Evaluating the Efficacy of Individualized Goal Setting in Traumatic Brain Injury Rehabilitation: Does Individualized Goal Setting at the Micro Level Achieve Meaningful Change in Global Outcome?

by

Nicholas Mark Bogod
B.A., McMaster University, 1993
M.Sc., University of Victoria, 1999

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

DOCTOR OF PHILOSOPHY

In the Department of Psychology

We accept this dissertation as conforming to the required standard

© Nicholas Mark Bogod, 2005
University of Victoria

All rights reserved. This dissertation may not be reproduced in whole or in part, by photocopying or other means, without the permission of the author.

Supervisor: Dr. Catherine A. Mateer

ABSTRACT

In today's financially restrictive health-care climate there is increasing onus on health care providers to demonstrate that their methods of intervention are effective. In brain injury rehabilitation, there is a lack of well-established outcome measures and the move towards evidence-based rehabilitation practice is in its infancy. Although a common method of rehabilitation is to deconstruct long-term rehabilitation goals into smaller, more manageable goals, the relationship between improvement on these smaller goals and global outcome lacks empirical basis. This study examined the relationship between improvement on small goals as measured by Goal Attainment Scaling (GAS), and improvement on a comprehensive and a more focused measure of global outcome administered at intake and discharge. The study took place at an in-patient residential brain injury program, Skeleem Recovery Centre (SRC), on Vancouver Island. GAS was used to quantify four goals for each participant, and produced a numerical index of improvement on these goals. The Mayo-Portland Adaptability Inventory – IV (MPAI-IV) and the Supervision Rating Scale (SRS) were used to measure general and more focused outcome, respectively – the difference in ratings on these measures between intake and discharge were used as the index of improvement on the respective measures. Sixteen participants who had sustained traumatic brain injury were evaluated. The results indicated that improvement on small goals as measured by Goal Attainment Scaling was significantly associated with improvement in terms of outcome on the MPAI – IV and SRS difference scores. The MPAI-IV change was significantly predicted by GAS over and above SRS change. Investigation of the three MPAI-IV subscale difference scores revealed that GAS change was predictive of each subscale individually, but not when the variance associated with the other two subscales was partialed from the analysis, suggesting that they may be capturing similar information. The participants were classified as either mild-moderate or severe TBI based on injury characteristics (e.g., Glasgow

Coma Scale). Logistic regression techniques were used to investigate which measures would best predict severity. Due to limitations in sample size and only three participants falling in the mild-moderate brain injury group, the predictors could only be examined individually. Limitations of the study and future directions are discussed.

TABLE OF CONTENTS

Page
Abstract....ii
Table of Contents....iv
List of Tables....v
List of Figures....vi
Introduction....1
Traumatic Brain Injury....3
Traumatic Brain Injury Rehabilitation....5
Outcome Measurement in Traumatic Brain Injury....6
Selected COMBI Outcome Measures....12
Assessing Outcomes in Traumatic Brain Injury....17
Goal Attainment Scaling (GAS)....18
Global Versus Specific Outcome Measures....25
Selection of Outcome Measures....27
Skeleem Recovery Centre Outcome Measurement Project....28
Implementation and Calculation of GAS Scores....31
Research Predictions....35
Methods....37
Participants....37
Setting and Apparatus....39
Measures....40
Results....44
Data Screening....44
Descriptive Statistics....44
Demographic Variables....46
Analyses of GAS and the SRS and MPAI-IV....47
Analyses of GAS and the SRS and MPAI-IV Subscales....49
Prediction of Injury Severity by Logistic Regression....52
Discussion....54
Relationship between GAS, SRS, MPAI-IV with the Demographic Variables....55
Relationship between GAS, SRS and MPAI-IV Difference Scores....56
MPAI-IV Subscales....59
Predictors of Brain Injury Severity....60
Selection of Outcome Measures....62
Considerations Regarding the Outcome Measures....64
Use of Difference Scores....66
Considerations in the use of Goal Attainment Scaling....68
Challenges of Implementation....71
General Considerations and Limitations of the Study....73
Future Directions....75
References....77
VITA....87

LIST OF TABLES

Page
Table 1 – Six Steps for the Development and Implementation of GAS....................20
Table 2 – Classification of Mild, Moderate and Severe TBI....................................39
Table 3 – ImpGAS, MPAI-IV & SRS Mean, SD and Range...................................45
Table 4 – Bivariate Correlations: ImpGAS, SRS, MPAI-IV and Demographic Variables 46
Table 5 – Correlations Between ImpGAS, and the MPAI-IV and SRS Difference Scores...47
Table 6 – Prediction of the MPAI-IV Subscale Difference Scores by SRS change..50
Table 7 – Prediction of the three MPAI-IV Subscale Difference Scores by ImpGAS 50
Table 8 – The Relationship of ImpGAS to MPAI-IV Subscale Change Over and Above SRS...51
Table 9 – The Unique Relationship of Each MPAI-IV Subscale Difference Score with ImpGAS...52
Table 10 – Prediction of Severity Group (Mild-Moderate versus Severe Injury).....53

LIST OF FIGURES

Page
Figure 1 – GAS Team Planning Flowchart.....33
Figure 2 – Sample GAS item.....34

Evaluating the Efficacy of Individualized Goal Setting in Traumatic Brain Injury Rehabilitation. Does Individualized Goal Setting at the Micro Level Achieve Meaningful Change in Global Outcome?

In today's challenging financial climate, service providers in general are under greater and greater pressure to justify the value of the services they provide. Nowhere is this more pronounced than in the health-care sector (Ashley & Krych, 1990; Banja, 1999). The need for government health ministries to balance health care delivery and funding with fiscal restraint has resulted in increasing demands for health care providers to streamline services, providing only those treatments proven efficacious, most expedient, and most cost effective. This has produced a major impetus to define techniques, interventions and treatments that are "evidence-based." The move towards evidence-based medicine has had profound and far-reaching influence on research and practice in the area of health care (Grol, 2001). Today, most practitioners recognize that health care providers must either establish criteria and undertake the task of validating their methods of practice, or risk their funders doing so for them and abide by the consequences. Practitioners in areas of practice where the techniques are more difficult to quantify are thus charged with the task of defining, adapting and refining outcome measures that can produce reliable and valid data to support the efficacy of their techniques.

This is particularly relevant to brain injury rehabilitation due both to the degree of morbidity associated with more severe injuries, and to the high cost of rehabilitation, particularly in the case of severe injury (Cardenas et al., 2001). A meta-

analysis of the effectiveness of rehabilitation methods for traumatic brain injury (Chestnut et al., 1999) found that even the strongest studies reflected limits in research design, method of analysis, patient selection, and relevant outcome measures. In a follow-up article examining cognitive rehabilitation in TBI, Carney et al. (1999) lamented the lack of operationalization and standardization of outcomes and outcome measures, and recommended future research incorporate standard definitions around interventions, and relevant outcome measures. The lack of consistent measurement of outcomes is a significant problem for the TBI rehabilitation field, since intervention efficacy research conducted at one site is almost invariably measured with a different combination of instruments than at other sites.

Recent efforts have attempted to address some of these concerns, including funding of the TBI Model Systems of Care (TBIMS; Bushnik, 2003), and the creation of the Center for Outcome Measurement in Brain Injury (COMBI; Wright, Bushnik & O'Hare, 2000). The TBIMS represents a collaborative effort of TBI rehabilitation centres in the United States (US) to provide comprehensive TBI rehabilitation services across the lifespan, and to further knowledge and foster research. The COMBI is an internet-based resource that provides information on brain injury outcome measures. The outcome measurement resources available on the COMBI reflect those used and endorsed by the TBIMS rehabilitation programs. Additional measures are reviewed for merit and added periodically.

The ambitious goal of achieving a unified set of outcome measures used by a majority of TBI rehabilitation programs is a long way off. However, the TBIMS and

COMBI are increasing awareness of the issue, and providing ready accessibility to well-established outcome measurement tools.

Traumatic Brain Injury

Traumatic brain injury (TBI) is defined by the Brain Injury Association of America (BIAA, 2004) as "an insult to the brain, not of degenerative or congenital nature, caused by an external physical force that may produce a diminished or altered state of consciousness, which results in an impairment of cognitive abilities, physical, behavioural, and/or emotional functioning." The American Centre for Disease Control and Prevention (aCDC, 2004) describes TBI as "a blow or jolt to the head...which can disrupt the function of the brain." Other centres around the world have slightly different definitions of TBI. However, the descriptions above share two critical elements: damage to the brain by a physical force, and possible change in brain functioning as a result.

A survey of the causes of TBI in 1995-1996 reported by the BIAA (BIAA, 2001) indicated that 44% were related to transportation, 26% to falls, 9% to assaults, 8% to firearms, and the remaining 13% to other or unknown causes. Data from the United Kingdom suggests traffic-related accidents account for approximately 40% of TBI (Das-Gupta & Turner-Stokes, 2002). These data are not available for a Canadian sample, but likely reflect similar causes with the exception of firearm related injuries in the U.S. data.

The aCDC found that 1.5 million Americans sustain a TBI each year. Of this group, 50,000 die, and 80,000 are placed on long-term disability (aCDC, 2001). Kraus

(1993) used composite data from all U.S. studies published before 1990 to develop an overall estimate of incidence of 200 TBI cases per 100,000 persons per year, age-adjusted to the 1990 U.S. population. In a US sample (using 1990-1993 census data from Missouri, Colorado, Oklahoma, and Utah), the highest incidence was among 15-24 year old males, and females >75 years of age, with 15-24 year old females having the second highest incidence (CDC, 1997).

Total annual incidence of TBI in the United Kingdom is reported at approximately 300 per 100,000 (Das-Gupta & Turner-Stokes, 2002). In Canada, data are scarcer, but some information is available. Willer & Moscato (1996) analyzed the 1986 Canadian census and reported a national TBI prevalence rate of 74.3 out of 100,000 adults. Data from the Ontario Brain Injury Association suggests yearly incidence rates of 115 per 100,000 in Ontario, with 24 per 100,000 identified as needing neurorehabilitation. Data provided by the Insurance Corporation of British Columbia is difficult to interpret as they report only police-attended collisions, and only report by most severe injury type as rated by the attending police officers. On this basis, concussive injuries in 1999 represented 623 of 28,117 police-attended motor vehicle accidents (MVA), or approximately 2% of such accidents. The US data suggests that MVA's represent the cause of less than half of all TBI's. Further, the ICBC data is unreliable, as in many cases concussive injury may be seen as secondary in severity to other injuries (e.g., amputations, bleeding, fractures). Unsupported data provided by the British Columbia Brain Injury Foundation indicates that there are 14,

000 new brain-injuries in BC each year, and that the associated health-care costs of TBI Canada-wide exceed one billion dollars per year.

As the data above indicate, TBI is a relatively prevalent source of injury that results in many fatalities, and a high degree of disability. Hospitalization and fatality costs related to TBI in the United States are estimated at 48.3 billion dollars per year. Little data is available around the cost of rehabilitation for survivors, but the combined financial and morbidity costs are likely far higher. In fact, TBI is the leading source of injury and neurological disability among young adults in the United States (Guilmette & Paglia, 2004). Survivors of traumatic brain injury often have a broad array of injury sequelae, including bladder problems, paresis and contractures, seizures, agitation and confusion, problems with memory, verbal and physical aggression, sexual disinhibition, lack of awareness of deficits, depression, etc. (Das-Gupta & Turner-Stokes, 2002). In addition, TBI is highly heterogeneous, resulting from a wide range of pathologies including axonal shearing, focal injuries, and space occupying hematomas (Ballen et al., 2003). This heterogeneity demands a high level of flexibility and adaptability from rehabilitation programs as they attempt to provide safe and effective rehabiliation to this diverse group.

Traumatic Brain Injury Rehabilitation

In terms of brain injury rehabilitation, the field is still in its infancy, with most programs having come into existence within the last 15 years. In general, it appears that rehabilitation programs have found that individualized treatment plans must be constructed from the ground up with every new admission (Pender & Fleminger,

1999). This is not to say that commonalities do not emerge. Rather, the unique contributions of each person's: character; premorbid strengths and diatheses; genetics and capacity for spontaneous neural reorganization; developmental and social history; family supports and circumstances; financial circumstances/socioeconomic status; level of education; focal and/or diffuse brain damage; and the unique interaction between these myriad factors all have influence on the person's presentation, and speed and degree of recovery post-injury. This is well summarized by Pender and Fleminger (1999, p. 347-348), who stress that "within neurorehabilitation settings it is inevitable that our interests are directed at the performance of one individual over time and charting his or her progress in treatment." Although there are numerous anecdotal reports of treatment efficacy in the literature, few large sample double-blind placebo controlled trials have been conducted (Chestnut et al., 1999), and little evidence has been generated to predict which approaches will be effective for subgroups of the brain injured population (Bajo & Fleminger, 2002). This is due to the recent emergence of the field, the heterogeneity of the brain injured population (Sohlberg & Mateer, 2001), and the lack of established outcome measures that have demonstrated reliability and validity when applied to the brain injured population as a whole (Chestnut et al., 1999).

Current State of Outcome Measurement in Traumatic Brain Injury

A review of the literature reveals many studies using a variety of different outcome measures to demonstrate improvement, and measure effectiveness of rehabilitation techniques. A number of studies have used combinations of the

Functional Independence Measure (FIM), the Functional Assessment Measure (FAM), the Disability Rating Scale (DRS), and the Community Integration Questionnaire (CIQ) as their primary outcome tools. Hammond et al. (2001) examined change over time in 1160 subjects using the FIM and the DRS. The authors concluded that the DRS was more sensitive to changes over a short period of time than the FIM, and also superior at detecting long-term deficits. Gurka et al. (1999) evaluated the relationship of the FIM and FAM with the CIQ and the Return to Work Scale (RTW) in 88 patients with severe TBI's. At 24-month follow-up they found that the FAM motor scale was the only significant predictor of the CIQ, and that the FAM cognitive score was the best predictor of RTW status. They found that the FAM subscales produced only modest gains in prediction of employment status and community integration at 24 months post-discharge. Semlyen, Summers & Barnes (1998) used the Barthel Index, FIM, and Newcastle Independence Assessment Form (NIAF) to compare multidisciplinary versus single discipline approaches to the rehabilitation of 56 sequential admissions for severe brain injury. The authors concluded that the multidisciplinary approach was more effective on the basis of the outcome data. However, they encountered ceiling effects with both the FIM and Barthel Index.

Another study (Gray & Burnham, 2000) examined outcomes of inpatient rehabiliation of a mixed brain-injury sample of 349 survivors using the FIM, FAM, and DRS. The authors found that many patients demonstrated improvement in their ratings on the three measures. However, no additional improvement was measured with the FIM and FAM for patients after they had been resident for 12 months in the

program. Finally, Hall et al. (1996) examined the characteristics of the DRS, FIM, FAM, and CIQ in a sample of 612 adults with TBI. Ratings were collected at admission and discharge from acute inpatient rehabilitation, and at one and two years post-injury. The authors found a substantial ceiling effect for the FIM, for the FIM+FAM combination in 1/3 of patients, and in the CIQ home and social integration subscales. The DRS showed less ceiling effect during all time frames than the former measures.

