Lincoln, N.B. & **dasNair, R.** (2008). Outcome measures in cognitive rehabilitation, In D.T. Stuss, G. Winocur, & I.H. Robertson (Eds.), *Cognitive Neurorehabilitation: Evidence and application (2nd Ed).* New York, NY: Cambridge University Press. Pp91-105. ISBN-10: 0521871336

Outcome measurement in cognitive neurorehabilitation

Nadina Lincoln and Roshan das Nair

Introduction

The aim of this chapter is to consider the criteria for selecting outcome measures for evaluating the effects of cognitive neurorehabilitation. The *International Classification of Function, Disability and Health* (ICF) (World Health Organisation, 2001) is used as a framework for deciding what to measure. The properties of the ideal outcome measure are discussed. Examples of outcome measures commonly used in clinical studies are provided and their strengths and limitations considered. The focus is on self-report measures rather than neuro-psychological tests as these reflect the effect of cognitive rehabilitation on daily life.

Outcome

- Outcomes may be assessed at the levels of impairment, activity or participation.
- Activity measures are the most important outcomes for cognitive rehabilitation.
- Quality of life is best assessed as component domains rather than a single measure.

Rehabilitation may be considered in terms of process, structure and outcome (Donabedian, 1966). Process consists of the activities which are designed to improve the functioning of the individual, such as the treatment techniques used by members of the multidisciplinary team to foster recovery or adaptation. Structure refers to the facilities provided to enable the treatments to be administered, such as the environment, staff and equipment. Outcome refers to the result of the rehabilitation endeavor. It is the endpoint against which the effectiveness of rehabilitation is judged. In cognitive rehabilitation, the aim is to help the patient to function to the maximum level of ability possible within the constraints of deficits resulting from brain damage. In addition, that individual should be as contented and satisfied with his or her condition as is possible, and so should the relatives. The assessment of outcome is the means by which we determine whether rehabilitation has achieved these aims.

Measurement of outcome

The ICF (World Health Organisation, 2001) is recognised as providing a useful framework for the selection of appropriate outcome measures (Heinemann, 2005; Jette & Haley, 2005; Mermis, 2005; Wade, 2003). The concepts are as follows.

(a) Body functions and deficits: these are impairments or the loss or abnormality of psychological, physiological or anatomical structure or function. They include cognitive deficits such as disorders of memory, attention and language.

(b) Activity limitation is the difficulties an individual may have in executing activities, including learning and applying knowledge; self-care; domestic life; interpersonal interactions; and community, social and civic life. It includes the effects of cognitive impairments on daily life, such as difficulties in telling the time and losing

items around the home, and the disruption to interpersonal relationships that may occur following a head injury.

(c) Participation is the involvement in life situations at a societal level. It includes the social, cultural, economic and environmental effects of activity limitation.

An impairment, such as visual inattention, usually will give rise to an activity limitation, such as the inability to dress independently, which in turn may affect participation, through loss of personal independence. An effective rehabilitation programme would reduce the impairments and the activity limitations which are consequent upon those impairments. In many instances, it is not possible to ameliorate the impairment, but nevertheless significant gains may be made by attempting directly to increase activity. Although it would be desirable to improve participation, this can rarely be achieved directly and may not entirely be within the remit of the rehabilitation service. Provided the activity limitation has been reduced, the assumption is that there will be a beneficial effect on participation.

Quality of life is an elusive outcome which relates to participation. Most assessments of quality of life incorporate several domains, which include both activities and participation. It is beyond the scope of assessment procedures to assess adequately all domains which contribute to quality of life and produce a satisfactory single quality-of-life measure. For this reason it does not seem practical even to attempt to assess general quality of life, but rather only the specific domains, even though improving the quality of life must be the ultimate goal of rehabilitation.

Selection of measures

- The psychometric properties of measures need to be considered.
- Measures of motor, sensory and cognitive impairment provide standardised descriptions of patients.
- Cognitive tests measure cognitive impairment and not the effect in daily life.

In order to choose an outcome measure it is important to know that the measure meets the requirements of a measurement tool. These requirements are:

(a) Validity: any measure must measure what it purports to measure. For example, measures of activities of daily living should include those activities people would consider to be essential for independence in daily life. Measures should relate to other measures of the same underlying ability and include all the relevant aspects of the attributes they measure. Hypotheses generated based on the measure should be upheld. For example, one would predict that head-injured people will do less well on a measure of memory than normal participants, and if this is supported it indicates that the measure has construct validity.

(b) Reliability: any outcome measure should provide the same information if used by different assessors (inter-rater reliability) or by the same assessor on different occasions (intra-rater reliability). If the assessment is to be used to monitor change it needs to have minimal practice effects and to show no variation simply as a result of repeating the assessment (test-retest reliability).

(c) Sensitivity: outcome measures in rehabilitation need to be sensitive to change, i.e., able to detect change in ability when change has occurred, and responsive to differences between rehabilitation programs.

(d) Practicality: the selection of outcome measures is dominated by practical constraints. An outcome measure must be short, easy to administer and acceptable to patients. Those measures which are tiring, detailed, intrusive or repetitive will not be tolerated. It must also be easy to communicate the findings to others.

Many studies evaluating the effects of rehabilitation have used single-case experimental designs. Measures are repeated frequently and therefore need to be very short, in order not to induce fatigue and not to interfere with the treatment program. They also need to have minimal practice effects so

that stable baselines can be achieved. Alternative versions of tests may be used but few standardised tests have sufficient alternative versions available to be suitable for monitoring progress in single-case experimental designs. When using a randomised controlled trial to evaluate an intervention, it is necessary to use measures which have been used in other studies, so that comparisons between trials are possible. It is also important to include sufficient patients to be sure that a small difference in outcome has not been missed, yet from a patient's perspective even a small gain in function may be worthwhile. A common strategy to resolve this is to conduct a meta-analysis of several trials, which is facilitated by the use of a common outcome measure.

Measures of impairment

Cognitive rehabilitation is designed to improve cognitive abilities. Most cognitive impairments can be assessed by a range of measures but few have been designed as measures of outcome. Some cognitive assessments are intended as screening devices, to identify cognitive impairments which require further evaluation. Others are diagnostic tools, to detect cognitive impairment and differentiate particular impairments from each other. Assessments for screening or diagnostic purposes may not be suitable measures of outcome. Their validity will be based on their ability to classify and therefore those which give clear cut-off points will be most robust. In contrast outcome measures need to have continuous scales such that they are sensitive to changes in ability rather than whether a patient fits a particular category. Measures of specific cognitive impairments can be found in later chapters and therefore are not reviewed in detail here. However, impairments other than cognitive deficits may affect a patient's response to a cognitive rehabilitation program, and it is therefore appropriate to assess these. They will not be used as outcome assessments, as few cognitive rehabilitation techniques would be expected to decrease, for example, motor impairments. However, the treatment of visual inattention could be predicted to reduce sensory inattention or visual field deficits. For this reason, such assessments are considered.

