Ecological aspects
of cognitive assessment

Ecological aspects
of cognitive assessment

Proefschrift

Ter verkrijging van de graad van doctor
aan de Universiteit Maastricht,
op gezag van Rector Magnificus, Prof. mr. G.P.M.F. Mols,
volgens het besluit van het College van Decanen,
in het openbaar te verdedigen op 29 april 2009 om 14.00 uur

door

Sharon Franciska Maria Bouwens

Geboren op 6 februari 1982 te Oirschot

Promotor
    Prof. dr. F.R.J. Verhey

Co-promotor
    Dr. C.M. van Heugten

Beoordelingscommissie
    Prof. dr. J.M.G.A. Schols (voorzitter)
    Prof. dr. M. Limburg
    Dr. R.W.H.M. Ponds
    Prof. dr. E. Salmon (Université de Liège)
    Prof. dr. D.T. Wade

Contents

Chapter 0

Prologue

This thesis deals with the ecological aspects of cognitive assessment. The following cases illustrate common issues involved in the investigation of the association between the cognitive assessment used to identify cognitive disorders and the daily life functioning (in relation to the individual patient) of people with brain dysfunction.

Case A

The first case concerns a 42-year-old male who is a former forklift truck driver with a medium level of education. He suffered from contusio cerebri accompanied by a right frontoparietal subdural haematoma. He lives alone. His cognitive assessment shows deficits with regard to: episodic memory for verbal information; episodic memory for visuo-spatial memory; selective and sustained attention; basal speed; inhibition; flexibility; and task shifting. In daily life, he experiences memory problems (e.g. forgets to feed his dog, to open his mail and to lock his car). Also he has severe problems with planning and organising (e.g. he often fails to keep appointments and to pay his bills). He receives professional ambulatory support for the problems he experiences in daily life in the form of assistance with his mail and financial issues.

This case illustrates a positive association between deficits identified by a cognitive assessment and the experience of problems in daily life.

Case B

The second case concerns a 54-year-old male who is a former chief executive with a high level of education. He had a recurrent right-sided cerebro vascular accident and was treated for a brain aneurysm. He is divorced, lives independently and has two daughters. His cognitive assessment shows deficits with regard to: episodic memory for verbal information; episodic memory for visuo-spatial memory; selective and sustained attention; basal speed; planning; flexibility; and task shifting. However, in daily life he does not have any evident cognitive problems.

This case illustrates that severe deficits identified by a cognitive assessment do not necessarily translate into daily life problems.

Case C

The third case concerns a 49-year-old female who is a former nurse with a medium level of education. She had a subarachnoidal haemorrhage. She is married and lives with her husband. Her cognitive assessment does not show any deficits. In daily life, however, she has problems with: executive functioning (e.g. using the mobile phone); memory (e.g. recalling previous conversations and appointments); and she experiences difficulties with multi-tasking.

This case illustrates that a normal score on a cognitive assessment does not necessarily indicate the absence of daily life problems.

These three patients have cognitive disorders due to brain dysfunction, but each reveals a different association between the results of a cognitive assessment and daily life functioning. These cases give rise to the following questions:

- What is the association between cognition and daily life functioning in people suffering from cognitive disorders due to brain dysfunction?
- What is the ecological validity of neuropsychological tests?
- How can treatment be tailored to the needs of the individual patient?

These questions form the basis for the research described in this thesis.

Chapter 1

General introduction

Defining the problem

A growing number of people worldwide are affected by neurological disorders associated with brain injury and dementia [1]. These brain dysfunctions can lead to physical deficits, as well as to deficits in neuropsychological domains such as cognition, emotion, and behaviour [2, 3]; thus they often have a far-reaching effect on patients' daily life functioning.

The International Classification of Functioning, Disability, and Health (ICF) of the World Health Organization [4], distinguishes three components related to functioning and disability: body function and structure (impairment); activities (limitation) and participation (restriction); and environmental factors (context). Impairments (e.g. concentration loss) can lead to limitations (e.g. not paying attention when someone is talking), and these limitations may lead to restrictions (e.g. not being able to work as a receptionist). Personal and environmental factors may further influence the association between the levels. Figure 1 illustrates the interaction of these components of functioning.

However, as reported in case B in Chapter 0, a deficit identified by a neuropsychological test (impairment) does not necessarily indicate a deficit in daily life functioning (restriction). Therefore, the extent to which daily life is affected by brain dysfunction varies from patient to patient. Furthermore, it is unclear what the association is between the domains which are affected in patients with brain dysfunction such as acquired brain injury and dementia.

Figure 1. Interaction of components of functioning (adapted from the World Health Organization, 2001)
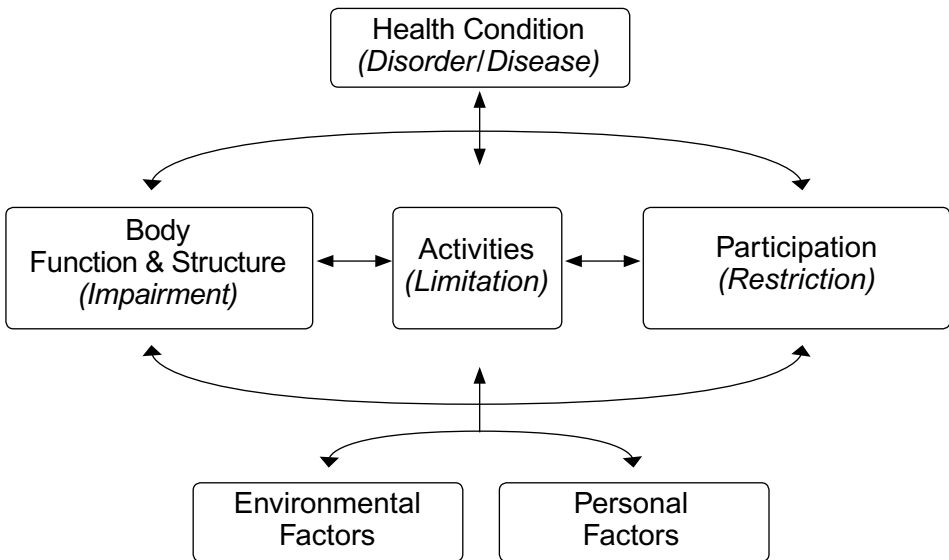
Association between domains and its importance for research and clinical practice

The domains affected in brain dysfunction are measured in several ways. Cognitive deficits are mostly assessed with neuropsychological tests [5], such as the Auditory Verbal Learning Test [6] which measures episodic memory. Emotional problems are usually assessed with questionnaires, such as the Symptom Check List 90 [7], and behavioural problems with informant-based interviews, such as the Neuropsychiatric Inventory [8]. The extent to which any of these domains can explain everyday functioning is unclear.

Due to the growing availability of neuroimaging techniques, neuropsychologists nowadays are more frequently asked to predict patients' daily life functioning, than to diagnose and localize brain impairments [9]. However, neuropsychological tests were developed to support the diagnosing of brain damage and not to predict daily life functioning [10]. Thus, although the cognitive pattern that arises from the neuropsychological assessment is now used to guide the clinician in planning treatment and future care, it is as yet unclear to what extent neuropsychological tests are capable of reflecting the daily life problems of patients with brain dysfunction. In other words, the ecological validity of cognitive tests is unclear.

High ecological validity of neuropsychological tests is important since the predictions based on these tests may have significant consequences for the lives of the patients and their caregivers [11]. Furthermore, the diagnostic and rehabilitation process often takes place in clinical, standardized settings. In these settings exact predictions of everyday life functioning cannot be made, since everyday life functioning depends on individual capacities that interact with specific environmental settings and social support networks and these domains are not covered in the clinical settings [12]. These factors influence the association between neuropsychological tests and daily life functioning.

Daily life functioning is often measured with the help of methods that are derivatives of everyday functioning (e.g. informant-based questionnaires such as the Lawton Instrumental Activities of Daily Living-scale [13], or the Blessed Dementia Scale [14]). These instruments are not based on direct observation but are informant-based and therefore challenge the validity of research findings with regard to the affected domains in patients with brain dysfunction. A better understanding of the association between the domains affected in these patients is important, both from a scientific and clinical perspective. From a scientific perspective, a better understanding will lead to the development of more sensitive instruments for detecting deficits that affect patients' daily life functioning. From a clinical perspective, a better understanding of the association will help clinicians to improve their treatment, formulate attainable treatment goals, and make future care more effective. It will also help patients to be better prepared for future functioning in society.

Patients' goals

As illustrated in Figure 1, besides the complex association between body function and structure, or activities and participation, the personal needs and preferences of the individual patient also play an important role in the successful treatment and future care of patients. Hence, two people with comparable neuropsychological profiles may have different perceptions of their ideal daily life functioning: the first patient may be happy when he is able to remember appointments with the help of an agenda, while for the other patient this outcome would be her worst nightmare. These personal preferences or individual goals are often not taken into account by clinicians and researchers and may bias both scientific and clinical results. From a scientific point of view, the effect of a treatment may be erroneously described as positive when patients improve as indicated by traditional, standardized measures, even though individual preferences are ignored. From a clinical point of view, clinicians may have different treatment goals from those of patients, which may bias treatment outcome and satisfaction levels.