Overall, significant concerns emerge across a number of studies about the ceiling effects of the FIM, FAM and CIQ. When measuring outcome from TBI, this is a significant concern, as the utility of an outcome measure will be directly tied to its ability to measure the incremental improvements characteristic of recovery. Thus, measures that frequently suffer ceiling effects have a reduced utility in adequately monitoring recovery from TBI through its acute to post-acute stages. As Bohac, Malec & Moessner (1997) identify, the FIM and FAM have proven useful as outcome measures for acute inpatients, but fail to measure the cognitive and behavioural impairments that are typically the main focus of post-acute rehabilitation efforts.

Stilwell et al. (1999) identify that large studies have typically used the Glasgow Outcome Scales (GOS), or extrapolated outcome from indicators such as RTW. However, as Stilwell et al. (1999) point out, the limited range of the GOS results in little utility after the acute period of recovery. This is a similar problem to the ceiling effects of the former measures in that the limited categories of the measure are insensitive to all but drastic changes in functioning. In addition, RTW is not a

good indicator as it is confounded by the prevailing employment climate of the time and location, and the degree of availability of supported or graduated RTW programs. Bohac et al. (1997) found that unidimensional brain injury outcome measures such as the GOS, DRS, and Rancho Levels of Cognitive Functioning Scale (RLCFS) failed to adequately capture the multidimensional nature of brain injury outcomes.

Others have tried to develop new measures to capture meaningful aspects of recovery from brain injury. For example, Kolitz, Vanderploeg & Curtis (2003) proposed a new measure of neurobehavioral change in traumatic brain injury called the Key Behaviors Change Inventory (KBCI). The scale is composed of eight subscales that attempt to capture a variety of difficulties including inattention, impulsivity, apathy, unawareness, etc. The scales contain a series of questions that are rated by others on a 4-point Likert format. The authors' initial validation was performed using 75 undergraduate volunteers, 20 members of the Multiple Sclerosis (MS) Society, and 25 collateral informants for individuals with TBI. The authors found that the scale was sensitive to typical behavioural changes after TBI, and that a combination of subscales differentiated MS from TBI. However, they acknowledge that further work was required to demonstrate the reliability and validity of the KBCI. Examples of other measures in current development and validation include:

- The Wisconsin HSS Quality of Life Inventory (WI HSS QOL), which was designed to assess level of need satisfaction after traumatic brain injury, conceptualized in Maslow's theory of human needs (Collins, Lanham & Sigford, 2000).

- The Sydney Psychosocial Reintegration Scale (SPRS), which was designed to quantify handicap in persons with TBI (Tate et al., 1999, 2004).
- The Community Integration Measure (CIM), a new measure of community integration level that attempts to improve on the original Community Integration Questionnaire (McColl et al., 2001).
- The Health of the Nation Outcome Scale (HoNOS–ABI), which targets psychiatric and psychological sequelae of brain injury.

As these examples illustrate, efforts to construct measures that capture meaningful information that can predict and/or measure important outcome relevant characteristics in TBI rehabilitation are ongoing. This reflects both the extreme lack of consensus in selection of outcome measures, the poor validation studies available for many existing measures, and the difficulty in capturing the multidimensional nature of brain injury outcomes. This is supported by Pender & Fleminger's (1999) report of an unpublished review that found no consensus on outcome measures among published outcome studies from programs providing behavioural and neuropsychological TBI rehabilitation. The magnitude of this problem is illustrated by The American National Institute of Health's evidence-based practice report on TBI rehabilitation (Chestnut et al., 1998) that reviewed 3000 original research articles and selected 363 for examination for scientific rigor and statistical validity. Chestnut and his colleagues found numerous flaws in the research literature including: a lack of standardized outcome measures; inadequate descriptions of rehabilitation interventions; a lack of randomized trials; spontaneous recovery confounding

improvement results; and unclear or absent control groups. The NIH consensus statement on traumatic brain injury rehabilitation (1998) recommended a strong focus on better outcome studies with better and more standardized outcome measures.

In their editorial introduction to a dedicated issue of Neuropsychological Rehabilitation examining outcome measurement in brain injury rehabilitation, Fleminger and Powell (1999) identified that in-patient cognitive rehabilitation units have a common theme of preference for individualized measurement derived from goal planning, or recorded through behavioural intervention programs. However, they criticize this approach on the basis that this type of outcome measurement may not translate to improvements in independence or quality of life. Although they do not suggest that every rehabilitation centre should use the same outcome measures, they do argue the need for identification of measures that could be used with reasonable consistency on inpatient cognitive and behavioral units (Fleminger & Powell, 1999).

As mentioned earlier in the introduction, the Center for Outcome Measurement in Brain Injury (COMBI) was formed through a National Institute on Disability and Rehabilitation Research (NIDRR) grant (Wright, Bushnik & O'Hare, 2000). The COMBI represents a collaborative project of eight traumatic brain injury model system centres in the United States (Bushnik, 2003), and provides internet-based resources for TBI outcome measurement including rating scales and forms, administration guidelines, descriptions of scale properties, supportive references, training and testing materials, and contact information. The goal of the COMBI is to act as a resource and disseminate information on standardized outcome measures to

the TBI rehabilitation community. The COMBI represents one of the first efforts to unify centres around a core battery of well-validated assessment tools. The advantage of centres sharing the same assessment instruments cannot be overstated. The use of a shared pool of instruments provides a common metric and language for communication about patients within and across rehabilitation centres. It also creates greater opportunity for large, multi-centre outcome trials, and broader validation of the instruments selected. This is well articulated by Ponsford, Olver, Nelms, Curran & Ponsford (1999) who acknowledge that until rehabilitation programs agree on a unified set of outcome measures, little utility can be found in comparing one program with another.

Selected COMBI Outcome Measures

The COMBI site provides detailed information, including protocols and scoring instructions, for a number of established measures of TBI outcome:

The Agitated Behavior Scale (ABS; Corrigan, 1989) was developed for the assessment of agitation in survivors of TBI, with a goal of permitting serial assessments of agitation level to measure change over time. Fourteen items were selected from an initial pool of 39 items based on their assessed ability to capture the full domain of the agitation construct. Item ratings ranged from one to four, with one indicating the absence of behavior and four indicating severe unredirectable behavior. Reliability was examined in a sample of 35 participants with brain injuries. The results indicated that Cronbach's alpha and theta exceeded .80 for all raters. The ABS was found to account for between 36-62% of the variance when correlated with

simultaneous independent observations of agitation. Subsequent studies have demonstrated the ability of the ABS to differentiate confusion from inattention (Corrigan, Mysiw, Gribble & Chock, 1992), and measure change in cognitive status (Corrigan & Mysiw, 1988). More recently, Bogner, Corrigan, Stange & Rabold (1999) demonstrated acceptable interrater reliability in a sample of 45 survivors of TBI and 23 persons with progressive dementia. Ratings of the survivors of brain and the dementia sample by research assistants yielded correlation coefficients of .92 and .86, respectively. The authors concluded that the ABS is a reliable instrument for measuring agitation in survivors of brain injury and persons with dementia.

The Supervision Rating Scale (SRS; Boake, 1996) measures the level of supervision that a patient receives from caregivers. It provides a rapid and objective index of the degree of supervision required at any point in time. The SRS rates level of supervision on a 13-point Likert scale ranging from full-time supervision to full independence. Hart et al. (2003) examined the relationship between demographic variables, neuropsychological measures, and level of supervision (measured by the SRS) in a sample of 563 adults who had sustained traumatic brain injury. The results suggested that pre-injury education and measures of cognitive flexibility predicted functional independence after TBI. They also noted that the SRS appeared to be prone to ceiling effects in follow-up.

The Mayo-Portland Adaptability Inventory – IV (MPAI-IV; Malec & Lezak, 2003) is based on the Portland Adaptability Inventory (PAI; Lezak, 1987) that was constructed to capture meaningful behavioural and social problems experienced by

persons after brain injury. The MPAI follows the guidelines of the World Health Organization (WHO) distinctions among impairment, activity and participation (World Health Organization, 1997), characterized in rationally derived subscales. Following the WHO guidelines, ratings on each scale item are constructed to indicate whether performance is (i) within normal limits, (ii) mildly limited without affecting everyday functioning significantly, (iii) sufficiently limited that it does affect everyday functioning in varying degrees (this category is broken into three levels of limitation on the MPAI: mild, moderate and severe). The MPAI-IV represents the most recent of a series of refinements of the original MPAI. Initial psychometric investigation of the MPAI demonstrated convergent validity of the MPAI with the DRS (Spearman r = .81), and found that the MPAI scores were significantly different for different group classifications using the RLCFS (Malec & Thompson, 1994). Subsequently, Bohac, Malec & Moessner (1997) conducted principal components factor analysis (PCA) of the MPAI using a sample of outpatients with acquired brain injury. They derived an eight-factor model that accounted for 64.4% of the total variance. Validation of the cognitive factor was achieved when factor scores were correlated with various neuropsychological measures. Validation of the impaired self awareness/distress factor was achieved by examining the difference between staff and participant ratings. Two polar response patterns emerged: one group minimized their deficits compared to staff reports and showed little depression/distress; a second group overstated their deficits and were far more likely to also be depressed. The authors considered making refinements to content items and factor structure based on the results. Rating scale

analysis (RSA) of the MPAI was then conducted using data from 305 outpatients with brain injury (Malec, Moessner, Kragness & Lezak, 2000). The analysis suggested a refined scale that reduced the overall scale items from 30 to 22. The authors then applied the two scale versions to prediction of outcome from a post-injury day treatment program, with the hypothesis that the refined 22-item MPAI would be a better predictor than the original 30-item version. The authors found that the scores on the 30- and 22-item versions of the scale were highly correlated in their sample (r = .98). Their hypothesis that the 22-item version of the MPAI would be a better predictor was not supported, as the two versions of the scale were equally predictive of outcome. The authors concluded that the dimension of the MPAI, established using RSA, was an overall measure of severity of sequelae of brain injury composed of a mixture of impairments, disabilities and handicaps. The authors concluded that the study provided recommendations for improving the reliability through modification of item-rating scales and elimination of non-contributory items from the overall score.

Recently, Malec et al. (2003) performed further psychometric evaluation and revision of the MPAI in a US national sample of 386 survivors, most of whom had sustained moderate to severe brain injury. The authors performed Rasch, item cluster, principal components, and traditional psychometric analyses of internal consistency of the MPAI overall scale and subscales. The research was conducted using the 30-item MPAI (revision III). Rasch scaling indicated that an item evaluating child rearing was not useful for most of the survivors of TBI and it was eliminated from the analysis. Of the remaining 29 items, four items were identified as having ratings that were not well

distributed across the five-level rating scale: work/school, audition, pain, and transportation. For each item, reduction in the number of rating categories resulted in acceptable fit for the four items. Cronbach's alpha for the overall measure suggested acceptable internal consistency ($\alpha = .89$). The resulting 29-item instrument was designated the MPAI-IV. Some items were rearranged to correspond to rational groupings of categories: ability (sensory, motor and cognitive abilities); adjustment (mood and interpersonal interaction); and participation (social contacts, initiation, and money management). The authors indicate that the subscales were selected on a rational rather than psychometric basis because they corresponded to clinical experience and had value in clinical settings. The subscales all correlated strongly with the overall MPAI-IV score, and were moderately intercorrelated, suggesting some degree of independence. Although cluster analysis of the subscales did not exactly represent the rationally derived subscales, the authors retained the rational subscales as they felt they better reflected clinical theory and practice. The authors also note that previous investigations, and unpublished outcome data from the TBI Model Systems sites, have found a unitary underlying TBI outcome dimension measured by the MPAI-IV that includes indicators of ability, activity and participation. The authors also provided reference MPAI-IV data from their sample for comparison purposes. They concluded that the MPAI-IV appears to be a reliable clinical instrument for measuring the outcome of rehabilitation interventions.

Assessing Outcomes in TBI

Pender and Fleminger (1999) identify that outcomes can be assessed by direct interview, ratings of patients and/or caregivers (such as those described above), multidisciplinary goal planning, and direct observation. Specific characteristics of the patient population including length of stay, in-patient or community dwelling, severity of injury, and types of challenges to recovery and independence influence the choice of measurement modality. For example, direct observation methods may not adequately measure change if the person is community dwelling and observation of them is not easily accomplished across environments. The recent emphasis on team approaches to rehabilitation has encouraged interest in multidisciplinary systems of goal setting. A recent survey of interdisciplinary brain injury rehabilitation team members' satisfaction with goal planning meetings was conducted by Nair & Wade (2003), who surveyed 44 rehabilitation professionals of various disciplines from 21 different rehabilitation teams. The results indicated that team members were satisfied both with the process of goal planning meetings, and with the behavior of other participants. Satisfaction with overall outcome was related to the degree of the individual team member's sense of participation.

A recent published example (McMillan and Sparkes, 1999) describes a system of client-centered goal setting where long and short-term goals were established by a rehabilitation team with input from the client. Short-term goals were those that could normally be achieved in one or two weeks, and were typically steps towards reducing handicap. Long-term goals provided an overarching focus, and were typically in

effect for the duration of rehabilitation. The authors examined 100 consecutive neurorehabilitation cases and found a relationship between long term goals achieved and change on a standard disability outcome battery. Their approach is a common one among brain injury rehabilitation centres – breaking long-term goals down to smaller goals to both facilitate improvement on a more manageable basis, and to give patients a sense of accomplishment as they reach short term goals. What they did not investigate is whether degree of improvement on the short term goals was directly associated with improvement. In fact, a search of the literature did not produce any studies that examined the relationship between achievement of short-term goals and global outcome. The main reason for this is the difficulty in quantifying the degree of improvement on short-term goals. The heterogeneity of the brain injured population creates difficulties in terms of comparing individuals, or examining group effects, in terms of improvement on these goals. However, the literature does suggest a mechanism by which improvement on short-term goals might be quantified, and includes some references to its application to brain injury rehabilitation.

Goal Attainment Scaling

Goal Attainment Scaling (GAS) was developed by Kiresuk & Sherman (1968) for the purpose of evaluating mental health outcomes. GAS is a case-specific method where a small number of goals are scaled for each individual. Malec (1999) reports that GAS is useful in brain injury rehabilitation settings for: monitoring progress; structuring case-conferences; planning and decision making in ongoing rehabilitation; ensuring concise and purposeful information sharing with the patient, family, referral

and funding sources; guiding the delivery of reinforcement; and evaluating the rehabilitation program, both individually and globally. Malec (1999) also notes that the explicit focus on goal setting encourages self-awareness, and can assist in redeveloping the capacity for goal setting in the patient. In its original form (Kiresuk & Sherman, 1968), only one major goal is set for each life-role, but a number of short-term goals may be established representing specific tasks that are components within that role. However, in more recent practice, GAS has been adapted in many different ways to suit the precise needs of the monitoring and goal setting environment (e.g., Barrett, Wilson, & Long, 2003; Rockwood, Graham & Fay, 2002).

The GAS process is described as a "relatively straightforward six-step process" (Joyce, Rockwood & Mate-Kole, 1994). As illustrated in table 1 below, GAS involves determining goals and articulating expected levels of outcome in objective behavioural terms. The goals are rated on a 5-point Likert scale where "-2" represents an outcome within the specified time-frame that is "much less than expected," "0" represents the expected level of progress, and "+2" represents outcome levels that are "much better than expected." Although a weighting step is included in the GAS process, most centres that have adopted GAS as part of their rehabilitation programs have not implemented a system of goal weightings, due to the added complexity, better agreement among goal setters in terms of the nature versus the priority of goals (Greenville & Lyne, 1995), and uncertainty around whether goal weights should add up to one (Malec, 1999).

