you have been able to do? For example, would you like to be able to do a different kind of job or accomplish more around the home?**

Satisfaction

Importance

How important was your ability to work over the past month?

57. {Work Demands} **Do you have difficulty dealing with the demands of a work (school) setting? For example, this could include difficulty getting along with your boss (teachers) or co-workers (fellow students), or working under pressure or working for an hour without getting distracted, or getting upset when someone criticises your work.** ***(“Work setting” refers to any structured environment demanding work and social skills. Clients need not be working in a paid or traditional work setting in order for difficulties in these areas to cause problems in their lives.)***	**Y N N/R**

Over the past month, how often has your difficulty dealing with the demands of a work setting caused you problems (in your everyday life)?

Frequency

Impact

Over the past month, how much have these problems affected your everyday life?

For the next two questions, I will be asking how you and others feel about your safety when you are alone.

58. {Safety Concerns, by Others} **Do people do things that show they are concerned about**	**Y N N/R**

your <u>physical or personal safety</u> when you are alone? **For example, because you might do something dangerous without thinking, or because you might be taken advantage of?**	**Never Alone**

Over the past month, how often? Frequency

Over the past month, how much has this affected your everyday life? Impact

59. {Safety Concerns, by Self} **Are <u>you</u> ever concerned about your <u>physical or personal</u> safety when you are alone?**	**Y N N/R** Never Alone

Over the past month, how often? Frequency

Over the past month, how much has this affected your everyday life? Impact

Now I'd like to ask you about your current financial situation and whether your injury has caused you financial hardship, or whether you are worried about where money will be coming from in the future.

60. {Economic Security} **So, over the past month, how satisfied have you been with your <u>financial situation</u> and your future prospects?**

FREQ

Satisfaction

Frequency

Over the past month, how often has your financial situation caused problems for you?

For the last three questions I would like you to think about the "big picture". I'd like you to consider everything going on in your life.

61. {Other Problems} **Are there any *other* <u>problems</u> you've had over the past month that you think might be happening because of your injury?**	**Y N N/R**

Like what?

____

(Anything else?)

____

Which would you say concerns you most?

____

Over the past month, how often has this happened? ☐ Frequency

Over the past month, how much has this affected your everyday life? ☐ Impact

62. {Life Satisfaction} **When you think about all that has been happening in your life over the past month, how <u>satisfied</u> are you with your <u>life</u>?**

☐ Satisfaction

63. {Good Things} **We have spent some time talking about situations that are difficult for you. Now I'm interested in hearing some of the <u>things</u> that are <u>going well</u> or that you are happy about.**

Can you tell me some examples?

____

(Anything else?)

____

Taking all these and other good things things into account, how often, over the past month, have good things happened?

Frequency
Impact

Over the past month, how much positive impact have good things had on your everyday life?

We’ve covered many different areas in our conversation today. You’ve had some time to think about problems areas. Now, in your own words, could you tell me what your top 3 areas of concern are? (so we can take your words and relate them to our ratings). This could include areas we didn’t cover in the interview.

____

__

____

__

____

__

____

__

Now, can you tell me the three areas where you’ve noticed the most improvement in the last 6 months *(if injury occurred less than 6 months ago, say,* “after leaving the hospital”)?

____

____

_

____

_

____

_

Conversationally, putting away all the interview materials, pulling out the feedback form left unattached to the FOP, and turning over the Questionnaire to show blank page on back, ask something like the following:

"Now we're all done with the interview and I'd like to get some feedback from you about the Questionnaire and the interview."

Clinical Notes: (Impressions of client and their overall level of functioning, major areas of concern noted by interviewer, etc.)

---

[i] Only the 63 questions regarding the survivor were in this analysis, with the 3 questions in reference to the significant other omitted (Teasdale et al., 1997)

[ii] It has been shown that ratings of pre-injury functioning tend to represent an idealized view where less pre-injury difficulties are reported than is found in a non-injured control group, indicating that retrospective evaluations of pre-morbid functioning are unreliable (Deloche, Dellatolas & Christensen, 2000).

[iii] If client takes scale book off the table, say that you will be asking them to answer questions by pointing to places on the scale, and that you need the book on the table to be able to see the responses. Interviewer should keep control of page flipping, to control the interview pace.

Feedback

Were there any unclear questions?

Is there anything you would like to see changed?

Were there any areas missing?

Were you comfortable during the interview?