Aims of this thesis

The general aim of this thesis is to investigate the ecological aspects of cognitive assessment. This is achieved by exploring the association between the domains that are affected in people with brain dysfunction (e.g. dementia and acquired brain injury). In this thesis, the following research questions will be answered:

1.   What is the association between cognition and daily life functioning in psycho-geriatric patients with cognitive disorders? This question will be answered in Part I.
2a.   To what extent can neuropsychologists predict the daily life functioning of people with acquired brain injury solely on the basis of neuropsychological data? This question will be answered in Part II.
2b.   Does the integration of a neuropsychological assessment with direct observation of the daily life functioning of people suffering from acquired brain injury have added value for treatment and future care? This question will also be answered in Part II.
3.   Is individual goal-setting attainable in the case of people with cognitive disorders due to dementia or acquired brain injury? This question will be answered in Part III.

Outline of the thesis

This thesis consists of three parts, each addressing different ecological aspects of cognitive assessment.

Part I: Association between cognition and daily life functioning

Chapter 2 describes the association between cognitive status and the daily life functioning of people with different types of dementia. Chapter 3 investigates the association between cognition and the direct observation of daily life functioning in a psychogeriatric population with cognitive disorders, and examines the influence of neuropsychiatric symptoms on this association.

Part II: Ecological validity of neuropsychological assessment and the integration of domains

Chapter 4 provides an introduction to Part II of the thesis. It elucidates aspects of the ecological validity of neuropsychological tests, since much confusion exists with regard to this topic and its importance. Chapter 5 investigates the ecological validity of neuropsychological assessment by exploring the extent to which experienced neuropsychologists can predict the daily life functioning of people with acquired brain injury solely on the basis of neuropsychological data. Chapter 6 describes an international cooperative project called PROFINTEG that attempts to enhance the ecological validity by integrating neuropsychological data and observational data in patients with brain dysfunction.

Part III: Individual goal-setting

Chapter 7 provides a systematic review of the feasibility of Goal Attainment Scaling (GAS) [15] in a psychogeriatric population with cognitive disorders. GAS is a method which enables patients to set individual goals and it measures the level of attainment of these goals. Chapter 8 shows the results of a study on the effectiveness of a post-acute outpatient community re-entry cognitive rehabilitation programme for patients with acquired brain injury and their relatives. GAS was used as one of the outcome methods to investigate whether patients had attained their predefined goals after the intervention. Chapter 9 presents the results of a cross-sectional study on the applicability of GAS as an outcome measure in the case of cognitive rehabilitation programmes for people with acquired brain injury.

Chapter 10 summarizes the results of in this thesis in the form of a general discussion. Clinical implications are addressed and recommendations for further research are suggested.

References

1. World Health Organization. Neurological Disorders: Public health challenges. Geneva; 2007.
2. Ettenhofer ML, Abeles N. The significance of mild traumatic brain injury to cognition and self-reported symptoms in long-term recovery from injury. J Clin Exp Neuropsychol 2008;11:1-10.
3. Greenwood R. The consequences of brain injury: classification assessment of outcome. Neuropsychol Rehab 1999;9(3-4):231-240.
4. World Health Organization. International Classification of Functioning, Disability and Health. Geneva; 2001.
5. Lezak MD, Howleson DB, Loring DW, Hannay HJ. Neuropsychological assessment. 4th edition. New York: Oxford University Press 2004.
6. Rey A. L'examen clinique en psychologie [The clinical examination in psychology]. Paris: Presses Universitaires de France 1958.
7. Arrindell WA, Ettema JHM. Symptom Checklist [Handleiding bij een multidimensionele psychopathologie-indicator]. Lisse: Swets & Zeitlinger B.V. 2003.
8. Cummings JL, Mega M, Gray K, Rosenberg-Thompson S, Carusi DA, Gornbein J. The Neuropsychiatric Inventory: comprehensive assessment of psychopathology in dementia. Neurology 1994;44(12):2308-2314.
9. Kibby MY, Schmitter-Edgecombe M, Long CJ. Ecological validity of neuropsychological tests: focus on the California Verbal Learning Test and the Wisconsin Card Sorting Test. Arch Clin Neuropsychol 1998;13(6):523-534.
10. Sbordone RJ, Long CJ, editors. Ecological validity of neuropsychological testing. 2nd edition. Boca Raton: St. Lucie Press 1996.
11. Chaytor N, Schmitter-Edgecombe M. The ecological validity of neuropsychological tests: a review of the literature on everyday cognitive skills. Neuropsychol Rev 2003; 13(4):181-197.
12. Burgess PW, Alderman N, Evans J, Emslie H, Wilson BA. The ecological validity of tests of executive function. J Int Neuropsychol Soc 1998;4(6):547-558.
13. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. Gerontologist 1969;9(3):179-186.
14. Blessed G, Tomlinson BE, Roth M. The association between quantitative measures of dementia and of senile change in the cerebral grey matter of elderly subjects. Br J Psychiatry 1968;114(512):797-811.
15. Kiresuk T, Sherman RE. Goal Attainment Scaling: a general method for evaluating comprehensive community mental health programs. Community Ment Health J 1968;4(6):443-453.

Part I

Association between cognition
and daily life functioning

Chapter 2

Association between cognition and daily life functioning in dementia subtypes

Abstract

Objective
To investigate the association between cognition and daily life functioning in dementia subtypes.

Methods
Cross-sectional data were used from 615 patients with dementia who were referred to the Maastricht Memory Clinic of the Maastricht University Medical Centre. Pearson correlation coefficients were calculated between the Mini-Mental State Examination (MMSE; to measure cognitive status) and the Blessed Dementia Scale (BDS; to measure daily life functioning) for the following types of dementia: Alzheimer's disease (AD, n = 442); Vascular dementia (VaD, n = 113); frontotemporal dementia (FTD, n = 18); Parkinson's dementia (PD, n = 21); and primary progressive aphasia (PPA, n = 21). One-way ANOVA was used to test differences in age, MMSE scores and BDS scores across dementia subtypes.

Results
Scores on the MMSE showed strong correlation with BDS scores in cases of FTD (r = -0.80); moderate correlation in cases of AD, VaD, and PD (range r = -0.50 to 0.60); while no correlation was found in PPA cases.

Conclusions
The association between cognition and daily life functioning varied among dementia subtypes for AD, VaD, FTD and PD. Furthermore, the overall scores on both domains differ between dementia subtypes, indicating that different types of dementia are characterized by a specific pattern of cognitive status and daily life functioning. These findings underline the need for multidomain assessment in patients with dementia.

Introduction

Dementia, characterized by multidomain impairment, is a common disease in old age, and it can become manifest in several forms. Alzheimer's disease (AD) is the most common dementing disorder, followed by vascular dementia (VaD) [1]. Other forms like frontotemporal dementia (FTD) (including primary progressive aphasia), and Parkinson's dementia (PD) are less frequent. The severity of dementia is traditionally measured with cognitive tests [2-4]. However, the assessment of functional impairment due to cognitive impairment is a crucial element in the diagnosis of dementia [5, 6].

Various instruments are used to measure one or more domains of dementia, such as the Mini-Mental State Examination (MMSE) [7] and the Blessed Dementia Scale (BDS) [8]; both of which are recommended by the National Institute of Neurological Diseases and Communicative Disorders and Stroke, as well as the Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) [9].

Several studies have found that on average there is a strong association between cognition and daily functioning in people with dementia in general [10-12]. The association between cognitive status and daily life functioning in people with specific types of dementia has received less attention until now. Identifying the relation between cognitive status and daily life functioning for a specific dementia subtype may, however, be of diagnostic importance in practice. Furthermore, a better understanding of this association with regard to each dementia subtype is important, since it may improve insight into the different characteristics of the subtypes and may provide information about the presence of compensatory strategies for a patient's lack of abilities (i.e. daily life functioning may not be impaired if cognitive impairments are compensated adequately).

The aim of this study is to investigate the association between cognitive status and daily life functioning in people with different dementia subtypes.

Methods

Participants
Participants were patients with dementia from the Maastricht Memory Clinic (MMC) register of the Maastricht University Medical Centre (MUMC). Patients were referred by general practitioners, neurologists, psychiatrists, and community mental health teams, or by others who were mainly internists or geriatricians [13]. The study was based on the cross-sectional data of those people who met the criteria indicating one of the following subtypes of dementia: Alzheimer's disease (AD) [9]; Lewy body dementia (LBD) [14]; vascular dementia (VaD) [15]; frontotemporal dementia (FTD) [16]; Parkinson's dementia (PD) [14]; or primary progressive aphasia (PPA) [16]. Participants included also had to have an informant who could answer questions about their condition. Furthermore, their diagnostic phase data had to be complete and their informed consent obtained. The Ethics Committee of the MUMC approved the procedure.

Measures

The Mini-Mental State Examination (MMSE) [7] is a bedside test that measures cognitive status. The MMSE covers orientation in place and time, memory, attention, language, concentration, visuo-spatial skills, and praxis. The maximum score is 30, with a higher score reflecting a better global cognitive status. The cut-off score for dementia is 25 [17].

The Blessed Dementia Scale (BDS) [8] is a 30-minute, structured interview with an informant, who is close to the patient, about daily functioning, cognition and behaviour in people with dementia. The scale is based on the Dementia Scale (DS) and the Information-Memory-Concentration (IMC) Test. The former tracks changes in performance of everyday instrumental activities (items 1-8) and everyday basic activities (items 9-11), as well as changes in personality, interests and drive (items 12-22). The latter assesses orientation, memory and concentration.