Table 1.

Six Steps for the Development and Implementation of GAS

1. Goal selection
2. Weighting goals
3. Designation of follow up time period
4. Articulation of the "expected" level of outcome in objective behavioural terms
5. Articulation of other outcome levels
6. Assessment of GAS level on admission and at follow-up

Adapted from Malec, 1999

GAS has been used to set goals in a variety of populations including geriatric care and rehabilitation (Evans, Oakey, Almdahl & Davoren, 1999; Rockwood, Howlee et al., 2003; Rockwood, Joyce & Stolee, 1997; Stolee, Zaza, Pedlar & Myers, 1999), classroom outcomes for autistic students (Oren & Ogletree, 2000), school adjustment (Hughes et al., 2001), evaluation of psychotherapy outcomes (Shefler, Canetti & Wiseman, 2001), monitoring behavioural intervention (Mate-Kole et al., 1999; Sladeczek, 2001), chronic pain management (Fisher & Hardie, 2001; Zaza, Stolee & Prkachin, 1998), communication disorders (Schlosser, 2004), rehabilitation from lower-extremity amputations (Rushton & Miller, 2002), spinal rehabilitation outreach (Cox & Amsters, 2002), drug trials (Rockwood, Graham & Fay, 2002; Rockwood, Stolee, Howard & Mallery, 1996), evaluating health care (Kiresuk & Sherman, 1968; Turnbull, 1998), pediatric brain injury rehabilitation (Mitchell & Cusick, 1998), and sexual offender treatment (Stirpe, Wilson & Long, 2001).

A number of studies have used GAS for goal setting and outcome measurement in adult brain injury rehabilitation. Malec, Smigielski & DePompolo (1991) used GAS, the Portland Adaptability Inventory (PAI), and work status to measure the outcome of 16 participants in a post-acute brain injury rehabilitation program. The authors found that GAS scores were highest for those participants who had the best work outcomes, and PAI scores were lower (reflecting less impairment) at intake and discharge for successful program graduates. The GAS scores and PAI were modestly correlated. The authors concluded that the study supported the use of GAS as a "quantifiable, individualized measure that is useful for (1) monitoring patient progress, (2) structuring team conferences, (3) ongoing rehabilitation planning and decision-making, (4) concise, relevant communication to family, referral sources, and funding sources, and (5) overall program evaluation when used in the context of other objective outcome measures." Joyce, Rockwood & Mate-Kole (1994) used GAS and other standardized outcome measures to guide the rehabilitation of 16 in-patients, 13 of whom had sustained a TBI. The authors found that GAS change scores correlated highly with clinical judgements of efficacy of rehabilitation ($r = 0.81$), and modestly with standard measures of outcome. They also reported a high level of interrater reliability for GAS scores at admission ($r = 0.92$) and discharge ($r = 0.94$). The authors concluded that GAS had utility in measuring and evaluating rehabilitation in patients with brain injuries.

Malec and Moessner (2000) used GAS, MPAI, and the Vocational Independence Scale to evaluate the relationship between impaired self-awareness

(ISA) and distress, and the impact of a comprehensive brain injury day treatment program. Ratings of ISA and distress by rehabilitation staff and their relationship to the other outcome measures were examined in a sample of 62 consecutive graduates of the program. The authors found that ISA and distress diminished after program participation. GAS was modestly correlated with the other outcome measures. The authors found that ISA at program's end was a significant predictor of GAS score at program's end, accounting for 23.7% of the variance in GAS score. Distress was also a significant predictor of lower GAS scores, accounting for an additional 6.7% of GAS variance. ISA and distress were also significantly correlated with the MPAI, but not with vocational outcome. Overall, participants who demonstrated less distress and better self-awareness showed more positive change and skill development (as measured by the MPAI), and greater goal achievement (as measured by GAS).

Finally, Ponsford & Olver (1999) describe their program's approach to outcome measurement in TBI rehabilitation. The authors indicate that they use a role checklist and GAS to measure the progress made by program participants towards individual goals. They have documented gains and long-term outcome up to five-year follow-up using a structured questionnaire. They also use other standardized measures of outcome. They report that using GAS has improved both their capacity to focus on the most pressing needs of the participants in their program, and their ability to measure outcomes meaningfully. They do, however, identify that GAS is a very case-specific method, and recommend supplementation with a standardized program evaluation measure.

The GAS process has a number of strengths beyond its ability to evaluate longitudinal change. It offers (i) grading of goal attainment, (ii) comparability across clients and goals through quantification and aggregation, (iii) versatility across populations, interventions and fields, (iv) linkage tied to expected outcomes, (v) facilitation of goal attainment, and (vi) a focal point for team energies (Schlosser, 2004).

While the researchers above report positive results using GAS, there is criticism of the process in the literature. Cytrynbaum et al. (1979) have noted that subsequent studies of GAS have often ignored Kiresuk and Sherman's (1968) proposed methods for using GAS, such as (i) keeping the goal-setters and those that deliver the service independent; (ii) assigning patients randomly to treatment groups; and (iii) performing independent assessment of outcomes. Although many applications have recognized and included independent measures of outcome, little attention has been given to the first two concerns across the numerous settings where GAS has been used. This may be due to limitations in staff resources for independent goal setting, limited rehabilitation resources, and/or ethical issues associated with random assignment in brain injury rehabilitation. Additional concerns have been raised around the degree to which personnel are trained around the setting of goals, as the precision of improvement measured with the GAS system will be highly dependent on the quality of the goals identified and quantified (MacKay, Somerville & Lundie, 1996; Stephens & Haley, 1991). Other concerns have been expressed about the idiosyncratic nature of the GAS process and its vulnerability to bias. Although this

can be mitigated by Kiresuk and Sherman's (1996) suggestion of independent goal setters and goal implementers, as identified, this is rarely the situation. However, Bailey and Simeonsson (1998) have argued that bias can be mitigated to some extent by adequate team training on the GAS process and on actively striving to maintain objectivity. Further reduction in bias can be achieved by attaching objective measures to the evaluation of goal outcomes (Becker et al., 2000).

In summary, GAS has been used to facilitate the process of setting and measuring progress towards individually relevant goals in a broad variety of populations including rehabilitation of survivors of brain injury. It has not typically been used in isolation, but rather has been used as a means of articulating and monitoring progress towards person-relevant goals to facilitate improvement. Ponsford et al. (1999) indicate their adoption of this method and report that it has improved their capacity to focus on the most relevant needs of rehabilitation patients and measure outcomes meaningfully. They further report that it has provided a framework for restructuring case conferences to allow for the setting and review of goals and goal attainment.

Ottenbacher & Cusick (1993) caution that the appropriate use of GAS in clinical environments is to measure longitudinal change rather than functional status. They argue that GAS is best utilized as a tool to monitor change in attainment of goals within specific individuals. For these reasons, other standardized measures of outcome should be included in a program's overall rehabilitation and outcome

measurement system. The addition of outcome measures to specific goals has been identified as a means to reduce possible bias in the GAS process (Becker et al., 2000).

However, goal attainment as quantified by GAS has yet to be associated empirically with improvement on more global measures of outcome. Such a demonstration is very important. At present, most programs providing brain injury rehabilitation services act under the belief that working towards small incremental goals is a meaningful and effective way of improving overall outcome. This idea is attractive and has strong face validity, but requires empirical demonstration that it is truly associated with improvement on more global measures of outcome and change. Other factors such as program duration, spontaneous recovery, and/or common non-specific therapeutic factors of rehabilitation, could be equally contributory regardless of goal attainment.

Global Versus Specific Outcome Measures

The rehabilitation professional who wishes to evaluate outcomes has a broad selection of outcome measures from which to choose. Fleminger and Powell (1999) recommend that evidence-based rehabilitation needs to measure outcomes with instruments that possess the following characteristics: relevance to the patient or caregiver; and good quality in terms of demonstrated validity and reliability.

Outcome measurement should include instrument(s) that: capture broad and global aspects of functioning; capture specific components of functioning, and; assist in delineating and capturing progress towards individually identified goals (preferably on a metric that can be used to make rough comparisons across clients, and/or in terms

of success in facilitating goal outcomes for the population of the facility as a whole). More specifically, instruments should include the following components at minimum:

1. Global assessment measure(s) that capture most aspects of social, physical, cognitive/emotional, and occupational outcomes. These types of measures are time consuming as they cover numerous areas, but can give an indication of global outcome changes when administered on a quarterly or bi-annual basis. A single or a few well-constructed and well-validated instrument(s) with broad coverage should suffice for this aspect of outcome assessment/evaluation.
2. More tailored and specific instruments that can be utilized to measure the efficacy of interventions when specific aspects of a patient's presentation are targeted (e.g., irritability, yelling, difficulty with showering, problems with memory, mobility issues, neglect, psychosis). These can be utilized when a specific behaviour or challenge is a barrier to progress in rehabilitation and/or transition to a less restrictive or more independent state or circumstance. Generally, a collection of well-validated and target-specific measures covering the most common areas of rehabilitation focus could be utilized to measure more specific outcomes of targeted intervention, and discontinued once the target area is satisfactorily remediated (e.g., depressive symptoms, physical aggression, psychotic symptoms, etc.). Specific instruments constructed and demonstrated to capture key features in the area of interest can be used to monitor the variables or features that are most likely to be indicators of improvement/successful intervention. In addition, they can act as checklists or indexes of subcomponents to be focused on or

rehabilitated, and guide rehabilitation professionals to previously identified factors of importance.

3. Individualized goals with specific criteria for improvement and success. These would typically be utilized on a continuous or cyclical review/update basis for all individuals in treatment, and preferably would provide a common metric for scaling each patient's success in moving towards and achieving their identified goals. Ideally this would incorporate a common framework for delineating progression towards goals despite the diversity of the possible goals across patients. This would allow the program as a whole to examine how well they were succeeding in achieving rehabilitation goals, both individually and as a facility/treatment program.

Selection of Outcome Measures

Given that each rehabilitation candidate is unique and requires an individualized treatment plan, no unitary scale would be amenable to assessing the vast array of variables that predict post-injury presentation and outcome. The utility of such a measure, even were it possible to capture all of the important and influencing variables, would be too burdensome in terms of its length and in the time required for its completion to be of use. Further, demonstration of the validity and reliability of such a measure would be onerous if not impossible, and its functional utility, were it validated, is highly doubtful. However, a balance between more limited global assessment of performance in broad areas of functioning (e.g., activities of daily living, cognition, physical challenges, etc.), and measures that permit finer evaluation

and discrimination of specific aspects of post-injury presentation (e.g., attention difficulties, executive dysfunction, etc.) can be of great assistance to programs and the rehabilitation team.

Skeleem Recovery Centre Outcome Measurement Project

Skeleem Recovery Centre (SRC) provided community-based medium-term in-patient residential rehabilitation to individuals with acquired brain injuries. Over the preceding four years, SRC attempted to refine its intervention practices and develop a gold-standard model of goal setting and outcome measurement. To that end, SRC implemented GAS for goal setting and linked the outcomes to established and well-validated outcome measures. Two types of measures were selected: measures of more global outcome for measuring change in global aspects of functioning between admission and discharge; and specific measures that could be used to measure more specific GAS goals. The attachment of specific measures to GAS goals served to better quantify change over time, and to reduce some aspects of possible bias in the GAS process. A review of the recent literature around outcome measurement in traumatic brain injury revealed a broad selection of possible instruments to choose from (e.g., Pender & Fleminger, 1999). Pender and Fleminger (1999) note the inconsistencies among measures used to predict outcome in brain injury rehabilitation units. However, they identify and review outcome measures they perceive to be among the key instruments for assessing different areas of functioning (e.g., global outcome measures, behavioural measures, measures of independence, etc.). A broad review of the literature produces a smaller subset of outcome measures that have been

utilized and favourably reviewed by other influential scientists and rehabilitation professionals in this area (Alderman, Knight & Morgan, 1997; Boake, 1996; Corrigan, 1989; Corrigan & Bogner, 1994, 1995; Dodwell, 1988; Eames & Wood, 1985; Eames et al., 1996; Kiresuk & Sherman, 1968; Levin et al., 1987; Levin, O'Donell & Grossman, 1979; Malec, 1999; Malec, Smigielski & DePompolo, 1991; Malec & Thompson, 1994; Ponsford et al., 1999; Ponsford, Olver & Curran, 1995; Rockwood et al., 1996; Smith, 1981; Webb & Glueckhauf, 1994; Whitneck et al., 1992; Yudofsky et al., 1986). These include measures that have been broadly endorsed and adopted by the Model TBI Systems Centres, available through the COMBI website. Based on this review, a selection of measures were proposed and implemented as an initial battery for evaluation by the Skeleem clinical staff, administration, rehabilitation staff, and clients. The core battery included the Mayo-Portland Adaptability Inventory - 3 (MPAI-III) as a measure of global outcome across many areas of functioning, and the Supervision Rating Scale (SRS) as a measure of more focused change.

Since almost all survivors of brain injury admitted to SRC required at least a moderate degree of supervision, and improvement in supervision level represented in most cases a prerequisite condition for discharge, the SRS provided an excellent measure of baseline improvement. In contrast, the MPAI-III provided broad coverage of numerous functional areas.

An initial sampling of validated instruments used to target individual goals is detailed in the methods section. However, additional targeted measures were

introduced at SRC when the goal in question fell outside the purview of the initial selections.

In summary, outcome measurement in brain injury is notable for the lack of agreement among professionals and rehabilitation centres in terms of which measures to use and how to guide goal setting. Systematic approaches that guide the process of goal establishment, measurement and review can provide a useful framework and a unifying force for multidisciplinary team interventions at both the facility and individual client level. The addition of global outcome measures which are not sensitive to small changes, and are used to evaluate improvement over long time intervals, coupled with specifically targeted instruments that have the sensitivity to measure finer degrees of change in specific areas and interventions, provided a package that could be used to measure change on a number of levels. These included guiding the setting of goals around highly specific individualized interventions, and achieving outcomes and cost effectiveness predictions for the facility as a whole. However, the widely held belief that targeting interventions around change in smaller subcomponents is an effective way to effect global change has not been subjected to empirical evaluation. Fleminger & Powell (1999) identify that one cannot be sure that the improvements identified necessarily translate into improvements in independence or quality of life.

This study was proposed to examine the utility of working towards specific and individualized goals established and evaluated using GAS on a cyclical basis to effect improvement on two identified measures of change: more global change as

captured by the MPAI-III; and more focused change as captured by the SRS, using difference scores created by administering these measures at admission and again at discharge. Since changes in the SRS are a baseline condition for discharge, incremental changes should predict global change as measured by the MPAI-III beyond the more focused change measured by the improvement in SRS.

Calculation of Global Change Scores

Note: MPAI-III scores were converted to MPAI-IV scores after Malec et al. (2003). The total score (range 0-111, 29 items) for the MPAI-IV administered at discharge was subtracted from the score collected at intake to produce a difference score reflecting improvement on the measure. The same procedure was applied to the SRS, again producing a difference score reflecting improvement.

In addition, the MPAI-IV produced 3 sub-scores in the domains of physical/cognition (11 items), pain/emotion (4 items), and social participation (10 items). A change score for these three domains was calculated in the same manner as the total MPAI-IV difference score calculation.

Implementation and Calculation of GAS Scores.