Motor function may be assessed using summed indices, such as the Rivermead Motor Assessment (Lincoln & Leadbitter, 1979) and the Motor Assessment Scale (Carr *et al*., 1985), to indicate the level of motor impairment. The Rivermead Mobility Index (Collen *et al*., 1991) is a comprehensive assessment which includes sitting, transfers, walking and getting up stairs. It has been used as an outcome measure in rehabilitation studies on stroke (Wade *et al*., 1992) and is sensitive to change in people with multiple sclerosis (Vaney *et al*., 1996). The Expanded Disability Status Scale (Kurtzke, 1983) and Guys Neurological Disability Scale (Sharrack & Hughes, 1999), despite their names, are predominantly measures of impairments in multiple sclerosis and include aspects of motor function.

Sensory impairment is assessed as part of a clinical examination but there are only a few standardised scales available (Lincoln *et al*., 1998, Winward *et al*., 2002). Tactile inattention, proprioception and stereognosis may affect the outcome of cognitive rehabilitation, so it is important that they are assessed. Visual impairment will affect patients' ability to participate in cognitive rehabilitation but conventional acuity measurement techniques require language skills, and visual field assessment may be confounded by the presence of inattention (Walker *et al*., 1991).

Cognitive impairments are assessed by a wide range of psychological tests. Assessments are available to determine the severity of deficits in language, perception, memory, reasoning, attention, movement disorders and other cognitive functions. Standardised cognitive tests may assess "pure" levels of impairment specific to one cognitive domain. For example, memory can be assessed in terms of encoding ability or working memory capacity. However, such tests may not be of much value as outcome measures. In order for a cognitive assessment to be used as an outcome measure, it must be

sensitive to changes over time but have minimal practice effects. For example, although recognition memory tests are sensitive to differences between individuals, they are not appropriate for evaluating change unless there are parallel versions available. Many cognitive tests, while reliable over time, will show sufficient improvement simply as a result of practice to make them insensitive to small differences between interventions.

Activities

The main aim of cognitive rehabilitation is an increase in activity, which includes functional, cognitive, emotional and social activities.

- The most important outcomes for cognitive rehabilitation are cognitive activities, yet the measurement of these is poorly developed.
- Independence in activities of daily living and mood measures provide proxy measures for the outcome of cognitive rehabilitation.
- The Extended Activities of Daily Living (EADL) and Functional Independence Measure (FIM) are the most suitable measures of independence in activities of daily living.
- Mood should be assessed on questionnaires designed to detect change.

Limitations in functional activities

The ability to perform functional activities in everyday life may be assessed by scales of activities of daily living (ADL), including basic personal self-care skills and instrumental activities of daily living. The choice of scale is governed by practical considerations and there is little consensus on the 'best' measures (Jette & Haley, 2005). Most ADL scales consist of summed indices, but concerns have been expressed about treating such scales as ordinal, though some have been demonstrated to be acceptable using Guttman scaling, e.g., EADL and Barthel, or a Rasch model, e.g., Barthel and FIM.

The Barthel Index (Mahoney & Barthel, 1965) is a widely used measure of personal activities of daily living which has almost become the gold standard for stroke rehabilitation studies. It is sensitive to differences between rehabilitation interventions (Indredavik *et al*., 1991) and reliable when administered verbally or by post. It is sensitive in the inpatient phase of rehabilitation (Houlden *et al*., 2006) but less sensitive to change in community-based patients. The original index was scored on a 0–100 scale, but this implies a spurious degree of accuracy, and the revised 20-point version (Collin *et al*., 1988) has become the standard. The main limitation is its ceiling effect; therefore, rehabilitation studies may include a measure of instrumental activities of daily living in addition to the Barthel Index.

Two widely used measures of instrumental activities of daily living are the Nottingham Extended Activities of Daily Living (EADL) scale (Nouri & Lincoln, 1987) and the Frenchay Activities Index (Holbrook & Skilbeck, 1983). Each includes domestic activities such as preparing meals and washing up, mobility outside the home, leisure and social activities. Both scales have been found to be sensitive to differences between rehabilitation interventions (Forster & Young, 1996; Fuller *et al*., 1996). The EADL is suitable for multi-center studies as it has been validated for postal administration. Cognitive rehabilitation programs are more likely to have an effect on instrumental activities of daily living than on personal self-care skills. Therefore, these scales may be appropriate for assessing the generalisation of cognitive retraining to daily life skills.

The main measure of personal ADL used with a wide range of patients is the Functional Independence Measure (FIM) (Granger *et al*., 1986). This covers personal self-care and motor activities and has cognitive items on comprehension, expression, social interaction, problem solving and memory. Inter-rater reliability has been found to be higher for physical disability than communication and social cognition sections (Brosseau & Wolfson, 1994; Kidd *et al*., 1995; Pollak *et al*., 1996). Because the effect of a cognitive rehabilitation program is most relevant to these items, further improvement to the reliability is needed. Rasch analysis has indicated that the motor and cognitive items form two

distinct scales, though this has been questioned by Dickson & Kohler (1996), who identified six factors from a factor analysis of the FIM. Differential item functioning has also been identified in different patient groups (Dallmeijer *et al*., 2005), particularly in the motor domain and disordered thresholds, leading to the suggestion that the number of response categories should be reduced (Dallmeijer *et al*., 2005; Nilsson *et al*., 2005). The cognitive scale has shown ceiling effects in people with multiple sclerosis (van der Putten *et al*., 1999) and administration requires prolonged observation of the patient, which means the scale would be difficult to administer in the context of a randomised controlled trial by an independent assessor. Houlden *et al*. (2006) demonstrated comparable responsiveness between the Barthel and the FIM.

The Functional Assessment Measure (FAM), which was developed to assess the specific problems of brain-injured patients (Hall *et al*., 1993), contains additional items emphasising cognitive, communicative and psychosocial function. McPherson *et al*. (1996) found high inter-rater reliability, but greatest discrepancies occurred on cognitive, communication and behavioral items. Hobart *et al*. (2001b) compared the psychometric properties of the Barthel, FIM and FIM+FAM and found them similar, and highlighted that the FIM and FIM+FAM have significant redundancy of items and confer few advantages over the shorter simpler Barthel for patients receiving inpatient rehabilitation. For the FIM+FAM to be used to evaluate the outcome of cognitive rehabilitation, further work is needed to check the reliability and sensitivity.

The Rivermead Head Injury Follow-Up Questionnaire (Crawford *et al*., 1996) was developed as an outcome measure for patients with mild to moderate head injury. It is short, simple and can be administered by post or at interview. It was found to have good inter-rater reliability, to be sensitive to changes over time and to detect differences in outcome in a randomised controlled trial (Wade *et al*., 1997). The Brain Injury Community Rehabilitation Outcome 39 (BICRO39) (Powell *et al*., 1998) covers aspects of personal and social functioning for brain injury patients living in the community. It has good test-retest and inter-rater reliability and has also been found to be sensitive to the effects of intervention (Powell *et al*., 2002).

Measures of activity for people with multiple sclerosis seem to be few. The FIM (Brosseau & Wolfson, 1994) and Assessment of Motor and Process Skills (AMPS) (Doble *et al*., 1994) and Functional Assessment of MS (Cella *et al*., 1996) have been used, but there are few data to indicate the most appropriate measure for this group. The AMPS requires patients to perform tasks and so takes longer, but it covers instrumental activities of daily living, which may be more important to assess than personal activities of daily living, when the effects of cognitive rehabilitation are being evaluated and has adequate reliability (Doble *et al*., 1999).