For this study, we used items 1-11 of the DS, covering changes in performance of everyday instrumental and basic activities. The DS consists of items measuring both basic activities of daily life (BADL) and instrumental activities of daily life (IADL) functions [18, 19]. Changes in three BADL functions (eating, dressing and continence) are scored 0 to 3 and eight items assessing changes in IADL, such as performing household tasks, coping with money or finding one's way, are scored 0 (never), 0.5 (sometimes) or 1 (nearly always). The total score ranges from 0 (independent) to 17 (dependent), with a higher score reflecting a deterioration in daily life functioning. The MMSE and the BDS were both recommended by the NINCDS-ADRDA [9].

Procedure

Clinicians working at the Maastricht Memory Clinic (MMC) were trained in order to administer the MMSE and the BDS at the Maastricht University Medical Centre (MUMC) as part of the diagnostic phase at the MMC. Both rating scales were administered at the same visit. The caregiver was present during both assessments.

Statistical analyses

All analyses were calculated for dementia subtypes with an n >15. Demographic characteristics of the participants and descriptive statistics of the MMSE and the BDS were calculated. Normality tests were performed for the MMSE and the BDS. To test if dementia subtypes differed according to mean age, mean MMSE and mean BDS, One-Way ANOVAs were calculated with age, MMSE, and BDS as dependent variables and dementia subtypes as factor. Pearson correlation coefficients were calculated to examine the association between MMSE scores and BDS scores for dementia subtypes. Correlation coefficients between .30 and .49 were regarded as low, those between .50 and .69 as moderate, and those between .70 and .89 as high [20].

We hypothesized that the association between MMSE and BDS scores differed per dementia subtype with AD as reference group. To facilitate interpretation of the test results, the raw scores were converted into Z-scores because of large differences in score range (i.e. 0-30 for the MMSE and 0-17 for the BDS) and because the scoring of both tests is reversed (i.e. higher MMSE scores signify better performance, whilst higher BDS scores

signify worse performance). A repeated measures ANOVA with test score (MMSE Z-score, BDS Z-score) as within-subject variable and dementia subtype as between-subject variable was conducted to test this hypothesis. Statistical analyses were performed with the Statistical Package for Social Sciences (version 16.0) with an alpha level set at 0.05 for all analyses.

Results

Of the 965 people with dementia in the MMC register [13], 621 patients fulfilled the inclusion criteria of this study; 442 patients had AD, 6 patients had LBD, 113 patients had VaD, 18 patients had FTD, 21 patients had PD, and 21 had PPA. Since the number of patients with LBD was too low to perform meaningful analyses, we did not include these for further analyses, and 615 patients remained for the present study. Table 1 shows the patients' characteristics and mean scores on the MMSE and the BDS for the total group of people with dementia (N = 615) and for the groups per dementia subtype separately. Level of education did not differ between dementia subtypes.

Table 1. Patient characteristics and mean scores on the Mini Mental State Examination and the Blessed Dementia Scale

	N	Age	Female	MMSE	BDS
		Mean (SD, range)	n (%)	Mean (SD, range)	Mean (SD, range)
Diagnosis					
AD	442	72.3 (8.3, 49-95)	266 (60.2)	19.2 (5.7, 2-29)	3.5 (2.4, 0.5-14.0)
VaD	113	73.8 (7.2, 49-88)	51 (45.1)	18.9 (5.9, 5-30)	4.4 (3.1, 1.0-14.5)
FTD	18	63.0 (9.7, 49-77)	10 (55.6)	20.6 (6.9, 8-29)	2.9 (2.4, 0.5-8.5)
PD	21	73.9 (5.3, 59-82)	10 (47.6)	15.4 (4.7, 2-24)	5.9 (3.9, 1.5-15.0)
PPA	21	64.9 (6.8, 54-77)	10 (47.6)	20.6 (7.5, 4-28)	1.8 (1.4, 0.0-6.0)
Total	615	72.1 (8.3, 49-95)	347 (56.4)	19.1 (5.8, 2-30)	3.7 (2.7, 0.0-15.0)

MMSE: Mini Mental State Examination; BDS: Blessed Dementia Scale; AD: Alzheimer's disease; VaD: Vascular dementia; FTD: Frontotemporal dementia; PD: Parkinson's dementia; PPA: Primary progressive aphasia

The One-Way ANOVAs for mean age, MMSE and BDS for dementia subtypes all show a significant effect of dementia subtype on these mean scores (F = 11.578, p < 0.05; F = 2.954, p < 0.05; F = 9.974, p < 0.05 respectively) indicating that age and MMSE and BDS scores differ with regard to dementia subtype.

Figure 1. Association between the Mini Mental State Examination (MMSE) and the Blessed Dementia Scale (BDS) for five dementia subtypes. AD: Alzheimer's disease; VaD: Vascular Dementia; FTD: Frontotemporal Dementia; PPA: Primary Progressive Aphasia; PD: Parkinson's Dementia
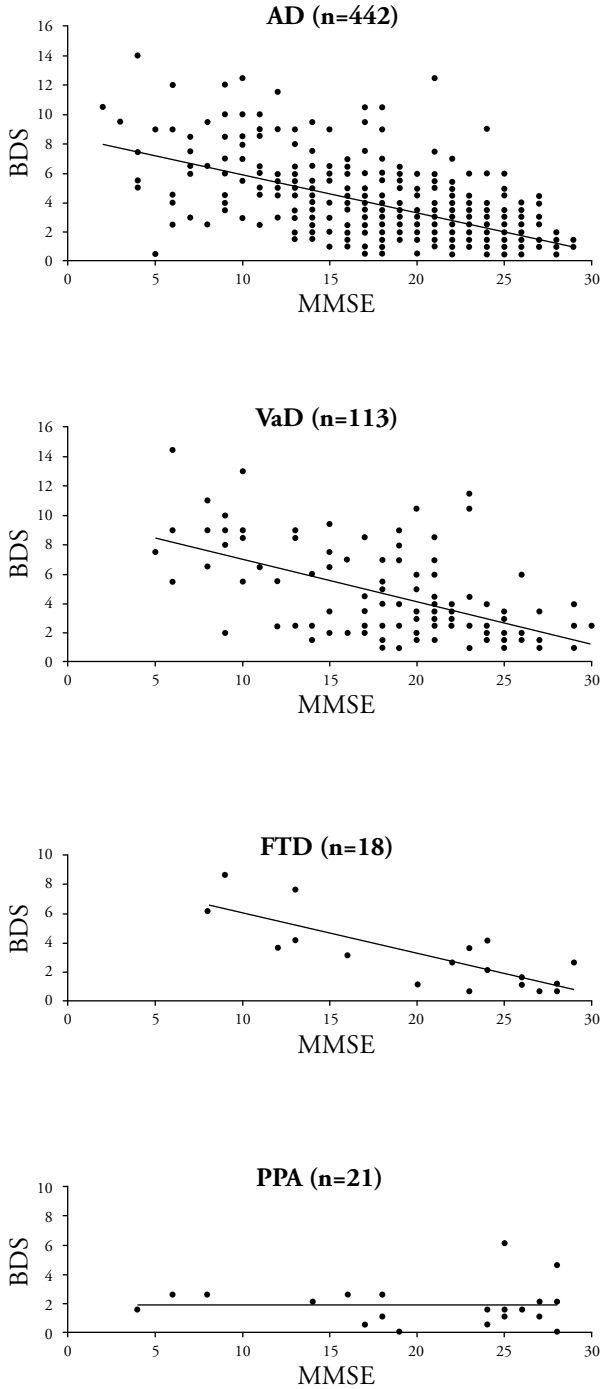
Figure 1. (continued)
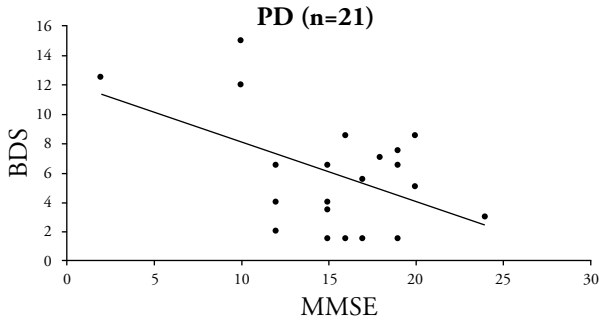
The Z-scores for both the MMSE and the BDS were normally distributed. The repeated measures ANOVA for mean MMSE Z-scores and mean BDS Z-scores for dementia subtypes show a significant effect of dementia subtype on the association between the MMSE and the BDS ($F = 7.123$, $p < 0.05$). This indicates that the association between cognitive status and daily life functioning depends on the specific dementia subtype when considering AD as reference group.

Discussion

We investigated the association between cognitive status (MMSE) and daily life functioning (BDS) of people with different types of dementia. We found that the association between cognitive status and daily life functioning is directly related to dementia subtype.

The extent of this association differs across dementia subtypes. This supports the importance of functional as well as cognitive assessment not only to assess patient care needs but also as a possible diagnostic tool. The moderate and high correlations between the MMSE and the BDS scores are expected since the interference of cognitive functioning on daily life functioning is part of the definition of dementia.

The absence of a significant correlation in the PPA group indicates that daily life functioning is not affected to the same extent as cognitive status, as indicated by the MMSE. An explanation for the absence of a significant correlation in this dementia subtype is that patients with aphasia score relatively worse on cognitive tests like the MMSE because most of the items in this test require verbal skills. However, these verbal skills do not affect the BDS because this is an informant-based rating of daily life functioning. Moreover, patients may have good compensatory abilities, which help them cope with aphasic problems in daily life. These findings are in line with Osher et al. [21], and Mioshi et al. [22]. Clinicians should take this verbal handicap into account when testing cognitive status in people with aphasia and make better use of non-verbal cognitive tests like the Aphasia Check List (ACL) [23]. The ACL is a 30-minute test battery for the assessment of aphasic and associated cognitive disorders. It includes nonverbal screening tests for three neuropsychological domains: memory, attention, and reasoning.