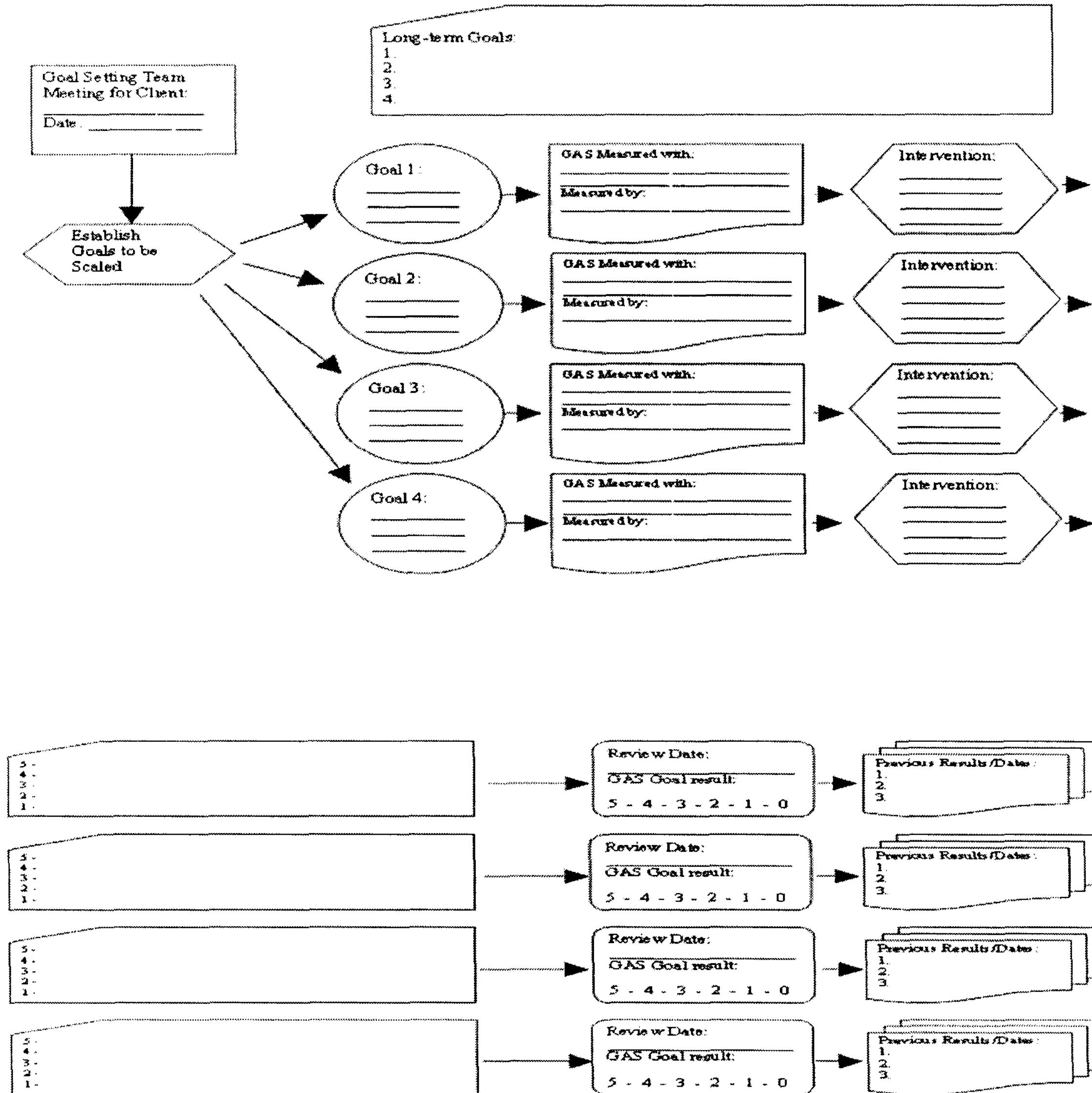
Each participant residing at SRC had four goals that were scaled using the GAS system in each six-week goal setting and review cycle. These goals were always selected by team consensus (including input from the participant) to represent the most clinically urgent goals at that time. In other words, the goals that were selected were the ones that were viewed as the greatest barriers to quality of life, independent living, reduced support costs, etc. Goals were drawn from any domain in which it was

possible to create an operationally defined set of projected outcomes. Examples of goals that were scaled across individuals at SRC included: reduction of psychotic symptoms as measured by the Brief Psychiatric Rating Scale; improvement in core strength and stability as measured by time able to maintain weight bearing standing balance without assistance; greater community exposure as measured by weekly time in community; reduction in agitation as measured by the Agitated Behavior Scale; and reduced tactile defensiveness as measured by the amount of time physical contact to the arm was tolerated, etc.

Each goal was articulated in terms of the goal itself, an intervention(s) was identified, an appropriate outcome measure was selected, and a clinical team member was designated as responsible for ensuring implementation and data collection. The outcome measure was then completed as a baseline from which predicted improvement levels were projected in objective, operationally defined, and measurable terms. Levels 1 (much less improvement than expected) through 5 (much greater improvement than expected) were articulated for each goal in this manner with a level 3 prediction representing expected improvement by the next review cycle (six weeks later). The flowchart below (see Figure 1) gives a graphical representation of the GAS components and format.

Figure 1.

GAS Team Planning Flowchart



Below is a sample goal of using a memory book articulated for the five levels (see figure 2). The data collected represents the successful independent notation of ten target memory items presented per week with a baseline success of 7% of the time and

expected success of 30-39% (Level 3 scaling projection) of the time after six weeks of intervention (identified as cueing use of memory book by staff).

Figure 2.

Sample GAS Item

Sample Scaled Goal: *"to use my memory notebook to compensate for my memory problems"*

- 5 – I use my memory book 50-59% of the time to record information I need to remember
- 4 - I use my memory book 40-49% of the time to record information I need to remember
- 3 - I use my memory book 30-39% of the time to record information I need to remember
- 2 - I use my memory book 20-29% of the time to record information I need to remember
- 1 - I use my memory book 10-19% of the time to record information I need to remember

---

Adapted from Malec, 1999.

At the conclusion of the six week cycle, the identified measures were repeated and achievement levels on the GAS scales were identified. Goals that required no further improvement or intervention were replaced by the next most pressing goal, while goals that required further intervention and improvement were rescaled for further intervention during the subsequent GAS cycle. Thus, every six weeks, data were

collected around progress on four goals for each participant. The system permitted conversion of GAS improvement to a common metric so that disparate goals, both within and across individuals, could be compared. This mirrored the typical goal setting and review approach of most rehabilitation programs, but also provided quantification on a common metric to permit comparison to global outcomes while still respecting individual needs and challenges.

For the purposes of evaluating the efficacy of individual goal setting, the four GAS scores produced for each participant in a six-week cycle were summed, and then divided by four, to produce a mean improvment score for that cycle ranging from zero to five. At discharge the combined scores for each cycle were summed and divided by the number of cycles to produce an average improvement per cycle for the duration of the person's residency in SRC. This resulted in a common "improvement per GAS cycle score" (ImpGAS) that accommodated the variability in the length of rehabilitation residency of given participants. This score was then used to evaluate the impact of individualized and specific goal setting on global outcomes.

The following predictions were made:

1. It was predicted that change on specific individual goals as quantified by the ImpGAS score would be associated with a reduction in supervision level as measured by the SRS, which is a focused outcome measure and a baseline condition for discharge. A regression analysis was performed where the ImpGAS score was used to predict the SRS difference score

2. It was predicted that change on specific individual goals as quantified by the ImpGAS score would be associated with improvement in the MPAI-IV Total Difference Score, the outcome measure designed to capture broad aspects of change in function. A regression analysis was used where the ImpGAS score was used to predict the MPAI-IV total change score.
3. Since improvement in supervision level represents a baseline condition for discharge, while change as measured by the MPAI-IV captures broader areas of potential improvement, the incremental change quantified by ImpGAS should better predict the more global change measure (MPAI-IV difference scores) beyond the more focused change measure (SRS difference scores). It was predicted that change on specific individual goals as quantified by ImpGAS score would predict MPAI-IV difference scores over and above change in SRS. This hypothesis was addressed by regressing MPAI-IV total difference scores on ImpGAS using change in SRS as a covariate.
4. Three sub-domains have been identified for the MPAI-IV. It was predicted that incremental change as measured by ImpGAS would be predictive of change on each of the three subscale difference scores. This was examined by separately regressing each of the subscale difference scores individually onto ImpGAS. The unique relationship between ImpGAS and the individual subscales difference scores was then examined by regressing each subscale difference score onto ImpGAS, with the other two subscale difference scores included as covariates.

5. The participants were divided into two groups, identified as mild-moderate and severe brain injury respectively, using severity criteria as identified by Lezak (1995). Where more than one indicator was available, the classification that predicted the greater severity was used. Logistic regression was performed to identify which variable(s) best predicted severity. Predictors of interest included ImpGAS score, MPAI-IV Total Difference Score, reduction in supervision on the SRS, and months in treatment program. It was predicted that longer treatment durations, lower ImpGAS scores, lower MPAI-IV Total Difference Scores and greater need for supervision (SRS) would have predictive value for greater severity. Due to the limited sample size and distribution of the sample (i.e., only 3 participants classified in the mild-moderate severity range), predictors for logistic regression were examined individually.

Methods

Participants

Participants were drawn from serial admissions during 2002 to SRC, a medium-stay in-patient residential brain injury treatment centre on Vancouver Island. Data was collected for 16 sequential admissions. There were no exclusion criteria for participation. The original study proposed collecting data on at least 20 participants, with an agreement to supplement the sample by treating individuals who were still in the program as if they had been discharged at the conclusion of the study. At that time, the exit ratings on the global outcome measures would be completed for those truncated cases as if they had been normally discharged. Unfortunately, despite being

acknowledged as a unique and model rehabilitation program in British Columbia, SRC was closed due to health care budgetary cutbacks. The time frame of the closure permitted normal discharge of the existing residents, but no new admissions were taken for approximately six months prior to the facility's closure. As a result, truncated cases were not available to supplement the sample, and the study was concluded approximately 12 months sooner than intended. This resulted in a sample of 16 participants for this study.

Human subjects approval was obtained from SRC and the University of Victoria. Consent was obtained for data use from each individual participant. Data collected was numerically coded to protect participant identities and stored in a locked cabinet. Participants were recruited by letter. There was no financial remuneration offered as the study evaluated the efficacy of the program participants were already receiving, and no additional requirements of time or effort were involved.

Thirteen males and three females participated in the study. Mean age for the sixteen participants was 39.4 years (SD = 9.58, Range = 27 – 65). Mean time post-injury was 47.4 months (SD = 19.92, Range = 17 – 83). Mean time in the rehabilitation program was 9.38 months (SD = 2.81, Range = 5 – 14). Etiology of injury included motor vehicle accident (13 participants), pedestrian stuck by a vehicle (one participant), sequelae of assault (one participant), and diffuse encephalopathy (one participant). Participants were classified into either mild-moderate, or severe TBI groups on the basis of GCS score, PTA, or LOC depending on which were available (see Table 2 for severity classification). Where more than one indicator of severity

was available, severity was classified using the indicator that provided the most severe classification.

Table 2.

Classification of Mild, Moderate and Severe TBI

TBI Classification	GCS	Length of PTA	Duration of Coma
Mild	13 or greater	60 minutes or less	20 minutes of less
Moderate	9 to 12	1 to 24 hours	less than 6 hours
Severe	less than 9	over 24 hours	over 6 hours

Note. Adapted from Neuropsychological Assessment (3rd ed.), (p. 173, 755), by M. D. Lezak, 1995, New York, Oxford University Press.

This resulted in the classification of three participants as mild-moderate TBI and thirteen participants as severe TBI.

Setting & Apparatus

The study took place on-site at SRC. Goal setting included participant input to the greatest possible degree, and was the product of team consensus. The clinical team consisted of an occupational therapist, physiotherapist, social worker, administrator, clinical coordinator, music therapist, counselor, behavioral specialist, nurses, rehabilitation assistants, and primary care front-line staff. Additional consultation on some goals was provided through a neuropsychiatry telemedicine program, and by the primary care physician on a weekly basis. Data collection was assigned to the most suitable member(s) of the clinical team (e.g., nurses evaluated

responses to medications and wound care, the behavioral specialist monitored behavioral interventions and outcomes, occupational therapy monitored changes in participation, physiotherapy monitored changes in mobility and range of motion).

Measures

Global assessment/outcome measures.

Mayo-Portland Adaptability Inventory – 4 (MPAI-IV) (Malec & Lezak, 2003).

The MPAI-IV is adapted from the Portland Adaptability Inventory (Lezak, 1987) and assesses global change in temperament and emotionality, activities and social behaviour, and physical capabilities. The MPAI-IV produces a total score, and three rationally derived subscale scores labelled Ability, Activity, and Participation. The MPAI-III was originally proposed and used to collect data for this study. However, Malec et al. (2003) conducted a thorough psychometric analysis and revision of the MPAI-III to improve the scale's psychometric properties. Malec & Lezak (2003) provide a recoding system to transform the MPAI-III to the MPAI-IV. All analyses performed in this study were recalculated to take advantage of the improved psychometric characteristics of the MPAI-IV. The MPAI-IV full scale internal consistency was reported as satisfactory, with Rasch Person Reliability =.88, Person Separation = 2.68, Item Reliability = .99, and Item Separation = 10.80. Cronbach's alpha for the full scale was .89. Person reliability for the three subcales ranged from .78 to .79, Item reliability from .98 to .99. Cronbach's alpha for the three subscales ranged from .76 to .83. The subscales were found to be moderately

intercorrelated (Pearson r = .49 to .65), and correlated strongly with the MPAI-IV total scale (Pearson r = .82 to .86) (Malec & Lezak, 2003).

Supervision Rating Scale (SRS) (Boake, 1996).

The Supervision Rating Scale provides a measurement of all forms of help that would require a caregiver to be in the physical vicinity of the patient. The patient is then allocated a numerical rating from "1" – independent, to "13" – in physical restraints. The rating is completed by a clinician on the basis of interviews with the patient and an informant who knows what supervision is actually received presently. Little reliability data is available for the SRS. Interrater reliability published by the scale author was satisfactory (intraclass correlation = .86, weighted Kappa = .64). The author also reported strong convergent validity between the SRS and ratings on the Disability Rating Scale and Glasgow Outcome Scale (Boake, 1996).

Examples of Selected Specific Assessment/Outcome Measures.

Agitated Behaviour Scale (ABS) (Corrigan, 1989).

The Agitated Behaviour Scale (ABS) provides an objective and rapid measurement system for assessing and monitoring important signs of agitation in patients that have sustained brain injury. The scale permits observational rating of 14 different signs of agitated behaviour on a four-point Likert scale (ranging from "1" – absent to "4" – present to an extreme degree).

Beck Depression Inventory – 2 (BDI – II) (Beck, Steer & Brown, 1996).

The Beck Depression Inventory – 2nd Edition (BDI – II) is a measure for evaluation of depression. Although many of the domains assessed by this measure may also tap

brain injury symptomatology (e.g., crying, agitation, concentration, etc.), it provides a useful tool to guide the clinician's evaluation of depression, and provides a converging line of evidence if depression is suspected.

Brief Psychiatric Rating Scale (BPRS) (Overall & Gorham, 1962). The BPRS is a widely-used seven-point Likert-scale measure evaluating 24 symptoms commonly associated with psychiatric disturbance.