In addition there are general rehabilitation outcome measures, which are predominantly measures of activity limitation. The Sickness Impact Profile (Bergner *et al*., 1981), the British version of which is the Functional Limitations Profile (Charlton *et al*., 1983), assesses the impact of sickness on daily activities and behavior. It provides subscales in 12 areas, including ambulation, body care and household management. It is lengthy and complex to administer.

Mood

Although mood disorders might be considered either as an impairment, i.e., a direct consequence of some underlying pathology, or as a consequence of impairments, in the context of cognitive rehabilitation they are probably best considered as emotional disabilities. Many mood scales are available but few have been validated for patients with neurological disorders. Those which have items affected by physical disability are likely to be insensitive to mood changes. Many mood questionnaires were developed as screening devices to detect significant levels of depression or anxiety. To evaluate the outcome of cognitive neurorehabilitation, measures need to be sensitive to change and therefore not all

screening questionnaires will be suitable for this purpose.

The Hospital Anxiety and Depression Scale (Zigmund & Snaith, 1983) is a short, easy to complete measure which provides separate scores for both anxiety and depression. Although it was designed for hospitalised medically ill patients, several items reflect physical disability. It is probably more suitable as a screening measure than for assessing outcome. In contrast, other scales, such as the Beck Depression Inventory II (Beck *et al*., 1996) and Center for Epidemiologic Studies Depression Scale (Radloff, 1977), were designed as measures of the severity of depressive symptoms. They are likely to be more sensitive to the effects of intervention than measures designed for screening purposes (Turner-Stokes & Hassan, 2002).

Cognitive rehabilitation would be expected to improve mood. If patients' cognitive performance improves following treatment, they will probably suffer from less depression, be less anxious and suffer less general distress. The most widely used measure of general psychological distress is the General Health Questionnaire (Goldberg & Williams, 1988). The GHQ28 and GHQ30 have been found to be sensitive to the effects of psychological interventions in randomised controlled trials (Juby *et al*., 1996, Watkins *et al*., 2007); the GHQ12 is shorter but its sensitivity to the effects of intervention in neurological patients has not been established.

One problem in assessing mood in neurological patients is that many have communication difficulties. The Stroke Aphasic Depression Questionnaire (Lincoln *et al*., 2000; Sutcliffe & Lincoln, 1998) was developed to assess mood in aphasic patients. Items that could be observed by relatives or nursing staff were taken from mood questionnaires and rephrased in terms of observable behaviors. The Visual Analog Mood Scales (Stern, 1997) provide a pictorial method of assessing mood. The Neuropsychiatric Inventory (Cummings *et al*., 1994) was developed to assess mood disorders in people with dementia and has been used as an outcome measure for pharmacological treatments in dementia (Ringman & Cummings, Chapter 19, this volume). These scales have good validity and reliability, but their sensitivity to change in response to cognitive rehabilitation has not been established.

Limitations in cognitive activities

Limitations in cognitive activities are a major concern for patients and carers and are primary targets for cognitive rehabilitation programs. (For further discussion see Cicerone, Chapter 7, this volume.) Outcome is assessed in several ways, including semi-structured interviews, questionnaires or patient observation. The most common strategy has been to employ a questionnaire that includes items on the behavioral manifestations of cognitive impairments. These may be completed by patients or carers, to determine the subjective effects of cognitive impairment on daily life. These are important indicators of the outcome of cognitive neurorehabilitation, but there are few scales available and not all possible cognitive deficits are included. Another problem is the reliability of eliciting this kind of information from individuals who may not be able to judge their functioning accurately due to their cognitive deficits. For example, frontal lobe damage after brain injury results in impaired metacognitive processing (Hanten *et al*., 2000), patients with MS underestimate their memory problems on questionnaires (Beatty & Monson, 1991) and patients with memory problems and epilepsy, tended to overstate memory problems on questionnaires, compared with "objective" tests of memory (Piazzini *et al*., 2001). Semi-structured interviews may provide valuable information regarding the patients' experience of cognitive rehabilitation, but may not be accurate measures of outcome.

Ecological validity

Cognitive assessments can use ecologically valid tests that assess cognitive functions in the context of everyday tasks. For example, the Rivermead Behavioral Memory Test (RBMT) (Wilson *et al*., 1985) has tasks, such as remembering where an

object was placed in a room or remembering to do things at appointed times. This task-based performance is related to a number of cognitive functions, and not a specific one, but provides a clinical assessment which approximates the patient's functioning in everyday life. Two approaches to ecological validity (Franzen & Wilhelm, 1996) have been adopted. One is developing tests with high face validity which simulate daily tasks (requiring the underlying cognitive functions to complete these tasks to be intact). The other relates performance on pre-existing (traditional) tests to daily functioning (Chaytor & Schmitter-Edgecombe, 2003). Assessments of cognitive function which are designed to reflect the cognitive skills needed in everyday life, such as the RBMT (Wilson *et al*., 1985), Behavioral Inattention Test (BIT) (Wilson *et al*., 1987), Behavioral Assessment of the Dysexecutive Syndrome (Wilson *et al*., 1996) and Test of Everyday Attention (TEA) (Robertson *et al*., 1994), maybe considered to measure limitations to cognitive activities rather than impairment. They include items which are likely to be predictive of everyday performance rather than assessing everyday performance itself. Although they have the advantage of ecological validity, unlike many assessments of cognitive impairment, they are probably not true measures of activity limitation, because they comprise artificial activities and not those which people necessarily perform on a daily basis. Ecologically valid tests of memory and attention (e.g., RBMT and TEA) have been found to be better at predicting functional disability than memory questionnaires (Higginson *et al*., 2000). However, there is a trade-off when developing ecologically valid tests: the more they approximate real-world scenarios, the less structured they are; and consequently have poor psychometric properties (Van Zomeren & Spikman, 2003). Assessments in the main cognitive domains will be considered.

Memory

Subjective memory impairment has been investigated using the Everyday Memory Questionnaire (EMQ) (Sunderland *et al*., 1983) in studies of stroke (Tinson & Lincoln, 1987), head injury (Sunderland *et al*., 1984) and MS (Taylor, 1990). The questionnaire has been used both for patients to assess their own problems and for relatives. Five factors, reflecting the underlying memory processes, have been identified in healthy individuals: retrieval, task monitoring, conversational monitoring, spatial memory and memory for activities (Cornish, 2000). The EMQ has validity in that it correlates moderately with tests of memory, and its reliability is acceptable, though not good. The Subjective Memory Assessment Questionnaire (Davis *et al*., 1995) is short and has been validated for stroke patients. The Memory Failures Questionnaire (Gilewski *et al*., 1990) contains more items in four subscales, including one on the use of mnemonics. There is conflicting evidence on the extent to which it correlates with prospective memory (Kinsella *et al*., 1996; Zelinski *et al*., 1990).