Furthermore, Mioshi et al. [22] also found that frontotemporal dementia has a negative effect on activities of daily living. Since other studies, including ours and that of Larner and Hancock [24], have found different results, future research is necessary to study the effect of FTD on daily life functioning. Besides, the high correlation between the MMSE and the BDS in patients with FTD should be analysed more deeply by looking for sub-items of the MMSE and the BDS that are particularly impaired in the FTD group.

An unexpected result of our study was the comparable correlation between the MMSE and the BDS for AD and PD. This may be due to overlap in old criteria for PD and AD (DSM-III-R and DSM-IV) [25], by which the specific characteristics of each dementia subtype were not taken into account. However, there is an emerging consensus that dementia subtypes differ in clinical characteristics, which is reflected in recent proposals for specific dementing disorders [14, 15, 26].

The main strength of our study lies in the use of a large sample of clinical data. As a consequence, the dataset contained information on a heterogeneous population as it appears in everyday clinical practice, which may improve the general applicability of the results. Our data are representative according to distribution of dementia subtypes, gender in the total group and across subtypes, mean age per subtype, MMSE scores and BDS scores [1, 27-30]. Although our sample is in general large and representative, the number of patients with specific types of dementia, such as FTD, PD and PPA are relatively small. Therefore, these results must be interpreted with some caution.

Harwood et al. [31] and Bouwens et al. [5] found influences of neuropsychiatric symptoms on the cognitive status and daily life functioning of people with dementia. We did not take neuropsychiatric data into account. The association between cognition and functioning further depends on the type of IADL measures used, and whether they are based on informant-reporting, self-reporting or clinician ratings [32]. Nowadays, IADL [33] is recommended instead of the BDS which was chosen on the basis of the NINCDS-ADRDA [9] from 1984.

Discrepancies between cognitive tests and instruments for measuring activities of daily life, as is the case in our PPA group, provide important information that can guide the clinician in choosing the appropriate treatment. Since PPA, PD and FTD are relatively rare subtypes of dementia with different patterns between cognitive status and daily life functioning, clinicians must pay special attention to the possible presence of one of these three subtypes when finding relative discrepancies between cognitive status and daily life functioning.

Conclusions

Different types of dementia are characterized by different patterns of cognitive status and daily life functioning. These findings underline the need for multidomain assessment in patients with dementia. Clinicians should take into account the different patterns between cognitive status and daily life functioning for each dementia subtype in order to provide appropriate care as well as adequate and accurate information. Discrepancies between cognition and daily life functioning have diagnostic relevance.

References

1. Qiu C, de Ronchi D, Fratiglioni L. The epidemiology of the dementias: an update. Curr Opin Psychiatry 2007;20:380-385.
2. Fratiglioni L, Grut M, Forsell Y, Viitanen M, Grafstrom M, Holmen K, et al. Prevalence of Alzheimer's disease and other dementias in an elderly urban population: relationship with age, sex, and education. Neurology 1991;41(12):1886-1892.
3. O'Connor DW, Pollitt PA, Hyde JB, Fellows JL, Miller ND, Brook CP, et al. The reliability and validity of the Mini-Mental State in a British community survey. J Psychiatr Res 1989;23(1):87-96.
4. Zhang MY, Katzman R, Salmon D, Jin H, Cai GJ, Wang ZY, et al. The prevalence of dementia and Alzheimer's disease in Shanghai, China: impact of age, gender, and education. Ann Neurol 1990;27(4):428-437.
5. Bouwens SF, van Heugten CM, Aalten P, Wolfs CA, Baarends EM, van Menxel DA, et al. Relationship between measures of dementia severity and observation of daily life functioning as measured with the Assessment of Motor and Process Skills (AMPS). Dement Geriatr Cogn Disord 2008;25(1):81-87.
6. Hendrie HC, Lane KA, Ogunniyi A, Baiyewu O, Gureje O, Evans R, et al. The development of a semi-structured home interview (CHIF) to directly assess function in cognitively impaired elderly people in two cultures. Int Psychogeriatr 2006;18(4): 653-666.
7. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975;12(3): 189-198.
8. Blessed G, Tomlinson BE, Roth M. The association between quantitative measures of dementia and of senile change in the cerebral grey matter of elderly subjects. Br J Psychiatry 1968;114(512):797-811.
9. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. Neurology 1984;34(7):939-944.
10. Barberger-Gateau P, Commenges D, Gagnon M, Letenneur L, Sauvel C, Dartigues JF. Instrumental activities of daily living as a screening tool for cognitive impairment and dementia in elderly community dwellers. J Am Geriatr Soc 1992;40(11):1129-1134.
11. Reed BR, Jagust WJ, Seab JP. Mental status as a predictor of daily function in progressive dementia. Gerontologist 1989;29:804-807.
12. Skurla E, Rogers JC, Sunderland T. Direct assessment of activities of daily living in Alzheimer's disease. A controlled study. J Am Geriatr Soc 1988;36(2):97-103.
13. Verhey FR, Jolles J, Ponds RW, Rozendaal N, Plugge LA, de Vet RC, et al. Diagnosing dementia: a comparison between a monodisciplinary and a multidisciplinary approach. J Neuropsychiatry Clin Neurosci 1993;5(1):78-85.

14. McKeith IG, Galasko D, Kosaka K, Perry EK, Dickson DW, Hansen LA, et al. Consensus guidelines for the clinical and pathologic diagnosis of dementia with Lewy bodies (DLB): report of the consortium on DLB international workshop. Neurology 1996;47(5):1113-1124.

15. Roman GC, Tatemichi TK, Erkinjuntti T, Cummings JL, Masdeu JC, Garcia JH, et al. Vascular dementia: diagnostic criteria for research studies. Report of the NINDS-AIREN International Workshop. Neurology 1993;43(2):250-260.

16. The Lund and Manchester Groups. Clinical and neuropathological criteria for frontotemporal dementia. J Neurol Neurosurg Psychiatry 1994;57(4):416-418.

17. Braekhus A, Laake K, Engedal K. A low, 'normal' score on the Mini-Mental State Examination predicts development of dementia after three years. J Am Geriatr Soc 1995;43(6):656-661.

18. Erkinjuntti T, Hokkanen L, Sulkava R, Palo J. The blessed dementia scale as a screening test for dementia. Int J Geriatr Psychiatry 1988;3(4):267-273.

19. Pena-Casanova J, Monllau A, Bohm P, Aguilar M, Sol JM, Hernandez G, et al. [Diagnostic value and test-retest reliability of the Blessed Dementia Rating Scale for Alzheimer's disease: data from the NORMACODEM project]. Neurologia 2005;20(7):349-355.

20. Hinkle DE, Wiersma W, Jurs SG. Applied statistics for the behavioral sciences. New York: Houghton Mifflin Company 1998.

21. Osher JE, Wicklund AH, Rademaker A, Johnson N, Weintraub S. The Mini-Mental State Examination in Behavioral Variant Frontotemporal Dementia and Primary Progressive Aphasia. Am J Alzheimers Dis Other Demen 2008;22(6):468-473.

22. Mioshi E, Kipps CM, Dawson K, Mitchell J, Graham A, Hodges JR. Activities of daily living in frontotemporal dementia and Alzheimer disease. Neurology 2007 12;68(24):2077-2084.

23. Kalbe E, Reinhold N, Brand M, Markowitsch HJ, Kessler J. A new test battery to assess aphasic disturbances and associated cognitive dysfunctions - German normative data on the aphasia check list. J Clin Exp Neuropsychol 2005;27(7):779-794.

24. Larner AJ, Hancock P. Re: Activities of daily living in frontotemporal dementia and Alzheimer disease. Neurology 2008;70(8):658.

25. American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision. Washington, DC, American Psychiatric Association 2000.

26. Dubois B, Burn D, Goetz C, Aarsland D, Brown RG, Broe GA, et al. Diagnostic procedures for Parkinson's disease dementia: recommendations from the movement disorder society task force. Mov Disord 2007;22(16):2314-2324.

27. Chow TW, Hynan LS, Lipton AM. MMSE scores decline at a greater rate in frontotemporal degeneration than in AD. Dement Geriatr Cogn Disord 2006;22(3): 194-199.

28. Juva K, Makela M, Erkinjuntti T, Sulkava R, Ylikoski R, Valvanne J, et al. Functional assessment scales in detecting dementia. Age Ageing 1997;26(5):393-400.

29. Neary D, Snowden JS, Gustafson L, Passant U, Stuss D, Black S, et al. Frontotemporal lobar degeneration: a consensus on clinical diagnostic criteria. Neurology 1998;51: 1546-1554.

30. Starr JM, Loanie J. The influence of pre-morbid IQ on mini-Mental Sate Examination score at time of dementia presentation. Int J Geriatr Psychiatry 2007;22:382-384.

31. Harwood DG, Barker WW, Ownby RL, Duara R. Relationship of behavioral and psychological symptoms to cognitive impairment and functional status in Alzheimer's disease. Int J Geriatr Psychiatry 2000;15(5):393-400.

32. Burton CL, Strauss E, Hultsch DF, Hunter MA. Cognitive functioning and everyday problem solving in older adults. Clin Neuropsychol 2006;20(3):432-452.

33. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. Gerontologist 1969;9(3):179-186.

Chapter 3

Relationship between measures of
dementia severity and observation of daily life
functioning as measured with the Assessment of
Motor and Process Skills (AMPS)

Abstract

Background
Cognitive impairment is mostly regarded as the core symptom of dementia, but several other domains (such as daily functioning) are as relevant to assess the severity of dementia. The relationship between these domains is unclear. The AMPS (Assessment of Motor and Process Skills) is a relatively unexplored instrument in people with dementia, measuring severity by direct observation.

Objective
To study the relationship between the AMPS and scores on several commonly used outcome measures for the assessment of dementia severity, and to examine the possible influence of neuropsychiatric symptoms on these relationships in patients with cognitive disorders.

Methods
Cross-sectional data of 118 patients with cognitive disorders were used; data on cognition (MMSE, CAMCOG), global severity (GDS), daily life functioning (IADL), and neuropsychiatric symptoms (NPI) were collected and analyzed using correlation and regression analyses. Different combinations of the severity measures were tested for their ability to predict the AMPS process ability scores.

Results
Scores on the MMSE, CAMCOG and GDS were moderately associated with the AMPS process ability score. These measures explained between 27% and 44% of the variance in the AMPS score. The presence of apathy influenced the association between the cognitive measures and the AMPS score.

Conclusion
Commonly used measures of dementia severity are only moderately associated with observation of performance on daily activities. This underlines the need for direct observation of daily activities in dementia patients. This relationship between several approaches of assessing dementia severity needs further study.

Introduction

Dementia is defined as a syndrome of memory impairment and other cognitive deficits severe enough to cause significant interference with activities of daily life [1]. It is important to be able to assess the severity of dementia in order to arrange appropriate care and to evaluate its effect. Several instruments are available, among which those measuring cognitive function, such as the Mini-Mental State Examination (MMSE) [2], informant questionnaires, such as the Instrumental Activities of Daily Living (IADL) [3], and global assessment taking into account a combination of aspects, such as the Global Deterioration Scale (GDS) [4] and the Clinical Dementia Rating (CDR) [5]. Cognitive impairment is most often seen as the core symptom of dementia. For instance, both in research and clinical settings, the MMSE score is used the most often, for example to define study groups, or to determine whether antidementia drugs are indicated [6]. In clinical practice, however, level of functioning in daily life might be of more relevance. Therefore several instruments that measure cognitive status, such as the MMSE, are taken as a measure of severity. However, dementia not only involves cognition but also other domains like functional ability and neuropsychiatric functioning, aspects that may be of more relevance in daily clinical practice. The relationships between these domains are complex and until now not well explored.

Functioning in daily life is ideally assessed by direct observation of patients performing activities important to their daily lives and in the ecologically appropriate environment [7]. The Assessment of Motor and Process Skills (AMPS) [8] is an instrument for the direct observation of daily activities in a structured manner. Activities differ in complexity and only activities relevant to the patient are selected, thus ensuring high ecological validity [9].

The AMPS is not influenced by possible bias in the caregiver's judgment because clinical staff observe and score the patient's performance [10, 11]. Especially the process score is a sensitive indicator of the probable need for assistance for community living [12].

The AMPS has been used to assess the functioning of a variety of clinical populations, such as brain injured patients [13-15], and psychiatric patients [16, 17]. The instrument has also been studied in patients with dementia [18-21]. The aim of our study is to provide a deeper understanding of how different domains of dementia severity relate to performance on daily activities. In addition, we evaluate the possible effects of neuropsychiatric symptoms on the relationship between the AMPS and other commonly used measures.

Methods

Participants

The current study was part of the larger MEDICIE study (Maastricht Evaluation of a Diagnostic Intervention for Cognitively Impaired Elderly), a randomized clinical trial into the clinical and health-economic effects of an integrated multidisciplinary approach for people with mild to moderate dementia (Diagnostic Observation Centre for PsychoGeriatric patients, DOC-PG) compared with traditional care [22]. Our study contains cross-sectional data of patients who were included in the intervention-group of the MEDICIE study, who had an informant

available, of whom data at baseline were complete, and who gave informed consent. The Ethics Committee of the Maastricht University Hospital approved the procedure.

Measures

Skills necessary for daily functioning were measured with the Assessment of Motor and Process Skills (AMPS) [8]. This is an observational instrument, offering 83 tasks, categorized into personal activities and household or instrumental activities. These tasks are classified in terms of difficulty, and culture-dependent or gender related activities. The AMPS consists of a motor ability score, measuring impairments in motor behaviour, and a process ability score, measuring functional skills. For motor ability, scores range from -3.00 to 4.00 (cut-off +2.00), and for process ability they range from -4.00 to 3.00 (cut-off +1.00). Higher scores reflect better functioning. Patients were observed while performing two tasks that were chosen as the most familiar to them, in an ecologically natural space (kitchen). Two relevant tasks for each patient were selected. Administering the AMPS took about one hour for each patient.

Global level of deterioration was rated with the Global Deterioration Scale (GDS) [4]. The GDS is subdivided into 7 stages, ranging from no cognitive deterioration at all (stage 1) to severe cognitive and functional deterioration (stage 7).

Cognitive status was measured with the Mini-Mental State Examination (MMSE) [23] and the cognitive component (CAMCOG) of the revised Cambridge Examination for Mental Disorders of the Elderly (CAMDEX) [24]. The MMSE includes items of orientation in place and time, memory, attention, language, concentration, visuospatial skills, and praxis. The maximum score is 30, with a higher score reflecting a better global cognitive status. The CAMCOG was also used to measure the cognitive status of the patient. The subscale CAMCOG of the CAMDEX consists of seven cognitive domains, and assesses 1) orientation, 2) attention and calculation, 3) language, 4) memory, 5) praxis, 6) abstract thinking, and 7) perception. The maximum score is 105, with a low score reflecting severe cognitive problems.

The Instrumental Activities of Daily Living (IADL) [3] was used to assess the performance of everyday tasks, such as using the telephone, shopping, preparing food, housekeeping, laundering, mode of transportation, responsibility for one's own medications, and handling one's own finances. Each item has 3–5 response options, with higher scores being indicative of greater independence in performing the activity.

Neuropsychiatric symptoms were assessed with the Neuropsychiatric Inventory (NPI) [25, 26]. The NPI is a retrospective (to 1 month) informant-based rating scale developed to assess psychopathology in patients with neurodegenerative diseases. It is a semi-structured interview assessing 12 neuropsychiatric symptoms commonly observed in dementia: 1) delusions, 2) hallucinations, 3) depression, 4) anxiety, 5) agitation, 6) euphoria, 7) disinhibition, 8) irritability, 9) apathy, 10) aberrant motor behaviour, 11) night-time behaviour disturbances, and 12) appetite/eating abnormalities. The severity and frequency of each symptom are scored on the basis of semi-structured questions administered to the patient's caregiver. The continuous score for each symptom is obtained by multiplying severity (1-3) by frequency (1-4). The higher the score, the more problems the patient has. The summed symptom scores give the total NPI score. A symptom was defined as clinically relevant when the score of the individual NPI-item (severity*frequency) was ≥4.

Procedure

For this cross-sectional study, an old age psychiatrist administered the MMSE and the GDS, and a trained assistant administered the CAMCOG, the IADL and the NPI. The AMPS was administered by a certified occupational therapist who was blind to the scores on other measures. Each patient performed two familiar gender related tasks. Although the AMPS was administered in the kitchen of the occupational therapy unit, patients could choose non-kitchen tasks, such as ironing or watering plants. The caregiver was present during all assessments. All tests were administered within two weeks.

Statistical analyses

Descriptive statistics of MMSE, CAMCOG, IADL, GDS and AMPS motor and process scores were calculated. Pearson's correlation coefficients were computed to examine the relationships between the different scales (MMSE, CAMCOG, IADL, GDS, MMSE & IADL, CAMCOG & IADL, and GDS & IADL) and the AMPS process score. Correlation coefficients between .30 and .49 were regarded as low, between .50 and .69 as moderate, and between .70 and .89 as high [27]. Stepwise regression analyses were performed to examine the contribution of the MMSE, CAMCOG, IADL, GDS, NPI, and NPI items separately to the AMPS process score, as outcome measure for interference with daily life. Then stepwise regression analyses between the MMSE and the IADL, the CAMCOG and IADL, and the GDS and the IADL were performed to examine the additional effect of the combination of cognition (MMSE, CAMCOG, and GDS) and activity level (IADL) on daily life functioning as determined with the AMPS. Above mentioned analyses were calculated for the total group.

Pearson's correlation coefficients were also computed to examine the relationships between the different scales (MMSE, CAMCOG, IADL) and the AMPS process score for each diagnostic category with n ≥ 20 (AD, VaD/mixed type), and for different levels of cognitive severity. On the basis of the median MMSE score in the dementia group (score of 19) the demented group was divided in moderate-severe dementia and mild dementia. Independent Sample T-tests were carried out to see if there were significant differences between the mean scores of the MMSE, CAMCOG, IADL and AMPS motor and process scores according to diagnosis and severity. Stepwise regression analyses were performed for each diagnostic category and for each level of severity as described above for the total group. Statistical analyses were performed with the Statistical Package for Social Sciences, version 11.5. The alpha level was set at 0.05.