Brief Test of Intellectual Functioning (BTIF) (Hurwitz, 2001). The BTIF provides a brief screening of higher intellectual functions that is similar to the mini-mental Status Examination (MMSE) but provides expanded assessment. The BTIF can be completed in 15 minutes and includes assessment of orientation, verbal and non-verbal immediate, remote and forced-choice recognition memory, calculation, drawing/constructional ability, repetition, naming and receptive language functions. It was selected in preference to the MMSE for its broader screening of cognitive functions.

Community Integration Questionnaire (CIQ) (Willer et al., 1993). The CIQ is typically conducted via interview or by patient self-rating and produces a global score ranging from 0 to 29, where a high score indicates greater community integration. Although researchers have noted ceiling effects (Hall et al., 1996), these are unlikely to be of concern when using the CIQ to assess changes towards greater community independence in an in-patient rehabilitation population.

O-Log (Jackson & Novack, 1995).

The O-Log is designed to permit ongoing assessment of orientation to place, time and circumstance. The items are each scored on a four-point Likert scale (permitting rating of "spontaneous," "cued," "recognition," and "no correct response" answers). The O-Log produces a summation score of orientation that ranges from 0 to 30 and can be useful for monitoring changes in orientation during acute recovery, and for ongoing monitoring of rehabilitation success in patients with difficulties in cognitive functioning, and in patients who are confused or delirious.

Overt Aggression Scale – Modified for Neurorehabilitation (OAS – MNR) (Alderman, Knight & Morgan, 1997).

The OAS – MNR was adapted from the Overt Aggression Scale (Yudofsky et al., 1986) to provide a standardized instrument for measuring aggressive behaviour in traumatic brain injury, and to provide a structure for information collection that facilitates the completion of a functional analysis of aggressive behaviours in the individual patient (Alderman et al., 1997). The OAS – MNR provides for rating of four aspects of aggressive behaviours using a four-point Likert scale. Behaviours are rated in the areas of verbal aggression, physical aggression against objects, physical aggression against self, and physical aggression against other people.

Scaling individual goals and outcomes.

Goal Attainment Scaling (GAS) (Kiresuk & Sherman, 1968).

SRC implemented GAS using a 1-5 scale (rather than the -2 to +2 in the original system). This removed the association of negative ratings with improvement and

permitted the inclusion of a "0" rating to designate "no gain" or worsening of behaviour. Every six weeks the four most urgent goals as established by clinical team consensus were scaled for each individual.

Results

Data Screening

Prior to analysis, measures used were examined for correct entry of data, any missing values, and to ensure that the properties of the variables met with the assumptions of analyses. The data set was complete for all participants. No outliers were identified (Tabachnick & Fidell, 1996). Some of the variables were not normally distributed, showing some deviation in terms of skewness and/or kurtosis.

Descriptive statistics: MPAI-IV, SRS & ImpGAS

Descriptive statistics for the MPAI-IV, SRS and ImpGAS are provided for the participant sample (see Table 3).

Table 3.

ImpGAS, MPAI-IV & SRS Mean, SD and Range

Scale	Mean	SD	Range
ImpGAS	1.76	1.07	0 – 3.53
Intake			
MPAI-IV Total	59.38	16.73	27 – 95
MPAI-IV Abilities	17.69	9.97	5 – 42
MPAI-IV Adjustment	26.19	6.13	14 – 35
MPAI-IV Participation	25.56	4.18	15 – 30
SRS	7.88	1.31	4 – 10
Discharge			
MPAI-IV Total	51.63	19.85	15 – 87
MPAI-IV Abilities	15.88	10.28	1 – 38
MPAI-IV Adjustment	21.50	7.23	8 – 32
MPAI-IV Participation	23.19	6.01	12 – 30
SRS	6.56	2.25	1 – 8
Difference Scores			
MPAI-IV Total	7.75	8.66	-7 – 31
MPAI-IV Abilities	1.81	2.32	-3 – 5
MPAI-IV Adjustment	4.69	4.91	-4 – 18
MPAI-IV Participation	2.38	2.96	0 – 10
SRS	1.31	1.74	0 – 6

Note, n= 16.

Demographic Variables

The bivariate correlations between ImpGAS, SRS and MPAI-IV and the demographic variables, collected from participant files, are reported in Table 4.

Table 4.

Bivariate correlations: ImpGAS, SRS, MPAI-IV and demographic variables

Variable	Age	Education	Gender	Severity	Time Post-injury	Months in rehabilitation
ImpGAS	-.39	-.29	.10	-.46	-.07	-.11
SRS intake	.11	-.19	-.21	.58*	.18	.10
SRS discharge	.23	.14	-.05	.86**	.32	.06
SRS difference	-.22	-.32	-.09	-.67*	-.28	.002
MPAI-IV intake	.04	-.22	.02	.60*	.60*	.24
MPAI-IV discharge	.08	-.16	.01	.66**	.52*	.09
MPAI-IV difference	-.10	-.07	.01	-.34	-.03	.25

Note, $n = 16$. Two-tailed p's *$p < .05$, **$p < .01$

SRS intake, SRS discharge, and SRS difference scores were significantly correlated with severity of injury. The MPAI-IV intake and discharge scores were significantly correlated with severity of injury and with time post-injury. There were no significant correlations between ImpGAS, SRS, or MPAI-IV in terms of age, education, gender, or number of months in rehabilitation. Consideration was given when the study was proposed to controlling for the participants' program duration, as it could translate into greater opportunity to achieve improved scores on the discharge outcome measures. However, ImpGAS provides some control of program duration since it represents mean improvement over programming cycles. In addition, as seen

in Table 4, months in program was not significantly correlated with any of the outcome measures. As a result, it was not included as a covariate in any of the subsequent analyses.

Analyses of ImpGAS and the SRS and MPAI-IV

Bivariate correlations: ImpGAS, and the MPAI-IV and SRS difference scores

The primary hypothesis in this study was that improvement in global outcome (SRS, MPAI-IV difference scores) could be predicted by small incremental changes in individualized goals (ImpGAS). First, the bivariate correlations between ImpGAS and the MPAI-IV and SRS difference scores were computed (see Table 5).

Table 5.

Correlations Between ImpGAS, and the MPAI-IV and SRS difference scores

Variable	1	2	3
1. ImpGAS	--		
2. MPAI-IV Difference Score	.724**	--	
3. SRS Difference Score	.617*	.585*	--

Note, $n = 16$. Two-tailed p's *$p < .05$, **$p < .01$

There were significant correlations between the ImpGAS score and the MPAI-IV and SRS difference scores. In addition, there was a significant correlation between the two outcome measures (difference scores). Nonparametric calculation of the bivariate correlations among the same variables produced similar findings to parametric techniques. The zero-order correlations suggest that the mean

improvement on individual goals as measured by the ImpGAS score was significantly correlated with improvements in outcome as measured by the MPAI-IV and SRS change.

The relationship between incremental change (ImpGAS) and measures of global change (MPAI-IV, SRS)

The relationship between incremental change and global change was evaluated first by separately regressing the SRS and MPAI-IV difference scores on ImpGAS. The bivariate correlations reported in Table 5 are significant and therefore the bivariate regressions will be significant as well. The purpose of performing the regressions was to obtain the adjusted R^2 and regression weights relating each global outcome measure to ImpGAS. The adjusted R^2 is important given the small sample size, and the regression weight provides details of the amount of change in each global measure associated with incremental change in ImpGAS. The prediction of the SRS difference score with ImpGAS was significant ($B = 1.003$, $p = 0.01$, adjusted $R^2 = .336$). The prediction of the MPAI-IV difference score with ImpGAS was also significant ($B = 5.855$, $p < .01$, adjusted $R^2 = .490$).

The relationship of ImpGAS to MPAI-IV change over and above SRS change

As stated earlier, the SRS difference score is a focused measure of change, while the MPAI-IV difference score is a broad-based measure capturing more global aspects of change. Therefore, the SRS was treated as a control variable, and change in global outcome as measured by the MPAI-IV difference score, after covarying the SRS difference score, was examined as predicted by ImpGAS. This hypothesis was

evaluated by regressing the MPAI-IV difference scores on ImpGAS, with SRS difference scores included as covariates. An overall relationship between the MPAI-IV difference scores and the predictor variables was found ($F(2, 13) = 8.099$, $p < .01$, adjusted $R^2=.486$). Examination of the relative contribution of ImpGAS to prediction of the MPAI-IV difference scores revealed that ImpGAS was a significant predictor of MPAI-IV difference scores, and accounted for a significant amount of unique variance in the MPAI-IV difference scores ($t(14) = 2.491$, $p < .05$, $B = 4.379$ (1.902), partial correlation = .568) above and beyond the variance associated with SRS change.

Analyses of ImpGAS and the SRS and MPAI-IV subscales

The MPAI-IV total score is derived from scores on three subscales that evaluate functioning under three broad areas: abilities; adjustment; and participation. Analyses of ImpGAS and the SRS and MPAI-IV change scores were conducted in the same way as analyses in the previous section, except that the MPAI-IV subscale difference scores were used rather than the MPAI-IV total change score.

The relationship of SRS change to MPAI-IV subscale score change

The difference scores computed for the three subscales were examined individually; first, in terms of their relationship with the measure of focal change (SRS difference score). Regression of the individual subscale change scores of the MPAI-IV on SRS change was performed (see Table 6), and revealed a moderate degree of relationship between SRS change and the three MPAI-IV subscale change scores.

Table 6.

Prediction of the MPAI-IV subscale difference scores by SRS change.

Variable	B	Std. Error	r^2	t	Significance
Abilities	.725	.298	.545	2.431	.029
Adjustment	1.333	.664	.473	2.006	.065
Participation	.949	.378	.557	2.512	.025

Note, $n = 16$.

The prediction of the MPAI-IV Abilities and Participation change scores by SRS change was significant, while the prediction of the Adjustment change score by SRS change approached significance.

The relationship of ImpGAS to MPAI-IV subscale change over and above SRS change

Further analyses were performed to examine the relationship between ImpGAS and the individual difference score subscales of the MPAI-IV. First, each subscale difference score was separately regressed on ImpGAS (see table 7).

Table 7.

Prediction of the three MPAI-IV subscale difference scores by ImpGAS

Variable	B	Std. Error	r^2	t	Significance
Abilities	1.490	.419	.689	3.556	.003
Adjustment	3.170	.886	.691	3.579	.003
Participation	1.847	.551	.667	3.350	.005

Note, $n = 16$.

The results revealed that all three of the MPAI-IV subscale change scores were significantly predicted by ImpGAS.

The prediction of the MPAI-IV subscale difference scores by ImpGAS was then examined by regressing the subscales on ImpGAS while including SRS change as a covariate (see Table 8).

Table 8.

The relationship of ImpGAS to MPAI-IV subscale change over and above SRS change

	Overall Model		ImpGAS as a Unique Predictor				
Variable	Adj. R2	Significance.	B	Std. Error	Partial r^2	t	Significance
Abilities	.421	.011	1.232	.540	.535	2.281	.040
Adjustment	.401	.014	2.959	1.164	.576	2.542	.025
Participation	.399	.014	1.445	.704	.495	2.052	.061

Note, n = 16.

The combination of SRS change and ImpGAS significantly predicted all three subscales. After controlling for the variance associated with SRS change, ImpGAS accounted for a significant amount of unique variance in the MPAI-IV Abilities difference score and the MPAI-IV Adjustment difference score, while prediction of the MPAI-IV Participation difference approached significance.

The relationship of ImpGAS to each MPAI-IV subscale independent of the remaining subscales

As reported in Table 7, ImpGas was significantly related to all three MPAI-IV subscale change scores. In order to determine whether any of the subscales shared a unique relationship with ImpGAS, each subscale was regressed on ImpGAS with the remaining two MPAI-IV subscales included in the regression equation as covariates (see table 9).

Table 9.

The unique relationship of each MPAI-IV subscale difference score with ImpGAS

Variable	B	Std. Error	Partial r^2	t	Significance
Abilities	.746	.562	.357	1.326	.210
Adjustment	.549	.996	.157	.552	.591
Participation	.518	.661	.220	.783	.449

Note, $n = 16$.

When each of the MPAI-IV subscale difference scores was examined with the variance associated with the remaining two MPAI-IV subscale difference scores entered as covariates, none of the subscales was significantly predicted by ImpGAS.

Logistic regression prediction of severity

It was hypothesized that treatment durations, ImpGAS scores, MPAI-IV Total Difference Scores, and supervision level (SRS) might reasonably predict severity of injury. Logistic regression was used to examine the capacity of ImpGAS, SRS difference score, MPAI-IV difference score, and months in program to predict severity

of injury (mild-moderate versus severe). Because the MPAI-IV and SRS administered at intake and at discharge were found to be correlated with severity (see Table 4), and since these measures provide a direct evaluation of the participants functioning at a fixed point in time, they might also be expected to be predictive of severity of injury. As such, additional analyses of their prediction of severity level were also included. However, due to limitations in the sample size, and the fact that only three participants were classified in the mild-moderate severity range, the logistic regression analyses were conducted one predictor at a time (see Table 10).

Table 10.

Prediction of Severity Group (Mild-Moderate versus Severe Injury)

Variable	Odds Ratio (-2 Log)	B	Std. Err.	Wald	Significance
ImpGAS	11.602	-1.544	.978	2.489	.115
MPAI-IV Intake	6.481	.272	.180	2.290	.130
MPAI-IV Discharge	6.197	.187	.109	2.938	.087
MPAI-IV change	13.728	-.098	.080	1.498	.221
SRS Intake	9.480	2.327	1.704	1.864	.172
SRS Discharge*	---	---	---	---	---
SRS change	8.836	-1.119	.607	3.397	.065
Months in Program	14.485	.238	.253	.884	.347

Note. n = 16. *Maximum iterations reached without solution

None of the variables evaluated significantly predicted severity of injury. Change in supervision level (SRS change) and MPAI-IV discharge score approached

significance and classified all but one of the 16 participants into their correct severity group (13/13 Severe and 2/3 mild-moderate correctly classified). Due to limitations noted above, combinations of predictors could not be evaluated.

Discussion

The primary purpose of this study was to evaluate the widely held belief among brain injury rehabilitation professionals that working towards small incremental goals is an efficient and meaningful way to achieve larger long-term goals and improvements in global rehabilitation outcome. Demonstrating that a rehabilitation program or service is making improvements in long-term goals and global outcome is important for a number of reasons. Most importantly, it is a reflection on the overall rehabilitation program's effectiveness in doing its job - in this case, helping people to recover from the sequelae of brain injury. Secondly, intensive brain injury rehabilitation services are often highly labour-intensive, and invariably very expensive. It is important to be able to demonstrate to survivors of brain injury, their family members, and program funders, that your methods of performing rehabilitation work, and are worth the financial investment. Particularly in today's restrictive financial climate, which is characterized by service reductions and increasing health care costs, being able to demonstrate the effectiveness of a rehabilitation program is becoming essential. There is an increasing expectation that health-care delivery programs demonstrate the effectiveness of their approach, typified by the movement towards evidence-based practice. The onus falls on the programs

themselves to either demonstrate that their practices are effective, or risk loss of funding, imposed program restructuring, and/or imposition of practice methods.