In general, memory questionnaires appear to have adequate reliability, but low validity, particularly when completed by patients (Ruisel, 1991). To complement memory questionnaires, and to compensate for some of their limitations, memory "performance" tests are frequently used. Such tests, which approximate real-life scenarios, may be more valid measures of rehabilitation outcome than traditional memory tests. The Rivermead Behavioral Memory Test (RBMT) (Wilson *et al*., 1985) was developed to assess everyday memory problems in individuals with acquired brain damage. The extended version (RBMT-E) (Wilson *et al*., 1998), with two parallel forms, is sensitive to milder memory deficits. It has age norms (years 11–95), and has been translated into many languages.

Visual neglect

The behavioral manifestations of visual neglect have also been assessed by questionnaires. Towle & Lincoln (1991) developed the Problems in Everyday Living Questionnaire as a subjective measure of visual neglect. The patient has to report how often problems, such as bumping into door frames and making errors when dialling the

telephone, have occurred. The Catherine Bergego Scale (Azouvi *et al*., 1996) contains ten items which the patient has to rate according to their severity. It has been found to have good inter-rater reliability and validity. Test-retest reliability and sensitivity to change need to be checked. Neither scale has been demonstrated to be sensitive to differences between interventions. An alternative approach has been to ask patients to carry out practical tasks and to observe their performance. Zoccolotti *et al.* (1992) described a scale which differentiated tasks involving the exploration of external space, dealing cards and serving tea, from those which related to one's own body, using a comb or razor. The scales were found to have high inter-rater reliability and internal consistency, and concurrent validity in relation to conventional impairment measures.

An ecologically valid test of unilateral visual neglect that reflects patients' daily life is the Behavioral Inattention Test (BIT) (Wilson *et al*., 1987). This short, easy to administer test has six conventional tests (such as line cancellation) and nine behavioral tests (such as telephone dialling). Hartman-Maeir & Katz (1995) found construct and predictive validity of most of the behavioral subtests as functional measures of neglect. To assess the effects of domain-specific interventions, tests of neglect for body space, peripersonal space and loco-motor space may be used. Robertson *et al*. (1998) used a variant of the Hair Combing Task (Zoccolotti & Judica, 1992) to measure neglect of body space, the Baking Tray Task (Tham & Tegnér, 1996) to assess peripersonal space and a tailor-made navigation task to assess locomotor neglect. The Shapes Task (Maddicks *et al*., 2003), in which the patient has to name 20 shapes on a wall three meters away, has also been used for this purpose. Although some neglect tests have been designed for diagnosis, imaginative adaptations can render them suitable outcome measures.

Attention

van Zomeren and Spikman (2003) discussed impairments of attention in terms of hemi-neglect, mental slowness, lack of control (focused and divided attention) and poor sustained attention. Some traditional tests, e.g., the Stroop Color Word and Trail Making, have been used to assess these, but as tests of impairment, they can only be used as proxy measures of rehabilitation outcome. An alternative has been to use questionnaires.

Changes associated with attention training have been assessed with the Attention Questionnaire (Sohlberg *et al*., 2000), the Dysexecutive Questionnaire (Wilson *et al*., 1996) and the Attention Rating and Monitoring Scale (ARMS) (Cicerone, 2002). The latter consists of 15 items measuring concentration, mental effort, and cognitive symptoms associated with attentional difficulties. An observational rating scale, the Moss Attention Rating Scale (MARS) (Whyte *et al*., 2003) is completed by occupational and physical therapists. The inter-rater reliability is good but further evaluation of reliability and validity is needed.

The Rating Scale of Attentional Behavior (Ponsford & Kinsella, 1991) showed moderate correlations with neuropsychological measures of attention, good internal consistency and intra-rater reliability but agreement between raters working in different contexts was less satisfactory. Scores showed change over time with treatment but the correspondence to neuropsychological measures of attention was low. Discrepancies seemed to occur as a result of emotional factors and expectations of the therapists. This highlights the difficulty of validating such scales, and Ponsford & Kinsella (1991) suggested that more concrete descriptions of scale items might reduce this subjectivity.

The Test of Everyday Attention (TEA) (Robertson *et al*., 1994) is an ecologically valid test of selective attention, sustained attention and attentional switching. It consists of eight subtests that approximate attentional tasks carried out by people in daily life, such as listening to winning numbers in a lottery, searching telephone directories and scanning maps. It takes about 60 minutes to complete and normative data exists for ages 18–80. It has three parallel forms and high test-retest reliability. Discriminative validity has been established in brain injury (Chan, 2000).

Executive functions

People with impairment of executive function display problems in initiating and stopping behaviors, shifting set, paying attention and being aware of themselves and others. Cognitive rehabilitation of executive dysfunctions has aimed to reduce the barriers to participation and activity limitations. Worthington (2005) recommended that outcomes be measurable as "socially meaningful (as opposed to statistically significant) change" (p. 259). A rating scale for problem-solving behaviors was developed by Von Cramon *et al*. (1991) to evaluate the behavioral effects of treatment. Aspects of problem-solving behavior were rated according to the-frequency of their occurrence. The scale was found to be reliable and sensitive to improvements.

The Behavioral Assessment of the Dysexecutive Syndrome (BADS) (Wilson *et al*., 1996) assesses problems related to planning, organising, initiating, monitoring and adapting behavior. It comprises six tests which simulate real-life scenarios. The Dysexecutive Questionnaire (DEX) has 20 items on a Likert scale that describe behaviors related to the dysexecutive syndrome. Patient and carer versions exist. Reliability has been evaluated by the authors, and validity by other small studies (Norris & Tate, 2000). Significant correlations have been reported between executive tests, such as Wisconsin Card Sorting (0.37) and phonemic fluency (0.35), and ratings on the DEX (Burgess *et al*., 1998), and with the Disability Rating Scale (0.52) (Hanks *et al*., 1999).

Language

The effects of language problems on everyday life have been investigated in detail. The Communicative Activities for Daily Living (Holland *et al*., 1999) presents language tasks through role play. This is more sensitive to communication strengths than traditional testing but is not a naturalistic observation. The Functional Communication Profile (Sarno, 1969) and Edinburgh Functional Communication Profile (Skinner *et al*., 1984) provide ratings of everyday language behavior but are very subjective. The Profile of Functional Impairment in Communication (Linscott *et al*., 1996) contains a detailed analysis of communication skills but requires an experienced assessor. Aphasia batteries have been employed as measures of change over time, but these are not sensitive (Weniger, 1990), with some subtests being sensitive but having few stimuli, and overall scores being too generic (Nickels, 2005).

Assessment of general cognitive functions

In addition there are scales which do not attempt to assess specific cognitive domains but consider cognitive ability in general.

The Level of Cognitive Functioning Scale (LCFS) (Hagen *et al*., 1972) (also referred to as Rancho Los Amigo Scale) has been used to assess cognitive functioning in post-coma patients. This scale has an 8-level classification of functioning, ranging from "no response" to "purposeful-appropriate." It has good test-retest and inter-rater reliability and concurrent and predictive validity (Gouvier *et al*., 1987). However, the categories are broad, making detection of small changes difficult.