Results

Out of the 137 patients from the MEDICIE baseline dataset, 118 patients fulfilled the inclusion and exclusion criteria of this study. Seventy-nine patients were diagnosed as demented, of whom 54 had Alzheimer's disease [28], 19 vascular dementia or mixed type dementia [29], one had dementia due to multiple sclerosis and six had dementia not otherwise specified. Thirty-one patients had cognitive impairments without dementia (CIND), 6 suffered from primary depression or anxiety, and 1 had no diagnosis. Table 1 shows the patients' characteristics and means scores on all measures.

Table 2 shows the correlations, and unique and combined explained variance of the MMSE, CAMCOG, IADL, GDS, MMSE & IADL, CAMCOG & IADL, and GDS & IADL on the AMPS process scores. All correlations were significant and moderate. Although a significant relationship was found between the AMPS process scores and the total NPI score, the latter did not contribute significantly to the variance in the AMPS process score. Apathy was the only NPI item that, in combination with the MMSE scores and the CAMCOG scores, contributed weakly, but significantly, to the variance in the AMPS score (4% and 3% respectively). The range of the explained variance for the AMPS process score in the total group is 27% (MMSE) to 44% (GDS).

The mean scores of all measures did not differ significantly between AD and VaD/ mixed type. In order to examine the effect of diagnostic subtypes of dementia, correlations were calculated for the AD and VaD/mix group separately (Table 3). Correlations between measures and the AMPS process scores were in general moderate, with an exception of the correlation between GDS and AMPS process score, which was high (-0.82). The range of explained variance for the AMPS process score in the AD group is 20% (CAMCOG) to 37% (IADL & MMSE). The range in the VaD/mix group is 25% (MMSE) to 66% (GDS). As expected, the mean scores of the MMSE, CAMCOG, IADL, GDS and AMPS process scores differed significantly between groups with regard to severity of cognitive deterioration (moderate-severe dementia: MMSE score < 19, mild dementia: MMSE score ≥ 19, and CIND) (MMSE: t = 13.1, $p < 0.05$; CAMCOG: t = 7.3, $p < 0.05$; IADL: t = -2.2, $p < 0.05$; GDS: t = -4.1, $p < 0.05$; AMPS process score: t = 3.0, $p < 0.05$). Correlations between these measures and the process scores of the AMPS were then calculated for these three levels of severity (Table 4). Correlations were in general low to moderate, with the exception of the correlation between GDS and AMPS for the mildly demented group, which was high (-0.75).

No differences between groups were found on the AMPS motor scores. No differences were found on the IADL score and the AMPS motor scores. The explained variance for the AMPS process score in the severe dementia group is 23% (IADL). The range in the mild dementia group is 18% (MMSE) to 55% (GDS). The range in the CIND group is 21% (MMSE) to 27% (CAMCOG).

Discussion

We examined the interrelationship between commonly used instruments for measuring different aspects of dementia severity and the process scores of the AMPS, a measure based on direct observation of functioning, in a group of psychogeriatric patients. The main findings can be summarized as follows: cognitive measures (MMSE and CAMCOG) and the IADL were only moderately associated with the AMPS score, with an explained variance varying between 27% and 44%; the association between the cognitive measures and the AMPS process score was affected by the presence of apathy, but not of other neuropsychiatric symptoms; and this association was higher in patients with VaD, compared to AD, and in patients with mild dementia, compared to moderate-severe stages.

Table 1. Patient characteristics and mean scores on all measures

Diagnosis	Age Mean (SD) / range	Female N	Female %	MMSE Mean (SD) / range	CAMCOG Mean (SD) / range	IADL Mean (SD) / range	GDS Mean (SD) / range	AMPS process Mean (SD) / range	AMPS motor Mean (SD) / range	NPI total Mean (SD) / range
AD (n=54)	79.5 (6.8) 61-93	44	81.5	19.3 (4.9) 9-28	58.7 (15.5) 28-95	17.9 (5.9) 6-31	4.5 (0.9) 3-6	0.1 (0.9) -2.4-2.0	1.8 (1.0) -0.6-3.9	23.9 (16.7) 0-64
VaD/mix (n=19)	78.7 (4.1) 72-88	7	36.8	19.4 (5.4) 9-26	57.7 (19.9) 25-87	18.3 (5.0) 9-28	4.8 (0.9) 3-6	-0.2 (0.7) -1.3-1.4	1.4 (0.9) -0.4-2.9	23.0 (15.5) 3-58
Dem nos (n=6)	77.3 (5.4) 70-86	5	83.3	20.7 (3.9) 15-26	66.7 (8.3) 57-79	16.8 (8.5) 7-26	4.5 (1.0) 3-6	0.1 (0.6) -0.6-0.8	1.9 (1.1) 0.6-3.7	32.3 (21.3) 8-63
CIND (n=31)	75.5 (7.4) 55-87	17	54.8	24.8 (3.0) 18-29	82.1 (12.6) 47-102	14.4 (5.9) 6-25	3.4 (1.1) 1-5	1.0 (0.7) -0.2-3.0	2.1 (0.9) 0.0-3.9	20.3 (15.8) 0-55
Dep/anx (n=6)	79.2 (8.2) 65-85	2	33.3	22.2 (3.8) 18-28	75.5 (6.5) 68-86	16.8 (6.4) 9-28	3.8 (1.5) 2-6	0.7 (1.7) -1.6-2.7	1.5 (1.0) 0.4-2.7	28.2 (16.1) 7-44
Other (n=2)	72.5 (3.5) 70-75	1	50	23.5 (6.4) 19-28	70.0 (19.8) 56-84	16.0 (7.1) 11-21	3.0 (0.0) 3-3	0.5 (0.5) 0.1-0.8	1.2 (0.4) 0.9-1.4	15.5 (10.6) 8-23
Total group (n=118)	78.1 (6.7) 55-93	76	64.4	21.1 (5.0) 9-29	66.1 (18.1) 25-102	16.9 (6.0) 6-31	4.2 (1.1) 1-6	0.3 (0.9) -2.4-3.0	1.8 (1.0) -0.6-3.9	23.3 (16.4) 0-64

AD: Alzheimer's disease [28]; VaD/mix: Vascular dementia or dementia of the mixed type [29]; Dem nos: dementia not otherwise specified; CIND: Cognitive impairment, no dementia; Dep/anx: depression or anxiety; MMSE: Mini-Mental State Examination; CAMCOG: Cognitive part of the revised Cambridge Examination for Mental Disorders of the Elderly; IADL: Instrumental Acitivities of Daily living; GDS: Global Deterioration Scale; AMPS: Assessment of Motor and Process Skills; NPI: Neuropsychiatric Inventory.

Table 2. Correlation and stepwise regression to show the relationship between measures of dementia severity and observation as measured with the AMPS process score for the total group (N=118)

	Pearson R	R square	df (regression| residual)	F	B	t	p
MMSE	.54	.29	1|116	47.2	0.1	6.9	.000
CAMCOG	.58	.34	1|113	57.0	3.0^{-02}	7.6	.000
IADL	.52	.27	1|115	42.2	-7.9^{-02}	-6.5	.000
GDS	.66	.43	1|116	88.6	-0.6	-9.4	.000
NPI total	.19	.04	1|116	4.3	-1.1^{-02}	-2.1	.040
MMSE & IADL	.63	.40	2|114	37.6	-5.5^{-02}	-4.6	.000
CAMCOG & IADL	.64	.42	2|111	39.3	-5.1^{-02}	-4.1	.000
GDS & IADL	.67	.44	2|114	45.3	-2.9^{-02}	-2.2	.032
MMSE & NPI apathy	.57	.33	2|115	28.3	-5.2^{-02}	-2.6	.009
MMSE & NPI depression						-0.2	.874
CAMCOG & NPI apathy	.60	.36	2|112	31.5	-4.1^{-02}	-2.1	.041
CAMCOG & NPI depression						-0.1	.899

AMPS: Assessment of Motor and Process Skills; MMSE: Mini-Mental State Examination; CAMCOG: Cognitive part of the revised Cambridge Examination for Mental Disorders of the Elderly; IADL: Instrumental Activities of Daily living; GDS: Global Deterioration Scale; NPI: Neuropsychiatric Inventory.

Although the MMSE is widely used as a measure of cognitive status in dementia, we found that it has limited relevance predicting functional ability. The CAMCOG, although measuring cognition in more detail than the MMSE, did not provide more explained variance. Our findings correspond with those of Robinson & Fisher [20], who examined the relationship between the AMPS and tests of cognition - MMSE & CAMCOG - in people with dementia or mild memory impairment. They found that cognitive functioning accounted for 40% of the variance in functional ability. Therefore, a substantial part of the variance remains unexplained. This raises the question to what extent cognitive outcome measures may adequately reflect severity of dementia. According to clinical convention, measures such as the MMSE are taken traditionally as the most relevant to assess dementia severity. While this type of measure has been shown to correlate considerably with the amount of neuropathology, notably the number of neocortical neurofibrillary tangles, in patients with AD [30] and therefore may be regarded as valid measures of severity of the underlying disease, they are at best moderate predictors for daily life functioning. Deriving patients' functional status on the basis of a MMSE score should therefore take place cautiously, and attention

Table 3. Correlations between the severity measures and the AMPS process score for each diagnosis

	MMSE	CAMCOG	IADL	GDS
AD (n=54)	.47**	.44**	-.53**	-.56**
VaD/mix (n=19)	.52*	.56*	-.62**	-.82**

*p < 0.05; **p < 0.01
AD: Alzheimer's disease [28]; VaD/mix: Vascular dementia or dementia of the mixed type [29]; MMSE: Mini-Mental State Examination; CAMCOG: Cognitive part of the revised Cambridge Examination for Mental Disorders of the Elderly; IADL: Instrumental Activities of Daily living; GDS: Global Deterioration Scale; AMPS: Assessment of Motor and Process Skills.