The implementation of a unified system of goal setting and outcome measurement at Skeleem Recovery Centre (SRC) was undertaken for two reasons: to address the need to demonstrate the effectiveness of the SRC rehabilitation program, and to provide greater organization, systematization, and streamlining to clinical goal setting and rehabilitation intervention. Goal Attainment Scaling (GAS) was selected as the tool that would be used to organize and monitor individual goal setting. GAS provided a means to produce a common numerical metric across persons and goals that could accommodate the heterogeneity of the brain injured population and their diverse rehabilitation needs. It also provided a structured way to assign key personnel, measurement instruments, and specific interventions to rehabilitation goals. It allowed for prediction of expected improvement and a cyclical way for the SRC clinical rehabilitation team to follow up on the effectiveness of the interventions employed.

Relationship between ImpGAS, MPAI-IV and SRS with demographic variables

The significant correlations between the SRS (intake, discharge, difference) and the MPAI-IV (intake, discharge) and severity are in keeping with the reasonable expectation that measures of participants' functioning would differ according to severity of brain injury. The significant correlation between the MPAI-IV (intake, discharge) and time post-injury may reflect both degree of spontaneous recovery and amount of compensatory adaptation to deficits.

Relationship between ImpGAS and the MPAI-IV and SRS difference scores

The primary hypothesis of this study was that improvement on specific incremental individual goals as quantified by the average GAS improvement score (ImpGAS) would predict improvement in the MPAI-IV and SRS difference scores (subtracting intake scores from discharge scores).

All three measures were intercorrelated. This is not entirely surprising given that all three measures quantify improvement in functioning. The strong association between ImpGAS and the MPAI-IV difference score indicates a good relationship between improvement on small goals and global change. In addition, there was a moderate correlation between the SRS and MPAI-IV change scores. This provides some convergent validity between the two measures of global outcome. In addition, since the SRS change score was treated as a baseline measure of focused change, while the MPAI-IV change score was the measure of more global outcome, demonstrating that the MPAI-IV change score was predicted by incremental change over and above supervision requirements provided a more stringent test of whether incremental improvement predicts improvement in more global aspects of outcome.

It was hypothesized that incremental change (ImpGAS) would predict change on the two outcome measures. This hypothesis was supported, as both the MPAI-IV and SRS difference scores were significantly predicted by ImpGAS.

As discussed earlier, the SRS was used as a control variable. The SRS score at intake had little variability since most participants were admitted to SRC due to their need for at least 24-hour indirect supervision. The results indicated that the MPAI-IV

difference score was significantly predicted by ImpGAS, above and beyond variance associated with the SRS difference score. This speaks both to the strength of the prediction of the MPAI-IV change by ImpGAS, and to the fact that the MPAI-IV captures broader aspects of functioning, including aspects of functioning that relate to degree of supervision and level of independence, that overlap the SRS.

The implication is that the MPAI-IV's coverage is broad enough that an additional measure of degree of supervision required may not be necessary for evaluation of global outcome. However, the choice of measures depends somewhat upon the program's requirements. The reason the SRS was selected for use at SRC's rehabilitation program was that, in the opinion of the social worker and discharge planning team, the single most important factor for discharge to a less restrictive setting was improvement with respect to the level of supervision the person required. Degree of supervision required has a direct relationship both to the choice of discharge environments available (e.g., ranging from a locked facility, to a proprietary care home, to a boarding home, to semi- or independent living), and to reduction in care costs attributable to the success of the rehabilitation program. There are few things that are more directly tied to future cost of care than level of supervision required. An additional benefit of the SRS is its simplicity. It has thirteen possible supervision levels, each of which is well operationalized, and which can be identified at a glance. The measure takes less than a minute to complete. This may make it appealing for inclusion as a supplemental measure. The major weakness of the SRS is that it lacks the rich degree of information captured by the MPAI-IV. It would be hard to justify

using it as a sole measure of outcome, whereas such an argument could be made for the MPAI-IV with its broad coverage of sensory and motor function, cognitive abilities, mood, pain, fatigue, social interaction, awareness, initiation, relationships, self-care, leisure, transportation, money management, residence, etc.

The MPAI-IV captures all of the important domains of interest to most rehabilitation programs. Unlike the SRS, another clinician reviewing the MPAI-IV completed on a brain-injured program resident would have an excellent sense of the person's overall functioning, and of their abilities in almost all important domains related to program requirements, areas of strength and weakness, and environmental needs. An experienced brain injury rehabilitation professional could construct a preliminary rehabilitation program based on the MPAI-IV information alone. In its latest refinement, the MPAI-IV proved to be an extremely useful tool at SRC for monitoring all aspects of global outcome. The main disadvantage of the MPAI-IV compared to the SRS was the time required to complete the instrument – approximately 30 minutes of team consensus discussion. Although it requires intensive knowledge of the person about whom it is being completed, this is not inherently a liability of the measure, as it would be improper, unethical and likely ineffective to design and implement an individualized brain injury rehabilitation program without such knowledge.

The general finding that improvement between intake and discharge as measured by the MPAI-IV difference score was significantly predicted by improvement on small incremental goals, as measured using ImpGAS, is important. It

provides preliminary research support for the widely held belief in brain injury rehabilitation that deconstructing long-term goals into smaller more manageable units of progress, and working to achieve incremental improvement via these small units, is associated with improvements in global outcome. What this means is that in this study, participants who made better progress towards small incremental goals reflected that progress in their global outcome as measured by the MPAI-IV. Although this may seem like a somewhat obvious conclusion, there is scant empirical evidence for this anecdotal belief. In part, this is due to the fact that the brain injured population is highly heterogeneous, resulting in widely varying challenges and rehabilitation needs. The net result of this heterogeneity is that goals tend to be quite dissimilar from person to person. Thus, quantifying improvement beyond the individual level has proven difficult. This study represents a preliminary attempt to place individual rehabilitation goals on a common metric using GAS.

Investigation of the Three MPAI-IV Subscales

An examination of the relationship between the three subscales of the MPAI-IV: Adjustment, Abilities and Participation, and average amount of improvement as measured by ImpGAS revealed that although ImpGAS was significantly predictive of improvement on each subscale when investigated individually, no single subscale was a significant predictor over and above the variance associated with the other two subscales. This suggests the possibility that overall improvement as indicated by the total measure was composed of different patterns of subscale improvement among different individuals. This indeed proved to be the case with 10 participants showing

their highest degree of improvement on the Adjustment subscale, one on the Abilities subscale, and two on the Participation subscale. In addition, two participants tied for most improvement on the Abilities and Adjustment subscales, and one participant tied for most improvement on the Adjustment and Participation subscales.

However, an alternative explanation, given the high degree of intercorrelation of the three subscales, is that the MPAI-IV total change score represented a unitary factor related to global outcome from TBI. Because the subscales were rationally derived - created based on clinical experience and for purposes of clinical utility - rather than psychometrically, these results are not necessarily counter to previous findings. In fact, Malec et al. (2003) identified that they and other members of the TBI Model Systems programs have found in their own samples that the MPAI-IV appears to have one primary underlying factor representing global TBI outcome.

Predictors of Brain Injury Severity

Logistic regression was used to test the hypothesis that degree of improvement measured by ImpGAS score and MPAI-IV scores, SRS scores, and more months in the treatment program, would be predictive of severity of injury, with poorer scores reflecting more severe injury. Due to the overall limitation in sample size, and the additional limits imposed by having only three participants in the mild-moderate group, the predictors could not be entered in combination. Rather, the predictive value of each individual variable was examined individually. None of the variables examined significantly predicted group membership. However, the SRS change score and the MPAI-IV discharge rating both approached significance, and might reasonably

be expected to be significant predictors were there even a modest increase in sample size. Both of these variables independently produced correct classification of 13/13 severe brain injuries and 2/3 mild-moderate brain injuries. Since combinations of variables could not be analyzed, it is impossible to know whether a combination of these two predictors, or of additional non-significant predictors, might have correctly classified all participants by severity.

It is interesting that change in supervision level approached significance, while admission supervision level did not. The reason for the poor predictive value of intake supervision level may be due to bias produced by the selection of program participants at SRC. The program is for individuals who require residential inpatient services. As such, they are selected at intake for, among other things, their need for a certain degree of supervision that cannot be accomplished in community outpatient settings. Because of this, they are likely to be scored similarly in terms of supervision requirements at intake, regardless of injury severity, as a result of this selection process. In contrast, the difference score reflects the degree of improvement made between intake and discharge, which might be expected to be a superior discriminator of severity as those with milder injuries may be more likely to be discharged to less restrictive settings after treatment.

Similarly, the fact that the MPAI-IV score at discharge approached significance, while the MPAI-IV difference score did not, may be explained by the properties reflected in an unadjusted scale score versus a difference score. The unadjusted MPAI-IV score reflects a comprehensive summary of numerous abilities

and challenges faced by the participant at program discharge. Since this is a direct measure of many aspects of global functioning, it should be a good predictor of severity, since greater severity generally means more challenges, and therefore more areas of concern that would be noted on the instrument. In contrast, the difference score calculated between intake and discharge on the MPAI-IV reflects something quite different – amount of improvement in ability. For this reason, it is less probable that the MPAI-IV difference score would contribute meaningfully to the discrimination of persons with mild-moderate injuries versus severe injuries. They may improve similar amounts, even though their intake level of functioning may be substantially different.

In summary, the logistic regression analyses were conducted very tentatively as sample limitations restricted the analyses to single predictors. The analyses did suggest that the change in supervision requirements from intake to discharge, and the MPAI-IV administered at discharge, both approached significance and correctly classified 15/16 participants. In a larger sample they might reasonably be expected to be significant predictors of severity.

Selection of Outcome Measures

The global outcome measures used at SRC were selected on the basis of three major criteria: that they covered the domains of interest in outcome; that they were widely used and empirically supported; and that they were available in the public domain. A number of measures exist that fulfill some or all of these criteria, including the Glasgow Outcome Scales, FIM/FAM, CHART, etc. However, no consensus exists

among rehabilitation programs in terms of the measures selected (Pender & Fleminger, 1999). One of the recognitions in selection at SRC was that unification of the tools used by brain injury programs would be extremely valuable both in terms of standardization of measurement of global outcome, and the ability to directly compare outcome data with other programs.

The Center for Outcome Measurement in Brain Injury (COMBI) collaborative project of model system centers in the United States provides internet-based resources for outcome measurement, including rating scales and forms, administration guidelines, descriptions of scale properties, supportive references, training and testing materials, and contact information (Santa Clara Valley Medical Center, 2004). This represents one of the first efforts to unify centres of excellence in brain injury rehabilitation around a core battery of well-validated assessment tools. This was particularly useful in assisting selection of outcome measures at SRC. The COMBI site provided scale forms, interpretive information/manuals where available, and reference information for a variety of useful and well-validated outcome measures. Further, it provided information about the model systems programs that had already adopted these measures for outcome evaluation in their brain injury rehabilitation programs. In addition, most of the measures also fell within the public domain. Overall, the selection process was made far easier by the existence of the COMBI site, and the team selected outcome measures with some confidence about their psychometric properties. Most importantly, they were used with the knowledge that

other rehabilitation sites were also adopting similar measures to evaluate their program outcomes.

As data from these sites begin to be published it is hoped that direct comparisons can begin to be made between sites using the same set of outcome instruments. The best possible result would be that the data produced by the COMBI and model systems outcome measurement projects results in a common set of guidelines and outcome measures recognized and used as examples of best practice in brain injury rehabilitation outcome measurement. The more the information is disseminated, and the more sites that adopt the COMBI measures, the greater the benefit to rehabilitation professionals, program administrators, program funders, survivors of brain injury and their families. It will enable comparison of efficacy of various interventions across common metrics, which will aid in the refinement of the most cost effective and rapid rehabilitation techniques. It can also help to eliminate practices that are broadly implemented but relatively ineffective. It will, over time, permit meta-analyses of outcomes to be analyzed increasing the power of research on this highly heterogeneous population with the ultimate possibility of a unified set of interventions that have been found most effective across multiple settings and multiple brain injury sequelae.

Considerations Regarding the Outcome Measures

Measurement of outcome is increasingly being seen as a necessity to justify the high cost of brain injury rehabilitation efforts. However, most of the research that is produced around brain injury outcomes used clinical team members to complete the

outcome measures. This is unavoidable in almost all cases as completion requires an intimate knowledge of the range of abilities and challenges that might be addressed for a given survivor of brain injury. Further, there are few if any programs that have the resources to identify blind raters who know the survivor clinically, but have no affiliation with their care and rehabilitation. However, it must be acknowledged that unblinded raters who are invested in the rehabilitation process may introduce bias when rating a survivor's level of improvement after intervention. For this study we used team consensus on the primary outcome measures (MPAI-IV, SRS) to attempt to moderate the influence of any one rater. In addition, the admission ratings were not reviewed by any team member prior to the discharge ratings. Thus, although the possibility of introducing some bias at discharge cannot be completed mitigated, it was felt that the team ratings were conducted in as objective a way as possible, and that bias was minimized.

There are also issues related to the outcome measures selected. Although empirical support for various outcome measurement instruments is gradually accumulating, there still remains a lack of consensus about what measure(s) represent the gold-standard for outcome measurement. A by-product of this lack of unification is that information on the reliability and validity of the numerous measures available tends to be somewhat scarce. The measures selected for use in this study represent two ends of the spectrum in terms of available data around their respective psychometric properties: the MPAI-IV having undergone revision through advanced scale analysis procedures, while the SRS is supported by only some initial inter-rater

reliability and convergent validity data from its original publication source. The numerous other measures available are in various stages of validation, most falling somewhere between the two outcome measures selected for this study in terms of the psychometric data that is available.

The Use of Difference Scores

Where there is often scarce reliability and validity data for brain injury outcome measures in general, there is almost a complete absence of empirical data about the reliability of difference scores used in predicting brain injury outcome. Most studies of existing brain injury outcome measures attempt to use the scale score(s) to predict some important element(s) of outcome (e.g. employment status). Other measures use ratings by the survivor of brain injury and some "other" rater (e.g., staff member, family member) to create difference scores that are presumed to be underpinned by differing perspectives of the various parties, and discrepancies between a survivor and "other" rating have been argued to indicate a lack of awareness of deficits on the part of the survivor - assuming the survivor is underestimating their deficits compared to the other rater (Bogod, Mateer & MacDonald, 2003). However, the reliability of difference scores is not reported on. A limitation of difference scores, which must be considered in terms of this study, is that repeating a measure that has anything less than 100% reliability will increase the unreliability of the result (May & Hittner, 2003). In other words, the degree of error in measurement will be duplicated each time the measure is repeated. In the case of the MPAI-IV, the published data suggests a high degree of reliability, so one could argue

that administering it on two occasions should not dramatically influence the overall reliability of the data produced. The SRS, however, has very limited available reliability data, so if its initial reliability is weak, this would be compounded by creating a difference score using two administrations. Since the MPAI-IV was used as the primary instrument, with the SRS as a control variable, the data produced in this study is not thought to have sustained a substantial negative effect to its reliability via repeated measurement. It is also important to note that the effect of attempting to correlate a difference score with a third variable of interest will typically underestimate the degree of relationship between the third variable and the change score (Raykov, 1999). Essentially this will reduce the likelihood of significant findings in studies where change scores are used to predict other variables. This may result in increased Type II errors (failure to reject an incorrect null hypothesis), but suggests that relationships that produce significant findings under these circumstances may have been subjected to a more stringent statistical test of their relationship, and the resultant significant findings are certainly not invalidated as a result.