The Cognitive Failures Questionnaire (CFQ) (Broadbent *et al*., 1982) is a 5-point self-rating scale that determines the frequency with which cognitive slips (arising from failures in perception, memory, and motor functions) have occurred, with versions for both patients and significant others. It is less specific to memory and includes the behavioral consequences of other cognitive deficits. It has been used as a measure of outcome of treatment, including light therapy for neuropsycho-logical functions in seasonal affective disorder (Michalon *et al*., 1997) and cognitive assessment in stroke rehabilitation (McKinney *et al*., 2002). A similar scale, the Cognitive Log (Cog-Log) (Alderson & Novack, 2003), is suitable for assessing recovery of higher neuropsychological processes with inpatients. The Log assesses verbal recall, attention, working memory, motor sequencing and response inhibition, and is reported to have good inter-rater

reliability and high internal consistency (Alderson & Novack, 2003).

Limitation in social and occupational activities

Behavioral and psychosocial problems are common consequences of traumatic brain injury and need to be assessed, particularly in the later stages of rehabilitation (see Dawson & Winocur, Chapter 14, this volume). Assessment procedures have been criticised for their lack of rigorous evaluation (Hall, 1992). The Neurobehavioral Rating Scale (Levin *et al*., 1987) is based on behavior, symptoms and skills measured in a structured clinical setting. The Neurobehavioral Functioning Inventory can be used with informants to record their perceptions of everyday problems (Kreutzer *et al*., 1996). Neither scale has well-established reliability or validity nor shown to be a sensitive indicator of rehabilitation outcome. However, the scales would seem to be worth developing further as they tap the activity limitation associated with cognitive impairment.

Participation

- There are few measures of participation.
- The best developed measures are the London Handicap Scale and Short Form 36.
- Global measures which include impairment, activity and participation are unlikely to be sensitive to the effects of cognitive rehabilitation.

Although it is unlikely that cognitive rehabilitation will have a direct effect on participation, because multiple factors will determine participation and cognitive problems are only one of many, it may be useful to assess participation as an effect of an overall rehabilitation package. Several measures have recently been developed (Heinemann, 2005) and some overlap with indicators of quality of life. Participation is subjective and inherently more difficult to measure than activity limitation, particularly in people with cognitive impairments.

The London Handicap Scale (Harwood *et al*., 1994) generates a profile of handicaps in the six survival roles: orientation, physical independence, mobility, occupation of time, social integration and economic self-sufficiency and an overall severity score. Other potential measures include the Community Integration measure (McColl *et al*., 2001), Craig Handicap Assessment and Reporting Technique (Whiteneck *et al*., 1992) and MS Impact Scale (Hobart *et al*., 2001a).

Generic measures of quality of life may also be used as indicators of participation. The Short Form 36 (Stewart & Ware, 1992) is a short questionnaire with good construct validity (Ware *et al*., 1993). However, there are doubts about its applicability with the elderly (Hill & Harries, 1994) and it has floor effects in physically disabled groups (Freeman *et al*., 1996). Wade (2003) questioned whether measurement of quality of life was appropriate in the context of rehabilitation.

Future developments

- The psychometric properties of most measures need further evaluation.
- Activity measures are needed for the evaluation of outcome in single case experimental design studies.
- Researchers conducting randomised control trials should attempt to reach consensus on a few standard activity measures in order to facilitate meta-analyses.

For each of the measures mentioned, further work is needed to establish the validity and reliability of the scale. In particular, the validity is often only established with one diagnostic group. Validation studies should be carried out in several groups of patients to confirm the underlying construct. The reliability needs to be checked in a variety of situations (inpatient, outpatient, hospital, community), conditions and over a variety of time intervals. For most scales this task has hardly yet begun. Sensitivity to the effects of intervention will not be established until there are far more efficacy studies. Most single case

experimental design studies use measures of impairment to assess the effect of intervention. However, it is the effect on daily life that is of most concern to patients and their families. Therefore, activity measures need to be developed for use in this context. Few of the measures described above are sensitive to the small changes in ability that need to be detected and many are too long to be administered with the frequency that is necessary in single case experimental design studies. The alternative approach to treatment evaluation is the randomised controlled trial. Several measures described above have been found to be sensitive to differences between rehabilitation procedures in randomised controlled trials. However, there have been few well-designed, methodologically sound trials of cognitive rehabilitation (see Cicerone, Chapter 7, this volume; Lincoln & Bowen 2006). It is hard to recruit sufficient patients within a single center; therefore multi-center trials of cognitive rehabilitation are needed. This requires consensus about which outcome measure to use. In addition, consistency of outcome measures is important for meta-analysis. The main way forward, therefore, seems to be to agree on a group of outcome measures suitable for trials of cognitive neurorehabi-litation. These measures need to be refined in terms of their psychometric properties. If this is achieved in the context of research, it will also then be possible to use the measures for the audit of clinical services and the evaluation of the progress of an individual patient in a rehabilitation setting.

Conclusions

Standardised measures are available for evaluating the outcome of cognitive neurorehabilitation. These include measures at the levels of impairment, activity and participation. At each level, the measures chosen should be reliable, valid, sensitive to the effects of intervention and consistent with those used by other researchers. There is a particular need for the development of measures of limitations in cognitive activities.

REFERENCES

Alderson, A. L., & Novack, T. A. (2003). Reliable serial measurement of cognitive processes in rehabilitation: the cognitive log. *Archives of Physical Medicine and Rehabilitation*, *84*, 668–672.

Azouvi, P., Marchal, F., Samuel, C. *et al*. (1996). Functional consequences and awareness of unilateral neglect: study of an evaluation scale. *Neuropsychological Rehabilitation*, 6, 133–150.

Beatty, W. W., & Monson, N. (1991). Metamemory in multiple sclerosis. *Journal of Clinical and Experimental Neuropsychology*, *13*, 309–327.

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

Bergner, M., Bobbitt, R. A., Carter, W. B., & Gilson, B. S. (1981). The Sickness Impact Profile: development and final revision of a health status measure. *Medical Care*, ***19***, 787–805. Broadbent, D. E., Cooper, P. F., Fitzgerald, P., & Parkes, K. R. (1982). The Cognitive Failures Questionnaire (CFQ) and its correlates. *British Journal of Clinical Psychology*, ***21***, 1–16.

Brosseau, L., & Wolfson, C. (1994). The inter-rater relability and construct validity of the Functional Independence Measure for multiple sclerosis subjects. *Clinical Rehabilitation*, *8*, 107–115.

Burgess, P. W., Alderman, N., Evans, J., Emslie, H., & Wilson, B. A. (1998). The ecological validity of tests of executive function. *Journal of the International Neuropsychological Society*, *4*, 547–558.

Carr, J. H., Shepherd, R. B., Nordholm, L., & Lynne, D. (1985). Investigation of a new motor assessment scale for stroke patients. *Physical Therapy*, **65**, 175–180.

Cella, D. F., Dineen, K., Arnason, B. *et al*. (1996). Validation of the Functional Assessment of Multiple Sclerosis quality of life instrument. *Neurology*, *47*, 129–139.

Chan, R. C. K. (2000). Attentional deficits in patients with closed head injury: a further study to the discriminative validity of the test of everyday attention. *Brain Injury*, *14*, 227–236.

Charlton, J. R. H., Patrick, D. L., & Peach, H. (1983). Use of multi-variate measures of disability in health surveys. *Journal of Epidemiology and Community Health*, *37*, 296–304.