Table 4. Correlations between the severity measures and the AMPS process score for each level of severity

	MMSE	CAMCOG	IADL	GDS
CIND (n=31)	.44*	.52**	-.16	-.25
Dementia, moderate/severe (n=34)[a]	.15	.19	-.48**	-.15
Dementia, mild (n=45)[b]	.45**	.42**	-.55**	-.75**

*p < 0.05; **p < 0.01
[a] Dementia and MMSE < 19 is moderate/severe dementia; [b] Dementia and MMSE ≥ 19 is mild dementia; CIND: Cognitive impairment, no dementia; MMSE: Mini-Mental State Examination; CAMCOG: Cognitive part of the revised Cambridge Examination for Mental Disorders of the Elderly; IADL: Instrumental Activities of Daily living; GDS: Global Deterioration Scale; AMPS: Assessment of Motor and Process Skills; NPI: Neuropsychiatric Inventory.

should be paid to other factors affecting the relation between cognitive impairment and daily functioning, such as emotional functioning, premorbid functioning, environmental demands [9], cognitive profile and neuropsychiatric symptoms. Our finding that the presence of apathy affects this relationship underlines this view.

The moderate correlation between the IADL and the AMPS process score is in line with Doble et al. [31], who found a discordance of 46% between family informant ratings and the AMPS process score for assessing IADL functioning. This figure is lower than we originally would have expected, since both measures (AMPS and IADL) reflect daily life functioning. An explanation for this is that the IADL is a derivative of daily life functioning while the AMPS measures actual functioning. Observation is more ecologically valid than self- or informant-based questionnaires, and tasks measuring performance give a closer view of what a patient actually does – performance – instead of what a patient can do – competence [32]. Moreover, the IADL is informant based and might therefore be susceptible to caregiver bias [33].

In our study, the GDS showed higher correlations with the AMPS process scores than the cognitive outcomes. This may be explained by the more global nature of the GDS, which may be more relevant for the prediction of functioning than instruments focusing on one specific aspect of dementia (e.g. cognition). These findings are in line with those of Paul et al. [34], who found that the GDS was significantly related to ratings of instrumental activities of daily living.

In general, the relationship between AMPS process scores and the other measures were higher in the VaD group. This is in contrast to what we would have expected, as this type of dementia is characterized by a larger clinical heterogeneity due to more variation in localization of pathology in the brain compared to AD. It may be hypothesized that ischemic lesions in vascular demented patients also cause motor disturbances which also directly affect daily functioning, leading to a closer relationship between cognition and functioning in this particular group. As the VaD group was relatively small (n = 19), these results should be interpreted cautiously awaiting replication.

An advantage of our study is the standardized approach by which it was adopted, and the fact that the AMPS was administered independently of the other assessments. Second, we used data of a heterogeneous group of psychogeriatric patients, which may improve the generalizability of our findings. A limitation of our study is that although a professional observer scored the patient's performance, the activities were all performed in the kitchen of the university hospital. This could have influenced performance. Furthermore, the level of difficulty of the tasks was slightly higher in the female group.

Administering the AMPS may be problematic because it is more time-consuming compared to the traditional measures, and it requires specific training: occupational therapists are only allowed to use the instrument when adequately trained and certified. Therefore its use in routine diagnostic procedures may not be realistic. Our data underline, however, the notion that the outcome measures traditionally used for the assessment of severity of dementia reflect specific domains that interrelate in an unclear fashion. The combination of these domains may help health professionals provide appropriate care to improve functioning. Our study shows that direct observation of performance of activities assessed with the AMPS can provide additional relevant information about patients' functional ability problems, and may be helpful for a better understanding of these problems. In a study of recovery of specific cognitive functions and performance of daily life activities (determined with the AMPS) in severe brain injured patients, Linden et al. [15] concluded that the AMPS provides insight into a patient's rehabilitation different from that provided by cognitive tests and may be a better indicator of a patient's ability to resume independent living. Given the increased attention of occupational therapist interventions for people with dementia [21], we feel that the AMPS can also be relevant for this population.

Our study adds to the relatively unexplored relationship between domains of dementia. It is remarkable that there is relatively little theoretical background on this issue. Future research should focus on more insight in the associations between different constructs around dementia severity and daily life functioning.

References

1. American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision. Washington, DC, American Psychiatric Association 2000.
2. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975;12(3): 189-198.
3. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. Gerontologist 1969;9(3):179-186.
4. Reisberg B, Ferris SH, de Leon MJ, Crook T. The Global Deterioration Scale for assessment of primary degenerative dementia. Am J Psychiatry 1982;139(9):1136-1139.
5. Hughes CP, Berg L, Danziger WL, Coben LA, Martin RL. A new clinical scale for the staging of dementia. Br J Psychiatry 1982;140:566-572.
6. National Institute for Clinical Excellence. Guidance on the use of Donepezil, Rivastigmine and Galantamine for the treatment of Alzheimer's disease. London: National Institute for Clinical Excellence (NICE);2001.
7. Nygard L, Bernspang B, Fisher AG, Winblad B. Comparing motor and process ability of persons with suspected dementia in home and clinic settings. Am J Occup Ther 1994;48(8):689-696.
8. Fisher A. Assessment of motor and process skills. 3th edition. Fort Collins, CO: Three star press 1999.
9. Chaytor N, Schmitter-Edgecombe M. The ecological validity of neuropsychological tests: a review of the literature on everyday cognitive skills. Neuropsychol Rev 2003; 13(4):181-197.
10. Bootes K, Chapparo CJ. Cognitive and behavioural assessment of people with traumatic brain injury in the work place: occupational therapists' perceptions. Work 2002;19(3):255-268.
11. Kaitaro T, Koskinen S, Kaipio ML. Neuropsychological problems in everyday life: a 5-year follow-up study of young severely closed-head-injured patients. Brain Inj 1995; 9(7):713-727.
12. Bernspang B, Fisher AG. Differences between persons with right or left cerebral vascular accident on the Assessment of Motor and Process Skills. Arch Phys Med Rehabil 1995; 76(12):1144-1151.
13. Darragh AR, Sample PL, Fisher AG. Environment effect of functional task performance in adults with acquired brain injuries: use of the assessment of motor and process skills. Arch Phys Med Rehabil 1998;79(4):418-423.
14. Johansson U, Bernspang B. Predicting return to work after brain injury using occupational therapy assessments. Disabil Rehabil 2001;23(11):474-480.
15. Linden A, Boschian K, Eker C, Schalen W, Nordstrom CH. Assessment of motor and process skills reflects brain-injured patients' ability to resume independent living better than neuropsychological tests. Acta Neurol Scand 2005;111(1):48-53.

16. McNulty MC, Fisher AG. Validity of using the Assessment of Motor and Process Skills to estimate overall home safety in persons with psychiatric conditions. Am J Occup Ther 2001;55(6):649-655.

17. Baron KB. Clinical interpretation of "the Assessment of Motor and Process Skills of persons with psychiatric disorders". Am J Occup Ther 1994;48(9):781-782.

18. Doble SE, Fisk JD, MacPherson KM, Fisher AG, Rockwood K. Measuring functional competence in older persons with Alzheimer's disease. Int Psychogeriatr 1997;9(1):25-38.

19. Oakley F, Sunderland T. Assessment of motor and process skills as a measure of IADL functioning in pharmacologic studies of people with Alzheimer's disease: a pilot study. Int Psychogeriatr 1997;9(2):197-206.

20. Robinson SA, Fisher AG. A study to examine the relationship of the assessment of motor and process skills (AMPS) to other tests of cognition and function. British Journal of Occupational Therapy 1996;59(6):260-263.

21. Graff MJL, Vernooij-Dassen MJM, Thijssen M, Dekker J, Hoefnagels WHL, Olde Rikkert MGM. Community based occupational therapy for patients with dementia and their care givers: randomised controlled trial. BMJ, doi: 10.1136/bmj.39001.688843. BE (published 17 november 2006).

22. Wolfs CAG, Kessels A, Dirksen CD, Severens JL, Verhey FRJ. Integrated multidisciplinary diagnostic approach for dementia care: randomised controlled trial. 2008; Br J Psychiatry:192(4):300-5.

23. Kok RM, Verhey FRJ, Schmand B. Meetinstrumenten bij cognitieve stoornissen. Tijdschr Psychiatr 2004;46(10):665-669.

24. Verhey FR, Huppert FA, Korten EC, Houx P, de Vugt M, van Lang N, et al. Cross-national comparisons of the Cambridge Cognitive Examination-revised: the CAMCOG-R: results from the European Harmonization Project for Instruments in Dementia. Age Ageing 2003;32(5):534-540.

25. Cummings JL. The Neuropsychiatric Inventory: assessing psychopathology in dementia patients. Neurology 1997;48(5 Suppl 6):S10-16.

26. Cummings JL, Mega M, Gray K, Rosenberg-Thompson S, Carusi DA, Gornbein J. The Neuropsychiatric Inventory: comprehensive assessment of psychopathology in dementia. Neurology 1994;44(12):2308-2314.

27. Hinkle DE, Wiersma W, Jurs SG. Applied statistics for the behavioral sciences. New York: Houghton Mifflin Company 1998.

28. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. Neurology 1984;34(7):939-944.