Finally, it is important to note that the measurement of clinically meaningful change in a systematic and quantified way is difficult to achieve without measuring some characteristic(s), intervening, and then measuring the characteristic(s) again to look for improvement. Alternatives include outcome indicators such as return to work status, but these are confounded by general effects such as increasing time post-injury, and the prevailing work climate at the time of discharge. Only through measurement of change in specific functional areas can we: elucidate which interventions are

effective; identify the most effective interventions for certain challenges; and ensure change in the areas of function that have been established as the best predictors of targeted outcome indicators (e.g. return to work).

Considerations in the use of Goal Attainment Scaling

Goal Attainment Scaling in the format it was implemented at SRC was found to be extremely useful for managing individual brain-injured resident's goals. Clinical team meetings prior to the implementation of GAS were relatively haphazard, and goals that were set were often poorly implemented and unreliably followed-up. The team often felt reactive (e.g., as behavioural emergencies emerged) versus proactive. GAS had a number of benefits for the SRC clinical team. Firstly, it brought the clinical team together on a consistent meeting schedule with an agenda to set and follow-up goals for specific program participants – at SRC we found a six-week cycle of follow-up most manageable in terms of frequency of meetings and patient load. Secondly, it ensured that team members were in consensus around the goals for each person reviewed. Third, it required the identification and use of specific outcome measurement instruments. Fourth, it included assignment of the person responsible for data collection and follow-up with the outcome measurement instrument. Fifth, it demanded identication of a specific intervention. Finally, by requiring the team to anticipate expected improvement levels given the intervention and current level of measured functioning, it forced accountability – if the participant did not improve as expected, the team was forced to consider whether this was due to an ineffective intervention, poor implementation, insensitivity of the measurement instrument,

overestimation of probable outcomes, and/or insufficient time for the intervention to be effective. Overall, it brought a very high level of organization to the entire process of rehabilitation at SRC. Producing quarterly reports to funders was far easier, as goals, interventions and progress over that quarter were all fully documented by the GAS process. Overall, the entire SRC clinical team was pleased with the level of organization, the immediate access to outcomes of individual goals, and the accountability provided by the GAS process.

GAS also has a number of limitations and potential pitfalls. Firstly, GAS is only as effective as the quality of the goals set, the quality of the outcome measures, the effectiveness of the interventions, and the reliability of follow-up. Secondly, the idea that very disparate goals can be unified across a common group metric for group comparisons could be problematic in that some goals may be more important than others. The GAS process does provide a mechanism by which goals can be weighted, but this is essentially an arbitrary decision made by the team using GAS. Other professionals could differ considerably in the weightings applied to certain goals, or select completely different goals. At SRC, the use of weightings was not included as it was felt it added an additional level of complexity that might be a deterrent to clinical staff when the GAS process was proposed and explained to clinical team members prior to implementation. An additional consideration around converting all goals to a common metric for group comparison is that it represents a substantial collapsing of data, with the possibility that considerable amounts of meaningful variance are lost in the process. Another drawback of GAS is that it is quite labour

intensive for clinical team members. It required frequent meetings of the full clinical team, creation of two new programming positions for oversight of interventions and data collection, identification and measurement of all goals with appropriate instruments, and continuous follow-up and retuning of goals and interventions. Overall it proved to be considerably more time consuming than the previous, more reactive method of dealing with a resident's most pressing concerns as they came up. However, it was considerably more effective for goal management.

The final, and possibly largest consideration, relates to the estimation/projection of expected outcomes by the clinical team. There are two considerations in this process. The projection of how much a person should improve in a certain area of functioning given a certain intervention and a certain time frame is somewhat arbitrary. Initially it was difficult for the team to agree on how much improvement might be expected. Over time, the team became better at projecting fairly accurate outcomes. In some ways, being inaccurate in expected improvement estimates at follow-up helped the clinical team at SRC to refine their clinical projections, and to increase their knowledge about how quickly various areas of functioning recover given a specific intervention.

The other consideration is that the GAS process may be subject to bias. Because clinical team members (who are presumably invested in demonstrating that their interventions produce favourable outcomes on the goals selected) identify the levels of expected improvement, there is some potential for underestimating expected outcomes to ensure a favourable GAS improvement rating when the goals are

reviewed. Although this is mitigated to some extent by using team consensus, and also by attaching explicit measures to evaluation of scaled goals, the potential still exists and is hard to fully control for within the framework of GAS without the added procedure of blinding goal evaluators to the scaling level of goals. Although goals were not reviewed before re-evaluation, evaluators were selected based on the persons judged most suited to assessment of the goal type (e.g., physiotherapy for range of motion) and the program did not have the resources or the level of professional redundancy that would have been required to have blinded examiners. It is likely that most brain injury rehabilitation programs are under similar constraints.

Challenges of Implementation

There were a number of challenges in setting up a new structured goal setting and outcome measurement system managed using goal attainment scaling at SRC. Researching and selecting rehabilitation systems was discussed earlier, but was a significant undertaking in and of itself. Once it was decided that we would use GAS to monitor the overall rehabilitation process, and the primary outcome measures had been selected, transition and implementation of the new system became a major focus. The transition process started with a commitment to the new model for delivering rehabilitation services by the author and the executive director of SRC. The next phase was preparation of guideline materials, and provision of a series of information sessions to clinical team staff describing the process and encouraging feedback. As a result, some refinements were made to the process including the number of goals that would be set for each participant, the degree of participant involvement in their goal

setting, and the nature of goal selection. Two practice sessions were then conducted performing the goal attainment scaling process for mock participants with the entire clinical team. An information package was produced to share with clinical support partners (e.g., BC Neuropsychiatry Program – who provided neuropsychiatric consultation via a weekly telemedicine program) and funders of the program. Seminars about the system and how it would impact the way rehabilitation was conducted at SRC were then presented to front line rehabilitation staff. It was agreed that the two primary front line staff for each participant would attend the goal attainment meetings for that person. In addition, an internal hiring competition was conducted to fill two programming positions with front line staff. This was necessary to assist with the additional coordination of numerous goals, outcome measures and interventions. A target date was then established for implementation, and each admission after that date was provided rehabilitation services using the new system.

Overall, the greatest challenge after developing the system was the amount of time and energy expended preparing clinical and front-line staff for the transition, to ensure acceptance. The entire process took longer than six months from inception to implementation at SRC, which is a small residential rehabilitation centre with a maximum of 24 in-patient beds. Although the system was well-regarded by SRC front-line, clinical and administrative staff, program funders, survivors of brain injury and their families, it was a considerable undertaking, particularly to do as a side project while still attempting to fulfill all pre-existing clinical duties. The major lesson from the process was to enlist as many colleagues as possible to the process, to

involve all personnel at the earliest possible juncture, and to delegate responsibilities to other team members.

General Considerations and Limitations of the Study

A major limitation of this study was the small sample size. The original intention when the research was undertaken was to collect data on a sample of at least 20 participants, and to supplement this sample with 10 truncated cases of those persons still awaiting discharge at the study's conclusion. Because of the loss of funding and closure of the rehabilitation centre, the sample size was reduced to sixteen participants. There was sufficient time to complete the rehabilitation process with the participants, but new admissions were frozen so no truncated participants were available for inclusion. The small sample size presented limitations in a number of areas including: the number and scope of statistical analyses that could reasonably be computed; the chance of missing important findings (Type II error) due to small sample size and the resultant need for large effect sizes to find significance. However, although there were analyses that could not be run (i.e., logistic regression with multiple predictors), and findings that only approached significance, the primary hypothesis that GAS improvement would significantly predict MPAI-IV change was supported. This suggests that this finding was quite robust in this sample and that the effect size was reasonably large. Thus, despite the more stringent requirements for significance imposed by a small sample, some results of significance were found in relationships of interest.

An additional limitation of this sample is the generalizability of the findings. As this was a sample of convenience representing serial admissions to a specialized in-patient residential treatment centre, the sample may not be representative of the larger brain injury population as a whole. One reason is that the chance that this sample differs even from other samples being treated in similar centres is increased by the small number of participants. A larger sample would be more likely to be representative of this specific subcategory of the brain-injured population. An additional reason is that SRC was a specialized centre and survivors who were admitted tended towards the more severe end of the brain injury spectrum, and also tended to have had complicated recoveries characterized by failure to benefit sufficiently from other less intensive rehabilitation opportunities. As a result, the findings for this sample may not be representative of either a larger similar sample, or of individuals with brain injury of equal severity but with less complicated recovery trajectories. The limits imposed by the small sample size both in terms of analyses computed and in terms of generalizability to other brain injury samples suggest that replication of this study would be a necessary and important step to confirm that the findings could be duplicated in a larger sample, and for the purpose of increasing the generalizability of the results to more diverse and inclusive subsets of the brain injured population.

The lack of longer term monitoring of outcomes through periodic follow-up is another limitation of this study. Although improvement may be apparent between intake and discharge from a rehabilitation program, sustained improvements are really

the ultimate goal of a successful rehabilitation program. Due to the time constraints inherent in research conducted to complete doctoral dissertation requirements, long-term follow-up of outcomes was neither proposed nor conducted. However, a longitudinal rehabilitation study of this type would be strengthened by re-administration of the global outcome measures at one- and possibly even five-year follow-up, if possible. This would ensure that improvements noted at the end of an intensive rehabilitation program were maintained, and that generalization of skills learned in a controlled environment were successfully transferred to different and/or less restrictive settings.

Future Directions

Having been intimately involved in the process and implementation of brain injury rehabilitation, and after evaluating the effectiveness of the goal setting and outcome measurement system at SRC, a step that, in retrospect, could have been included, is evaluation of the impact of improvement in global functioning on the quality of life (QOL) of survivors of brain injury and their families. Although some of the measures assess components that may be expected to relate to QOL (e.g., supervision level, employment status, ability to perform self-care, transportation issues), QOL was not assessed directly from the survivors of brain injury engaged in rehabilitation and participating in this study. In a recent comprehensive review of all published research on QOL in survivors of TBI, Dijkers (2004) found that indicators such as achievements and subjective well being (SWB) for persons with TBI were reduced from pre-injury levels, and lower than other comparison groups. Interestingly,

injury severity was not necessarily predictive of SWB. The author also reported that QOL utility measures (that capture aspects such as cognition and other health and life domains) are absent from the literature. The study recommends exploring the ability of survivors of TBI to self-report on their QOL, developing utility instruments sensitive to the differences in impact of varying degrees of cognitive deficits on QOL, and performing qualitative studies exploring QOL experiences. Future research of this type should include a self-report QOL instrument that evaluates subjective changes in aspects of QOL before and after rehabilitation. Improving the perceived QOL of survivors of TBI is typically thought of as a primary focus of rehabilitation efforts, but evaluation of QOL has lagged behind the recognition of its importance as an outcome indicator.

References

Alderman, N., Knight, C., & Morgan, C. (1997). Use of a modified version of the overt aggression scale in the measurement of aggressive behaviours following brain injury. *Brain Injury, 11*, 503-523.

Alderman, N., Davies, J.A., Jones, C., & Padraig, M. (1999). Reduction of severe behaviour in acquired brain injury: case studies illustrating clinical use of the OAS – MNR in the management of challenging behaviours. *Brain Injury, 13* (9), 669-704.

Ashley, J.M., & Krych, D.K. (1990). Consideration in the purchase of post-acute rehabilitative services for the head injured. *Journal of Medical Insurance, 22*(3), 230-234.

Bajo, A. & Fleminger, S. (2002). Brain injury rehabilitation: what works for whom and when? *Brain Injury, 16*(5), 385-395.

Banja, J. (1999). Patient advocacy at risk: ethical, legal and political dimensions of adverse reimbursement practices in brain injury rehabilitation in the US. *Brain Injury, 13*(10), 745-758.

Banja, J., & Moore, G. (1997). Shifting burdens in managed care: but to whom? *The Journal of Head Trauma Rehabilitation, 12,* 91-95.

Barrett, M., Wilson, R.J., & Long, C. (2003). Measuring motivation to change in sexual offenders from institutional intake to community treatment. *Sexual Abuse: Journal of Research & Treatment, 15*(4), 269-283.

Beck, A.T., Steer, R.A., & Brown, G.K. (1996). *Beck Depression Inventory*, 2nd Ed. San Antonio, The Psychological Corporation.

Becker, H., Stuifbergen, A., Rogers, S., & Timmerman, G. (2000). Goal attainment scaling to measure individual change in intervention studies. *Nursing Research, 49,* 176-178.

BIAA (2004). http://www.biausa.org/Pages/types_of_brain_injury.html.

Blendon, R.J., Donelan, K., Leitman, R., Epstein, A., Cantor, J.C., & Cohen, A.B., et al., (1993). Physicians' perspectives on caring for patients in the United State, Canada, and West Germany. *The New England Journal of Medicine, 328*(14), 1011-1016.

Boake, C. (1996). Supervision rating scale: A measure of functional outcome from brain injury. *Archives of Physical Medicine and Rehabilitation, 77*, 65072. Bond, M.R. (1976). Assessment of psychosocial outcome of severe head injury. *Acta Neurochirurgica, 34*, 57-70.

Bogner, J.A., Corrigan, J.D., Stange, M., & Rabold, D. (1999). Reliability of the Agitated Behavior Scale. *Journal of Head Trauma Rehabilitation, 14*(1), 91-96.

Bogod, N.M, Mateer, C.A., & MacDonald, S.W.S (2003). Measurement of awareness of deficits after traumatic brain injury. *Journal of the International Neuropsychology Society, 9*(3), 450-458.

Bohac, D.L., Malec, J.L., & Moessner, A.M. (1997). Factor analysis of the Mayo-Portland Adaptability Inventory: structure and validity. *Brain Injury, 11*(7), 469-482.

Borgoro, S.R., & Prigatano, G.P. (2002). Early cognitive and affective sequelae of traumatic brain injury: a study using the BNI screen for higher cerebral functions. *Journal of Head Trauma Rehabilitation, 17*(6), 526-534.

Bushnik, T. (2003). Introduction: the Traumatic Brain Injury Model Systems of care. *Archives of Physical Medicine and Rehabilitation, 84*, 151- 152.

Cardenas, D.D., Haselkorn, J.K., McElligott, J.M., & Gnatz, S.M. (2001). A bibliography of cost-effective practices in physical medicine and rehabilitation: AAPM&R white paper. *Archives of Physical Medicine and Rehabilitation, 82*, 711-719.

Carney, N., Chestnut, R.M., Maynard, H., Mann, N.C., Patterson, P., & Helfand, M. (1999). Effect of cognitive rehabilitation on outcomes for persons with traumatic brain injury: a systematic reviews. *Journal of Head Trauma Rehabilitation, 14*(3), 277-307.

Centers for Disease Control (1997). Traumatic brain injury—Colorado, Missouri, Oklahoma, and Utah, 1990-1993. *MMWR, 46(1),* 8-11.

Chestnut, R.M., Carney, N., Maynard, H., Patterson, P., Mann, N.C., & Helfand, M. (1998). Agency for health care policy and research evidence-based practice report. *NIH Consensus Development Conference on Rehabilitation of Persons with Traumatic Brain Injury*. Programs and Abstracts, pp. 73-82.

Collins, R., Lanham, R.A., & Sigford, B.J. (2000). Reliability and validity of the Wisconsin HSS Quality of Life Inventory in traumatic brain injury. *Journal of Head Trauma Rehabilitation, 15*(5), 1139-1148.