Chaytor, N., & Schmitter-Edgecombe, M. (2003). The ecological validity of neuropsychological tests: a review of the literature on everyday cognitive skills. *Neuropsychological Review*, ***13***, 181–197.

Cicerone, K. D. (2002). Remediation of 'working attention' in mild traumatic brain injury. *Brain Injury*, **16**, 185–195.

Collen, F. M., Wade, D. T., Robb, G. F., & Bradshaw, C. M. (1991). The Rivermead Mobility Index: a further development of the Rivermead Motor Assessment. *International Disability Studies*, 13, 50–54.

Collin, C., Wade, D. T., Davis, S., & Home, V. (1988). The Barthel Index: a reliability study. *International Disability Studies*, **10**, 61–63.

Cornish, I. M. (2000). Factor structure of the Everyday Memory Questionnaire. *British Journal of Psychology*, **91**, 427–438.

Crawford, S., Wenden, F. J., & Wade, D. T. (1996). The Rivermead Head Injury Follow-up Questionnaire: a study of a new rating scale and other measures to evaluate outcome after head injury. *Journal of Neurology, Neurosurgery, and Psychiatry*, **60**, 510–514.

Cummings, J. L., Mega, M., Gray, K. *et al.* (1994). The Neuropsychiatric Inventory: comprehensive assessment of psychopathology in dementia. *Neurology*, 44,2308–2314.

Dallmeijer, A., Dekker, J., Roorda, L. *et al*. (2005). Differential item functioning of the functional independence measure in higher performing neurological patients. *Journal of Rehabilitation Medicine*, 37, 346–352.

Davis, A. M., Cockburn, J. M., Wade, D. T., & Smith, P. T. (1995). A subjective memory assessment questionnaire for use with elderly people after stroke. *Clinical Rehabilitation*, 9, 238–244.

Dickson, H. G., & Kohler, F. (1996). The multi-dimensionality of the FIM motor items precludes an interval scaling using Rasch analysis. *Scandinavian Journal of Rehabilitation Medicine*, **26**, 159–162.

Doble, S. E., Fisk, J. D., Fisher, A. G., Ritvo, P. O., & Murray, T. J. (1994). Functional competence of community dwelling persons with multiple sclerosis using the assessment of motor and process skills. *Archives of Physical Medicine and Rehabilitation*, 75, 843–851.

Doble, S. E., Fisk, J. D., Lewis, M., & Rockwood, K. (1999). Test-retest reliability of the assessment of motor and process skills in elderly adults. *Occupational Therapy Journal of Research*, **19**, 203–215.

Donabedian, A. (1966). Evaluating the quality of medical care. *Evaluating the Quality ofMedical Care*, 44, 166–206.

Forster, A., & Young, J. (1996). Specialist nurse support for patients with stroke in the community: a randomised controlled trial. *British Medical Journal*, **312**, 1642–1646.

Franzen, M. D., & Wilhelm, K. L. (1996). Conceptual foundations of ecological validity in neuropsychology. In R. J. Sbordone & C. J. Long (Eds.), *Ecological Validity of Neuropsychological Testing* (pp. 91–112). Delray Beach, FL: GR Press/St Lucie Press.

Freeman, J. A., Langdon, D. W., & Thompson, A. J. (1996). Health related quality of life in people with multiple sclerosis undergoing in-patient rehabilitation. *Journal ofNeurologic Rehabilitation*, **10**, 185–194.

Fuller, K. J., Dawson, K., & Wiles, C. M. (1996). Physiotherapy in chronic multiple sclerosis: a controlled trial. *Clinical Rehabilitation*, **10**, 195–204.

Gilewski, M. J., Zelinski, E. M., & Schaie, K. W. (1990). The Memory Functioning Questionnaire for assessment of memory complaints in adulthood and old age. *Psychology and Aging*, 5, 482–490.

Goldberg, D., & Williams, P. (1988). *A User's Guide to the General Health Questionnaire*. Windsor, UK: NFER-Nelson.

Gouvier, W. D., Blanton, P. D., LaPorte, K. K., & Nepomuceno, C. (1987). Reliability and validity of the Disability Rating Scale and the Levels of Cognitive Functioning Scale in monitoring recovery from severe head injury. *Archives of Physical Medicine and Rehabilitation*, **68**, 94–97.

Granger, C. V., Hamilton, B. B., & Sherwin, F. S. (1986). *Guide for Use of the Uniform Data Set for Medical Rehabilitation*. Buffalo, NY: Department of Rehabilitation Medicine, Buffalo General Hospital.

Hagen, C., Malkmus, D., & Durham, P. (1972). *Levels of Cognitive Functioning*. Downey, CA: Rancho Los Amigos Hospital.

Hall, K. M. (1992). Overview of functional assessment scales in brain injury rehabilitation. *Neurorehabilitation*, 2, 98–113.

Hall, K. M., Hamilton, B. B., Gordon, W. A., & Zasler, N. D. (1993). Characteristics and comparisons of functional assessment indices: Disability Rating Scale, Functional Independence Measure and Functional Assessment Measure. *Head Trauma Rehabilitation*, 8, 60–71.

Hanks, R. A., Rapport, L. J., Millis, S. R., & Deshpande, S. A. (1999). Measures of executive functioning as predictors of functional ability and social integration in a rehabilitation sample. *Archives of Physical Medicine and Rehabilitation*, **80**, 1030–1037.

Hanten, G., Bartha, M., & Levin, H. S. (2000). Metamemory following pediatric traumatic brain injury: a preliminary study. *Developmental Neuropsychology*, **18**, 383–398. Hartman-Maeir, A., & Katz, N. (1995). Validity of the Behavioural Inattention Test (BIT): relationships with functional tasks. *American Journal of Occupational Therapy*, 49, 507–516.

Harwood, R. H., Rogers, A., Dickinson, E., & Ebrahim, S. (1994). Measuring handicap: London Handicap Scale, a new outcome measure for chronic disease. *Quality Health Care*, 3, 11–16.

Heinemann, A. W. (2005). Putting outcome measurement in context: a rehabilitation psychology perspective. *Rehabilitation Psychology*, **50**, 6–14.

Higginson, C. I., Arnett, P. A., & Voss, W. D. (2000). The ecological validity of clinical tests of memory and attention in multiple sclerosis. *Archives of Clinical Neuropsychology*, 15, 185–204.

Hill, S., & Harries, U. (1994). *Assessing the Outcome ofHealth Care for the Older Person in Community Settings: Should we use the SF-36?* Leeds, UK: Clearing House on Health Outcomes, Nuffield Institute for Health.

Hobart, J., Lamping, D., Fitzpatrick, R., Riazi, A., & Thompson, A. (2001a). The Multiple Sclerosis Impact Scale (MSIS-29): a new patient-based outcome measure. *Brain*, 124, 962–973.

Hobart, J. C., Lamping, D. L., Freeman, J. A. *et al.* (2001b). Evidence-based measurement: which disability scale for neurologic rehabilitation? *Neurology*, 57, 639–644.

Holbrook, M., & Skilbeck, C. E. (1983). An activities index for use with stroke patients. *Age and Ageing*, **12**, 166–170.

Holland, A. L., Frattali, C. M., & Fromm, D. (1999). *Communication Activities for Daily Living (2nd Edition).* Austin, TX: Pro-E.