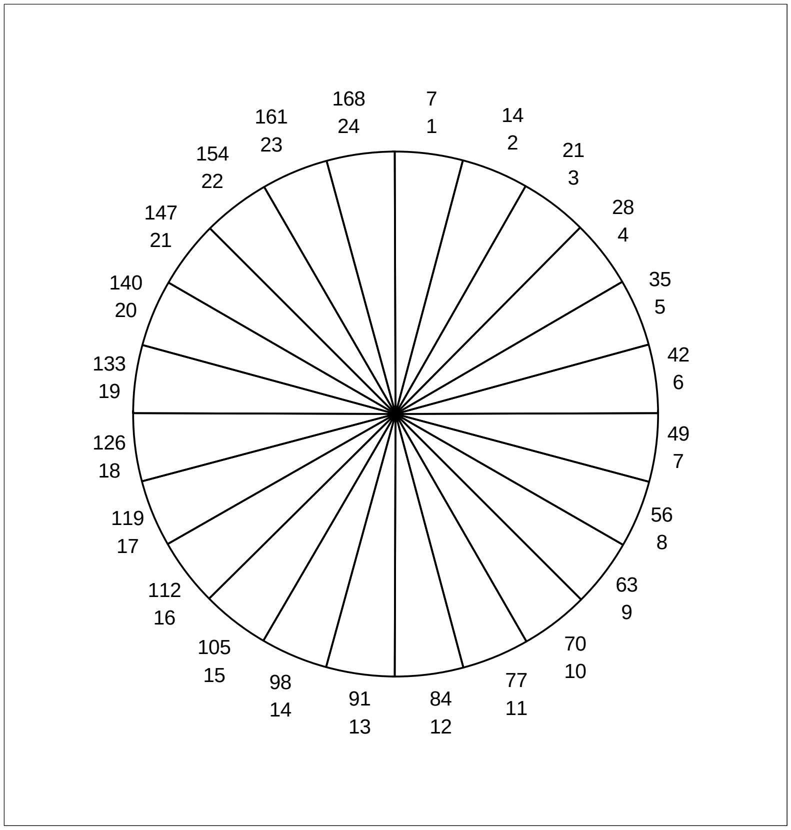
Was there anything that made you uncomfortable?

Did you feel you gained anything from the interview?

Do you have any general comments?

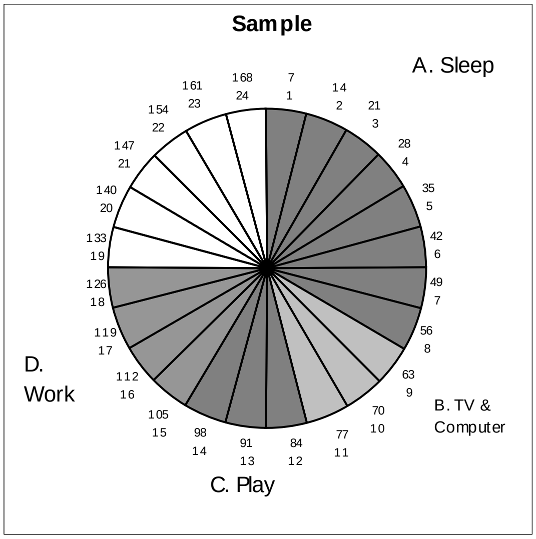
Activity Hours WorkSheet

This worksheet is provided as a visual aid for you to use if you are having difficulty obtaining realistic estimates of time use (e.g., adding up to more than 24 hrs/day) In such circumstances, consider filling out the 24-hour clock with your client. Shade in the proportion of the clock as the client reports the hours spent engaged in the activity. Label each shaded region A-D. (See sample.) It is not necessary to account for every hour in the day.



	Activity	Activity	Hrs/Day
A	Sleep (& napping)		
B	Personal Care (eating, toilet etc.)		
C	TV and Computer play		
D	Other Leisure		
	1		
	2		
	3		
	4		

	5		
	6		
		Total Play	
E	Work		
	1		
	2		
	3		
	4		
	5		
	6		
		Total Work	



TABLES TO RECORD DATA FOR QUESTIONS 9, 51, 52, 53 and 55

Hr/week	1	2	3	4	5	10	15	20	25	30	35	37.5	40
Hr/day	.15	.25	.5	.5	.5	1.5	2	3	3.5	4	5	5.5	5.5

	Hrs/week	Hrs/Day	Days/Wk	% Help Req'd
8. Sleep/Napping (p. 7)	N/A		N/A	N/A
51. Personal Care Activities (p. 20)	Hrs/week	Hrs/Day	Days/Wk	% Help
1. All Personal Care (or just Bathroom)				
2. Eating				
3. Other.				
52. TV and Games (p.20)	Optional		N/A	N/A
53. Leisure Activities (p.20)	Hrs/week	Hrs/Day	Days/Wk	% Help
1.				
2.				
3.				
4.				
5.				
6.				
7.				
8.				
55. Work Activities (p.21)	Hrs/week	Hrs/Day	Days/Wk	% Help
1.				

2.				
3.				
4.				
5.				
6.				
7.				
8.				
If times seem excessive, convert hrs/wk to hrs/day and total here. If total hours exceed 24, use Activity Hours worksheet to revise estimates with client.	Total			