Corrigan, J.D. (1989). Development of a scale for assessment of agitation following traumatic brain injury. *Journal of Clinical and Experimental Neuropsychology, 11*(2), 261-277.

Corrigan, J.D., & Bogner, J.A. (1995). Assessment of agitation following brain injury. *Neurorehabilitation, 5*, 205-210.

Corrigan, J.D., & Bogner, J.A. (1994). Factor structure of the agitated behaviour scale. *Journal of Clinical and Experimental Neuropsychology, 16*(3), 386-392.

Corrigan, J.D., & Mysiw, W.J. (1988). Agitation following traumatic head injury: equivocal evidence for a discrete stage of cognitive recovery. *Archives of Physical Medicine and Rehabilitation, 69*, 487-492.

Corrigan, J.D., Mysiw, W.J., Gribble, M., & Chock, S.K. (1992). Agitation, cognition and attention during post-traumatic amnesia. *Brain Injury, 6*, 155-160.

Cox, R., & Amsters, D. (2002). Goal attainment scaling: an effective outcome measure for rural and remote health services. *Australian Journal of Rural Health, 10*, 256-261.

Cytrynbaum, S., Birdwell, G.Y., Birdwell, J., & Brandt, L. (1979). Goal attainment scaling: A critical review. *Evaluation Quarterly, 3*, 5-40.

Dijkers, M.P. (2004). Quality of life after traumatic brain injury: a review of research approaches and findings. *Archives of Physical Medicine and Rehabilitation, 85*(2), S21-S35.

Dodwell, D. (1988). The heterogeneity of social outcome following head injury. *Journal of Neurology, Neurosurgery, and Psychiatry, 51*, 833-838.

Eames, P., & Wood, R. (1985). Rehabilitation after severe brain injury: a follow-up study of a behaviour modification approach. *Journal of Neurology, Neurosurgery, and Psychiatry, 48*, 613-619.

Eames, P., Cotterill, G., Kneale, T.A., Storrar, A.L., & Yeomans, P. (1996). Outcome of intensive rehabilitation after severe brain injury. *Brain Injury, 10*, 631-650.

Evans, D.J., Oakley, S., Almdahl, S. & Davoren, B. (1999). Goal attainment scaling in a geriatric day hospital: team and program benefits. *Canadian Family Physician, 45*, 954-960.

Fisher, K., & Hardie, R.J. (2002). Goal attainment scaling in evaluating a multidisciplinary pain management program. *Clinical Rehabilitation, 16*, 871-877.

Fleminger, S. & Powell, J. Editorial. *Neuropsychological Rehabilitation, 9*(3-4), 225-230.

Folstein, M.F., Folstein, S.E., & McHugh, P.R. (1975). Mini-mental state: a practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Residents, 12*, 189-198.

Gordon, J.E., Powell, C., & Rockwood, K. (1999). Goal attainment scaling as a measure of clinically important change in nursing-home patients. *Age and Ageing, 28*, 275-281.

Greenville, J., & Lyne, P. (1995). Patient-centred evaluation and rehabilitative care. *Journal of Advanced Nursing, 22*, 965-972.

Grol, R. (2001). Improving the quality of medical care. *Journal of the American Medical Association, 284,* 2578-2585.

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.

Haas, J.F. (1997). The impact of recent changes in health care insurance on medical practice. *Journal of Head Trauma Rehabilitation, 12,* 1-9.

Hall, K., Mann, N., High, W.M., Wright, L., Kreutzer, J.S., & Wood. (1996). Functional measures after traumatic brain injury: Ceiling effects of FIM, FIM + FAM, DRS and CIQ. *Journal of Head Trauma Rehabilitation, 11*, 27-39.

Hammond, F.M., Grattan, K.D., Sasser, H., Corrigan, J.D., Bushnik, T., & Zafonte, R.D. (2001). Long-term recovery course after traumatic brain injury: a comparison of the functional independence measure and disability rating scale. *Journal of Head Trauma Rehabilitation, 16*(4), 318-329.

Hart, T., Novack, T., Englander, J., Fidler-Sheppard, R., & Bell, K.R. (2003). The relationship between neuropsychologic function and level of caregiver supervision at 1 year after traumatic brain injury. *Archives of Physical Medicine and Rehabilitation, 84*, 221-230.

Hughes, J.N., Hasbrouck, J.E., Serdahl, E., Heidgerken, A., & McHaney, L. (2001). Responsive systems consultation: a preliminary evaluation and implementation and outcomes. *Journal of Educational and Psychological Consultation, 12*(3), 179-201.

Hurwitz, T. (1999). Brief test of intellectual function (BTIF). British Columbia Neuropsychiatry Program. *Unpublished.*

Iglehart, J.K. (1999). The American health care system: Medicaid. *The New England Journal of Medicine, 340,* 403-408.

Jackson W., & Novack T. (1995) A new scale for repeated measurement of orientation in neurorehabilitation: The Orientation Log (O-Log). *National Academy of Neuropsychology*, San Francisco, CA.

Joyce, B.M., Rockwood, K.J. & Mate-Kole, C.C. (1994). Use of goal attainment scaling in brain injury in a rehabilitation hospital. *American Journal of Physical Medicine and Rehabilitation, 73*(1), 10-14.

Kiresuk, T.J., & Sherman, R.E. (1968). Goal attainment scaling: A general method for evaluating comprehensive community mental health programs. *Community Mental Health Journal, 4*, 443-453.

Kolitz, B.P., Vanderploeg, R.D., & Curtiss, G. (2003). Development of the Key Behaviors Change Inventory: a traumatic brain injury behavioural outcome assessment instrument. *Archives of Physical Medicine and Rehabilitation, 84*, 277-284.

Kraus JF. (1993). Epidemiology of head injury. In: Cooper PR, editor. Head injury. 3rd ed. Baltimore: Williams and Wilkins: 1-25.

Kuipers, P., Kendall, M., Fleming, J., & Tate, R. (2004). Comparison of the Sydney Reintegration Scale (SPRS) with the Community Integration Questionnaire (CIQ): psychometric properties. *Brain Injury, 18*(2), 161-177.

Kuttner, R. (1999). The American health care system: employer-sponsored health coverage. *The New England Journal of Medicine, 340,* 248-252.

Levin, H.S., High, W.M., Goethe, K.E., Sisson, R.A., Overall, J.E., Rhoades, H.M., Eisenberg, H.M., Kalisky, Z., & Gary, H.E. (1987). The neurobehavioral rating scale: Assessment of the behavioural sequelae of head injury by the clinician. *Journal of Neurology, Neurosurgery, and Psychiatry, 50*, 183-193.

Levin, H.S., O'Donell, V.M., & Grossman, R.G. (1979). The Galveston Orientation and Amnesia Test. A practical scale to assess cognition after head injury. *Journal of Nervous and Mental Diseases, 167*, 675-684.

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

Lezak, M.D. (1987). Relationship between personality disorders, social disturbances, and physical disability following traumatic brain injury. *Journal of Head Trauma Rehabilitation, 2*, 57-69.

MacKay, G., & Somerville, W. (1996). Reflections on goal attainment scaling (GAS): cautionary notes and proposals for development. *Educational Research, 38*(2), 161-172.

Malec, J.F. (1999). Goal attainment scaling in rehabilitation. *Neuropsychological Rehabilitation, 9*, 353-361.

Malec, J.F. (2001). Impact of comprehensive day treatment on societal participation for persons with acquired brain injury. *Archives of Physical Medicine and Rehabilitation, 82*, 885-895.

Malec, J.F., Cragness, M., Evans, R.W., Finlay, K.I., 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. (2000). Self-awareness, distress, and postacute rehabilitation outcome. *Rehabilitation Psychology, 45*(3), 227-241.

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., Smigielski, J.S., & DePompolo, R.W. (1991). Goal attainment scaling and outcome measurement in post-acute brain injury rehabilitation. *Archives of Physical Medicine and Rehabilitation, 72*, 138-143.

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*, 116-124.

Mate-Kole, C.C., Danquah, S.A., Twum, M., & Danquah, A.O. (1999). Outcomes of a nonaversive behavior intervention in intellectually impaired individuals using goal attainment scaling. *Nursing Research, 48*(4), 220-225.

May, K., & Hittner, J.B. (2003). On the relation between power and reliability of difference scores. *Perceptual and Motor Skills, 97*(3, Pt. 1), 905-908.

McColl, M.A., Davies, D., Carlson, O., Johnston, J., & Minnes, P. (2001). The Community Integration Measure: development and preliminary validation. *Archives of Physical Medicine and Rehabilitation, 82*, 429-434.

McMillan, T.M. & Sparkes, C. (1999). Goal planning and neurorehabilitation: The Wolfson neurorehabilitation centre. *Neuropsychological Rehabilitation, 9*(3-4), 241-252.

Mitchell, T. & Cusick, A. (1998). Evaluation of a client-centered pediatric rehabilitation programme using goal attainment scaling. *Australian Journal of Occupational Therapy, 45*, 7-17.

National Institutes of Health. (1998). NIH consensus panel endorses comprehensive rehabilitation strategy for brain injury.

Oren, T., & Ogletree, B.T. (2000). Program evaluation in classrooms for students with autism: student outcomes and program processes. *Focus on Autism and other Developmental Disabilities, 15*(3), 170-175.

Overall, J.E., & Gorham, D.R. (1962). The brief psychiatric rating scale. *Psychological Reports, 10*, 799-812.

Pender, N., & Fleminger, S. (1999). Outcome measures on inpatient cognitive and behavioural units: An overview. *Neuropsychological Rehabilitation, 9*, 345-361.

Ponsford, J., Olver, J., Nelms, R., Curran, C., & Ponsford, M. (1999). Outcome measurement in an inpatient and outpatient traumatic brain injury rehabilitation programme. *Neuropsychological Rehabilitation, 9*,(3/4), 517-534.

Ponsford, J.L., Olver, J.H., & Curran, C. (1995). A profile of outcome two years following traumatic brain injury. *Brain Injury, 9*, 1-10.

Raykov, T. (1999). Are simple change scores obsolete? An approach to studying correlates and predictors of change. *Applied Psychological Measurement, 23*(2), 120-126.

Rockwood, K., Graham, J.E., & Fay, S. (2002). Goal setting and attainment in Alzheimer's disease patients treated with donepezil. *Journal of Neurology, Neurosurgery and Psychiatry, 73*, 500-507.

Rockwood, K., Howlett, S., Stadnyk, K., Carver, D., Powell, C., & Stolee, P. (2003). Responsiveness of goal attainment scaling in a randomized controlled trial of comprehensive geriatric assessment. *Journal of Clinical Epidimiology, 56*, 736-743

Rockwood, K., Stolee, P., Howard, K., & Mallery, L. (1996). Use of goal attainment scaling to measure treatment effects in an anti-dementia drug trial. *Neuroepidemiology, 15*, 330-338

Rushton, P.W., & Miller, W.C. (2002). Goal attainment scaling in the rehabilitation of patients with lower-extremity amputations: a pilot study. *Archives of Physical Medicine and Rehabilitation, 83*, 771-775.

Santa Clara Valley Medical Center (2004). The Center for Outcome Measurement in Brain Injury. http://www.tbims.org/combi (October 11, 2004).

Schlosser, R.W. (2004). Goal attainment scaling as a clinical measurement technique in communication disorders: a critical review. *Journal of Communication Disorders, 37*, 217-239.

Semelyn, J.K., Summers, S.J., & Barnes, M.P. (1998). Traumatic brain injury: efficacy of multidisciplinary rehabilitation. *Archives of Physical Medicine and Rehabilitation, 79*, 678-683.

Shefler, G., Canetti, L., & Wiseman, H. (2001). Psychometric properties of goal-attainment scaling in the assessment of Mann's time-limited psychotherapy. *Journal of Clinical Psychology, 57*(7), 971-979.

Sladeczek, I.E., Elliot, S.N., Kratochwill, T.R., Robertson-Mjaanes, S., & Stiober, K.C. (2001). Application of goal attainment scaling to a conjoint behavioural consultation case. *Journal of Educational and Psychological Consultation, 12*(1), 45-58.

Smith, A. (1981). Goal attainment scaling: A method for evaluating the outcome of mental health treatment. In P. McReynolds (Ed.); *Advances in Psychological Assessment, Vol. 5*, (pp. 424-459). San Francisco: Jossey-Bass.

Sohlberg, M.M. & Mateer, C.A. (2001). *Cognitive Rehabilitation* (2nd Ed.). NY: The Guilford Press.

Stilwell, P., Stilwell, J. Hawley, C. & Davies, C. (1999). The national traumatic brain injury study: assessing outcomes across settings. *Neuropsychological Rehabilitation, 9*(3-4), 277-294.

Stirpe, T.S., Wilson, R.J. & Long, C. (2001). Goal attainment scaling with sexual offenders: A measure of clinical impact at posttreatment and at community follow-up. *Sexual Abuse: A Journal of Research and Treatment, 13*, 65-77.

Stolee, P., Zaza, C., Pedlar, A. & Myers, A.M. (1999). Clinical experience with goal attainment scaling in geriatric care. *Journal of Aging and Health, 11*(1), 96-124.

Tate, R., Pfaff, A., Veerabangsa, A., & Hodgkinson, A. (2004). Measuring psychosocial recovery after brain injury: change versus competency. *Archives of Physical Medicine and rehabilitation, 85*, 538-545.

Tate, R., Hodgkinson, A., Veerabangsa, A., & Maggiotto, S. (1999). Measuring psychosocial recovery after traumatic brain injury: properties of a new scale. *Journal of Head Trauma Rehabilitation, 14*(6), 543-557.

Turnbull, J. (1998). Evaluating health care using goal attainment scaling. *Nursing Standard, 12*(42), 35-38.

Webb, P.M., & Gluekhauf, R.L. (1994). The effects of direct involvement in goal setting on rehabilitation outcomes for persons with traumatic brain injuries. *Rehabilitation Psychology, 39*, 179-188.

World Health Organization ICIDH-2 (1997). International classification of impairments, activities, and participation. A manual for dimensions of disablement and functioning. Geneva, Switzerland: World Health Organization.

Whitneck, G.G., Charlifue, S.W., Gerhart, K.A., Overholser, D., & Richardson, G.N. (1992). Quantifying handicap: A new measure of long-term rehabilitation outcome. *Archives of Physical Medicine and Rehabilitation, 73*, 519-525.

Willer, B., & Moscato, B. (1996). National prevalence rates of traumatic brain injury among adults: the Canadian household and institutional health and activity limitation surveys, 1986-1987. *Neurorehabilitation, 6*(3), 219-228.

Willer, B., Ottenbacher, K.J., & Coad, M.L. (1994). The community integration questionnaire: a comparative examination. *American Journal of Physical Medicine and Rehabilitation, 73*, 103-111.

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.

Wright, J., Bushnik, T. & O'Hare, P. (2000). The centre for outcome measurement in brain injury (COMBI): An internet resource you should know about. *Journal of Head Trauma Rehabilitation, 15*(8), 734-738.

Yudofsky, S.C., Silver, J.M., Jackson, W., Endicott, J., & Williams, D.W. (1986). The overt aggression scale for the objective rating of verbal and physical aggression. *American Journal of Psychiatry, 143*, 35-39.

Zaza, C., Stolee, P., & Prkachin, K. (1999). The application of goal attainment scaling in chronic pain settings. *Journal of Pain and Symptom Management, 17*(1), 55-64.