Houlden, H., Edwards, M., McNeil, J., & Greenwood, R. (2006). Use of the Barthel Index and the Functional Independence Measure during early inpatient rehabilitation after single incident brain injury. *Clinical Rehabilitation*, **20**, 153–159.

Indredavik, B., Bakke, F., Solberg, R. *et al.* (1991). Benefit of a stroke unit: a randomized controlled trial. *Stroke*, 22, 1026–1031.

Jette, A., & Haley, S. (2005). Contemporary measurement techniques for rehabilitation outcomes assessment. *Journal of Rehabilitation Medicine*, 37, 339–345.

Juby, L. C., Lincoln, N. B., Berman, P. *et al.* (1996). The effect of a stroke rehabilitation unit on functional and psychological outcome: a randomised controlled trial. *Cerebrovascular Diseases*, 6, 106–110.

Kidd, D., Stewart, G., Baldry, J. *et al.* (1995). The functional independence measure: a comparative validity and reliability study. *Disability and Rehabilitation*, 17, 10–14. Kinsella, G., Murtagh, D., Landry, A. *et al.* (1996). Everyday memory following traumatic brain injury. *Brain Injury*, 10, 499–507.

Kreutzer, J. S., Marwitz, J. H., Seel, R., & Serio, C. D. (1996). Validation of a neurobehavioral functioning inventory for adults with traumatic brain injury. *Archives of Physical Medicine and Rehabilitation*, 77, 116–124.

Kurtzke, J. F. (1983). Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology*, **33**, 1444–1452.

Levin, H. S., High, W. M., Goethe, K. E. *et al.* (1987). The Neurobehavioral Rating Scale: assessment of the behavioral sequelae of head injury by the clinician. *Journal of Neurology, Neurosurgery, and Psychiatry*, **50**, 183–193.

Lincoln, N. B., & Bowen, A. (2006). The need for randomised treatment studies in neglect research. *Restorative Neurology and Neuroscience*, 24, 401–408.

Lincoln, N. B., & Leadbitter, D. (1979). Assessment of motor function in stroke patients. *Physiotherapy*, **65**, 48–51. Lincoln, N. B., Jackson, J. M., & Adams, S. A. (1998). Reliability and revision of the Nottingham Sensory Assessment for stroke patients. *Physiotherapy*, **84**, 358–365.

Lincoln, N. B., Sutcliffe, L. M., & Unsworth, G. (2000). Validation of the Stroke Aphasic Depression Questionnaire (SADQ) for use with patients in hospital. *Clinical Neuropsychological Assessment*, 1, 88–96.

Linscott, R. J., Knight, R. G., & Godefrey, H. P. D. (1996). The Profile of Functional Impairment in Communication (PFIC): a measure of communication impairment for clinical use. *Brain Injury*, **10**, 397–412.

Maddicks, R., Marzillier, S. L., & Parker, G. (2003). Rehabilitation of unilateral neglect in the acute recovery stage: the efficacy of limb activation therapy. *Neuropsychological Rehabilitation*, 13, 391–408.

Mahoney, F. I., & Barthel, D. W. (1965). Functional evaluation: Barthel Index. *Maryland State Medical Journal*, 14, 61–65.

McColl, M. A., Davies, D., Carlson, P., Johnston, J., & Minnes, P. (2001). The community integration measure: development and preliminary validation. *Archives of Physical Medicine and Rehabilitation*, **82**, 429–434.

McKinney, M., Blake, H., Treece, K. A. *et al.* (2002). Evaluation of cognitive assessment in stroke rehabilitation. *Clinical Rehabilitation*, **16**, 129–136.

McPherson, K. M., Pentland, B., Cudmore, S. F., & Prescott, R. J. (1996). An inter-rater reliability study of the Functional Assessment Measure (FIM + FAM). *Disability and Rehabilitation*, **18**, 341–347.

Mermis, B. J. (2005). Developing a taxonomy for rehabilitation outcome measurement. *Rehabilitation Psychology*, **50**, 15–23.

Michalon, M., Eskes, G. A., & Mate-Kole, C. C. (1997). Effects of light therapy on neuropsychological function and mood in seasonal affective disorder. *Journal of Psychiatry and Neuroscience*, 22, 19–28.

Nickels, L. (2005). Tried, tested and trusted? Language assessment for rehabilitation. In P. W. Halligan & D. Wade (Eds.), *Effectiveness of Rehabilitation for Cognitive Deficits* (pp. 169–184). Oxford, UK: Oxford University Press.

Nilsson, A. L., Sunnerhagen, K. S., & Grimby, G. (2005). Scoring alternatives for FIM in neurological disorders applying Rasch analysis. *Acta Neurologicalica Scandinavica*, 111, 264–273.

Norris, G., & Tate, R. L. (2000). The Behavioural Assessment of the Dysexecutive Syndrome (BADS): ecological, concurrent and construct validity. *Neuropsychological Rehabilitation*, 10, 33–45.

Nouri, F. M., & Lincoln, N. B. (1987). An extended ADL scale for use with stroke patients. *Clinical Rehabilitation*, 1, 301–305.

Piazzini, A., Canevini, M. P., Maggiori, G., & Canger, R. (2001). The perception of memory failures in patients with epilepsy. *European Journal of Neurology*, 8, 613–620.

Pollak, N., Rheault, W., & Stoecker, J. L. (1996). Reliability and validity of the FIM for persons aged 80 years and above from a multilevel continuing care retirement community. *Archives ofPhysical Medicine and Rehabilitation*, 77, 1056–1061.

Ponsford, J., & Kinsella, G. (1991). The use of a rating scale of attentional behavior. *Neuropsychological Rehabilitation*, 1, 241–257.

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

Powell, J., Heslin, J., & Greenwood, R. (2002). Community based rehabilitation after severe traumatic brain injury: a randomised controlled trial. *Journal of Neurology, Neurosurgery, and Psychiatry*, 72, 193–202.

Radloff, L. S. (1977). The CES-D Scale: a self-report depression scale for research in the general population. *Applied Psychological Measurement*, 1, 385–401.

Robertson, I. H., Hogg, K., & McMillan, T. M. (1998). Rehabilitation of unilateral neglect: improving function by contralesional limb activation. *Neuropsychological Rehabilitation*, 8, 19–29.

Robertson, I. H., Ward, T., Ridgeway, V., & Nimmo-Smith, I. (1994). *The Test of Everyday Attention*. Titchfield: Thames Valley Test Company.

Ruisel, I. (1991). Memory rating. *Studia Psychologica*, 33, 71–77.

Sarno, M. T. (1969). *The Functional Communication Profile Manual of Directions*. New York, NY: Institute of Rehabilitation Medicine, New York University Medical Center.

Sharrack, B., & Hughes, R. A. C. (1999). The Guy's Neurological Disability Scale (GNDS): a new disability measure for multiple sclerosis. *Multiple Sclerosis*, 5, 223–233.

Skinner, C., Wirz, S., Thompson, I., & Davidson, J. (1984). *The Edinburgh Functional Communication Profile*. Oxford: Winslow Press.

Sohlberg, M. M., McLaughlin, K. A., Pavese, A., Heidrich, A., & Posner, M. I. (2000). Evaluation of attention process training and brain injury education in persons with acquired brain injury. *Journal of Clinical and Experimental Neuropsychology*, 22, 656–676.

Stern, R. A. (1997). *Visual Analog Mood Scales*. Odessa, FL: Psychological Assessment Resources, Inc.

Stewart, A. L., & Ware, J. E. (1992). *Measuring Functioning and Well-Being: The Medical Outcomes Study Approach*. Durham, NC: Duke University Press.

Sunderland, A., Harris, J. E., & Baddeley, A. D. (1983). Do laboratory tests predict everyday memory? A neuropsychological study. *Journal of Verbal Learning and Verbal Behavior*, 22, 341–357.

Sunderland, A., Harris, J. E., & Gleave, J. (1984). Memory failures in everyday life following severe head injury. *Journal of Clinical Neuropsychology*, 6, 127–142.

Sutcliffe, L., & Lincoln, N. (1998). The assessment of depression in aphasic stroke patients: the development of the Stroke Aphasic Depression Questionnaire. *Clinical Rehabilitation*, 12, 506–513.

Taylor, R. (1990). Relationships between cognitive test performance and everyday cognitive difficulties in multiple sclerosis. *British Journal of Clinical Psychology*, 29, 251–253.

Tham, K., & Tegnér, R. (1996). The Baking Tray Task: a test of spatial neglect. *Neuropsychological Rehabilitation*, 6, 19–25.

Tinson, D. J., & Lincoln, N. B. (1987). Subjective memory impairment after stroke. *International Disability Studies*, 9, 6–9.

Towle, D., & Lincoln, N. B. (1991). Development of a questionnaire for detecting everyday problems in stroke

patients with unilateral visual neglect. *Clinical Rehabilitation*, *5*, 135–140.

Turner-Stokes, L., & Hassan, N. (2002). Depression after stroke: a review of the evidence base to inform the development of an integrated care pathway. Part 1: Diagnosis, frequency and impact. *ClinicalRehabilitation*, **16**, 231–247.

van der Putten, J. J. M. F., Hobart, J. C., Freeman, J. A., & Thompson, A. J. (1999). Measuring change in disability after inpatient rehabilitation: comparison of the responsiveness of the Barthel Index and the Functional Independence Measure. *Journal of Neurology, Neurosurgery, and Psychiatry*, **66**, 480–484.

van Zomeren, A. H., & Spikman, J. M. (2003). Assessment of attention. In P. W. Halligan, U. Kischka, & J. C. Marshall (Eds.), *Handbook of Clinical Neuropsychology* (pp. 73– 88). Oxford, UK: Oxford University Press.

Vaney, C., Blaurock, H., Gattlen, B., & Meisels, C. (1996). Assessing mobility in multiple sclerosis using the Rivermead Mobility Index and gait speed. *Clinical Rehabilitation*, **10**, 216–226.

von Cramon, D. Y., Matthes-von Cramon, G., & Mai, N. (1991). Problem solving deficits in brain-injured patients: a therapeutic approach. *Neuropsychological Rehabilitation*, *1*, 45–64.

Wade, D. T. (2003). Outcome measures for clinical rehabilitation trials: impairment, function, quality of life, or value? *American Journal of Physical Medicine and Rehabilitation*, **82**, S26–S31.

Wade, D. T., Collen, F. M., Robb, G. F., & Warlow, C. P. (1992). Physiotherapy intervention late after stroke and mobility. *British Medical Journal*, *304*, 609–613.

Wade, D. T., Crawford, S., Wenden, F. J., King, N. S., & Moss, N. E. G. (1997). Does routine follow up after head injury help? A randomised controlled trial. *Journal of Neurology, Neurosurgery, and Psychiatry*, **62**, 478–484. Walker, R., Findlay, J. M., Young, A. W., & Welch, J. (1991). Disentangling neglect and hemianopia. *Neuropsychologia*, *29*, 1019–1027.

Ware, J. E., Snow, K. K., Kosinski, M., & Gandek, B. (1993). *SF-36 Health Survey Manual and Interpretation Guide*. Boston, MA: The Health Institute, New England Medical Center. Watkins, C. L., Auton, M. F., Deans, C. F. *et al.* (2007). Motivational interviewing early after acute stroke: a randomized, controlled trial. *Stroke*, **38**, 1004–1009.

Weniger, D. (1990). Diagnostic tests as tools of assessment and models of information processing: a gap to bridge. *Aphasiology*, *4*, 109–113.

Whiteneck, G., Charlifues, W., Gerhart, K., Overholser, J. D., & Richardson, G. N. (1992). Quantifying handicap: a new measure of long term rehabilitation outcomes. *Archives of Physical Medicine and Rehabilitation*, *73*, 519–526.

Whyte, J., Hart, T., Bode, R. K., & Malec, J. F. (2003). The Moss Attention Rating Scale for traumatic brain injury: Initial psychometric assessment. *Archives of Physical Medicine and Rehabilitation*, **84**, 268–276.

Wilson, B. A., Alderman, N., Burgess, P., Emslie, H., & Evans, J. (1996). *Behavioural Assessment of the Dysexecutive Syndrome*. Bury St Edmunds, UK: Thames Valley Test Company.

Wilson, B. A., Clare, L., Baddeley, A., Watson, P., & Tate, R. (1998). *Rivermead Behavioural Memory Test – Extended Version (RBMT-E)*. Bury St Edmunds, UK: Thames Valley Test Company.

Wilson, B. A., Cockburn, J., & Baddeley, A. (1985). *The Rivermead Behavioural Memory Test*. Bury St Edmunds, UK: Thames Valley Test Company.

Wilson, B. A., Cockburn, J., & Halligan, P. W. (1987). *Behavioural Inattention Test*. Bury St Edmunds, UK: Thames Valley Test Company.

Winward, C. E., Halligan, P. W., & Wade, D. T. (2002). The Rivermead Assessment of Somatosensory Performance (RASP): standardization and reliability data. *Clinical Rehabilitation*, **16**, 523–533.

World Health Organisation (2001). *International Classification of Functioning, Disability and Health*. Geneva, Switzerland: WHO.

Worthington, A. (2005). Rehabilitation of executive deficits: Effective treatments of related disabilities. In P. W. Halligan & D. Wade (Eds.), *Effectiveness of Rehabilitation for Cognitive Deficits* (pp. 257–267) Oxford: Oxford University Press.

Zelinski, E. M., Gilewski, M. J., & Anthony-Bergstone, C. R. (1990). Memory Functioning Questionnaire: concurrent validity with memory performance and self-reported memory failures. *Psychology and Aging*, *5*, 388–399.

Zigmund, A. S., & Snaith, R. P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavia*, *67*, 361–370.

Zoccolotti, P., & Judica, A. (1992). Functional evaluation of hemineglect by means of a semi-structured scale: personal extrapersonal differentiation. *Neuropsychological Rehabilitation*, *2*, 33–44.

Zoccolotti, P., Antonucci, G., & Judica, A. (1992). Psychometric characteristics of two semi-structured scales for the functional evaluation of hemi-inattention in extrapersonal and personal space. *Neuropsychological Rehabilitation*, *2*, 179–191.