29. Roman GC, Tatemichi TK, Erkinjuntti T, Cummings JL, Masdeu JC, Garcia JH, et al. Vascular dementia: diagnostic criteria for research studies. Report of the NINDS-AIREN International Workshop. Neurology 1993;43(2):250-260.

30. Giannakopoulos P, Herrmann FR, Bussiere T, Bouras C, Kovari E, Perl DP, et al. Tangle and neuron numbers, but not amyloid load, predict cognitive status in Alzheimer's disease. Neurology 2003;60(9):1495-1500.

31. Doble SE, Fisk JD, Rockwood K. Assessing the ADL functioning of persons with Alzheimer's disease: comparison of family informants' ratings and performance-based assessment findings. Int Psychogeriatr 1999;11(4):399-409.

32. Doble SE, Dalhousie U. Measuring competence in the performance of activities of daily living: the forgotten half of Alzheimer's disease. Dissertation Abstracts International Section B The Sciences and Engineering 2001;62(2-B):762.

33. Nygard L, Winblad B. Measuring long term effects and changes in the daily activities of people with dementia. J Nutr Health Aging 2006;10(2):137-138.

34. Paul RH, Cohen RA, Moser DJ, Zawacki T, Ott BR, Gordon N, et al. The global deterioration scale: relationships to neuropsychological performance and activities of daily living in patients with vascular dementia. J Geriatr Psychiatry Neurol 2002; 15(1):50-54.

Part II

Ecological validity of neuropsychological
assessment and the integration of domains

Chapter 4

Aspects of the ecological validity of
neuropsychological tests: Introduction to Part II

Abstract

Neuropsychologists are frequently asked to make predictions about the daily life functioning of people with brain dysfunction largely on the basis of neuropsychological tests. The assumption is that these tests are ecologically valid. The tenability of this assumption, however, is unclear. This chapter provides a description of different aspects of the ecological validity of neuropsychological tests. Research suggests that the association between neuropsychological tests and daily life functioning is moderate. This implies that the daily life functioning of people with brain dysfunction cannot be explained solely on the basis of neuropsychological tests. Several factors (e.g. mood, testing environment) may influence the association between neuropsychological tests and daily life functioning. Direct observation of patients in their own environment may provide useful additional information that can be used in combination with neuropsychological test data to plan adequate treatment and future care for patients.

Defining the problem

The association between cognition and daily life functioning in patients with different types of dementia as well as in psychogeriatric patients was described in Chapter 2 and Chapter 3. This chapter focuses on aspects related to the ecological validity of neuropsychological tests.

People with brain dysfunction often experience problems in various areas of functioning. A neuropsychological assessment is concerned with identifying the cognitive, emotional, psychosocial, and behavioural consequences of brain dysfunction associated with acquired brain injury or dementia [1, 2]. One of the aims of the neuropsychological assessment is to predict the daily life functioning of people with brain dysfunction [3]. Neuropsychological tests are used to identify the specific cognitive disorders. This use is based on the assumption that neuropsychological tests reflect the cognitive problems met in daily life; in other words, neuropsychological tests are supposed to be ecologically valid.

Ecological validity is defined as the degree to which the behaviours observed and recorded in a study reflect the behaviours that actually occur. Ecological validity is often confused with external validity (which refers to the generalizability of a study's findings). While these forms of validity are closely related, they are independent; a study may possess external validity but not ecological validity, and vice-versa [4, 5]. Ecological validity issues will become more relevant due to the growing availability of neuroimaging techniques, as a result of which neuropsychologists nowadays are more frequently asked to predict how patients will function in their daily life, than to diagnose and localize brain impairments [3]. Hence, the role of the neuropsychologist has changed to some extent.

An improved understanding of the neuropsychological factors associated with impaired daily life functioning will support the early identification of patients' needs and may provide useful information for treatment and future care [6]. High ecological validity of neuropsychological tests is important since the predictions based on these tests may have significant consequences for the lives of patients and their caregivers [7]. Furthermore, for patients and their environment, information about the consequences of brain dysfunction for daily life is of more importance than the measurement of performance in artificial testing situations.

The ecological validity of neuropsychological tests received little attention until the 1980s. Heaton & Pendleton [8] were among the first to point to the need for more information about the possible impact of cognitive deficits on daily life activities and to the role that could be played in this by neuropsychological testing, which could provide meaningful information about the problems that people with cognitive disorders face in their everyday functioning. In 1996, Sbordone & Long [9] edited an elaborate textbook on findings concerned with the ecological validity of neuropsychological testing.

Chaytor & Schmitter-Edgecombe [7] wrote an extensive review of the literature addressing the ecological validity of neuropsychological testing. They described the extent to which a range of specific neuropsychological tests was related to the everyday cognitive functioning of people with cognitive disorders. When investigating the ecological validity of neuropsychological tests, the specific underlying cognitive deficits in daily life need to be identified. Therefore, the outcome measure with which the test is compared is important.

Little is known about the influence of the outcome measure in the investigation of the ecological validity of neuropsychological tests.

The aim of this chapter is to provide a description of the relevant aspects of the ecological validity of neuropsychological tests.

Definition of ecological validity

Different definitions exist to describe ecological validity. Table 1 provides an overview of commonly used definitions. Ecological validity does not apply to the test itself, but to the inferences that are drawn from the test [10, 11]. This means that a test can have diagnostic validity but not necessarily adequate ecological validity [7].

The construct of ecological validity

In their review, Chaytor & Schmitter-Edgecombe [7] summarized different constructs of ecological validity as proposed by others. As this summary clearly points to the different ways of conceptualizing the ecological validity of neuropsychological testing, their summary is presented below.

Chaytor & Schmitter-Edgecombe [7] refer to Franzen & Wilhelm [12] who introduced two ways to address the ecological validity of neuropsychological tests: verisimilitude and veridicality. Verisimilitude refers to the degree to which the cognitive requirements of a test theoretically resemble the cognitive requirements in the everyday world. Verisimilitude abandons traditional tests and develops new tests with ecological goals in mind. These tests have more face validity than traditional tests since they attempt to simulate cognitive daily life activities. Tests with verisimilitude do not make a distinction between cognitively impaired and non-impaired people, but catch the essence of cognitive skills necessary for daily life functioning. The primary goal of such tests is to identify people who have difficulty performing everyday tasks, regardless of brain pathology. Therefore, it is possible that these tests fail to detect brain damage when the patient is able to perform daily life activities [7]. Examples of tests with verisimilitude are the Behavioural Assessment of the Dysexecutive Syndrome (BADS) [13] and the Test of Everyday Attention (TEA) [14], and the Rivermead Behavioural Memory test (RBMT) [15].

In contrast, veridicality refers to the degree to which traditional neuropsychological tests are empirically related to measures of everyday functioning. Statistical techniques are used to relate performance on traditional neuropsychological tests to measures of everyday functioning. It is possible that traditional tests do accurately predict daily life cognitive abilities even though they were not developed to generate verisimilitude [7].

Next to verisimilitude and veridicality, a third way to establish the ecological validity of neuropsychological tests is to relate test scores in a given cognitive domain to scores on measures of everyday cognition within the same domain across situations [16]. An example of this is that tests for selective attention would be related to measures of everyday selective attention abilities across situations. Determination of the cognitive abilities required for daily life functioning is necessary for this method to work.

Table 1. Definitions of ecological validity

Reference	General definitions
Brunswik [48]	The degree to which one could generalize from results of controlled, systematic experiments to naturally occurring events in the real world
Tupper & Cicerone [49]	The degree to which results obtained in controlled experimental conditions are related to those obtained in naturalistic environments
Specific definitions of ecological validity of neuropsychological tests	
Lezak [50]	Refers to how well the neuropsychological assessment data predict future behaviour or behavioural outcomes
Ginsberg et al. [51]	The functional and predictive relationship between a person's performance on a set of neuropsychological tests and his or her behaviour in a variety of real world settings, such as home, work, school or the community
Franzen & Arnett [10]	The degree to which test performance corresponds to real world performance
Higginson et al. [52]	The degree to which clinical tests of cognitive functioning predict functional impairment
Ready et al. [53]	Ecological validity is indicated when associations are found between test scores and behaviours that occur external to the psychological evaluation in the real world
Nadolne & Stringer [54]	Within the assessment domain, ecological validity can refer to the extent to which a test includes materials drawn from the everyday environment, the extent to which performance in a clinical setting resembles performance in a naturalistic setting, or the ability of a test to predict performance on some criterion of everyday functioning
Loring et al. [55]	Ecological validity typically refers to how well the neuropsychological assessment data, for cognitive rehabilitation, predict future behavior or behavioral outcomes
Kvavilashvili & Ellis [56]	The degree of 'representativeness' of a task (the extent to which a clinical test corresponds in form and context to a situation encountered outside the laboratory), and the 'generalisability' of test results (the degree to which poor performance on the test will be predictive of problems outside the laboratory)
Wood & Liossi [24]	Ecological validity allows test results to predict social and functional abilities relevant to real world settings

So, although originally a certain neuropsychological test was not developed with ecological validity in mind, it may yet prove to be ecologically valid (i.e. veridicality); or a neuropsychological test might be ecologically valid since the test was originally developed with ecological validity in mind (i.e. verisimilitude).

Investigating the ecological validity of neuropsychological tests

When investigating the ecological validity of neuropsychological tests, two main aspects have to be taken into account: 1) the neuropsychological test itself; 2) the outcome measure. A gold standard for measuring the ecological validity of neuropsychological tests is lacking. The nature of the outcome measures used to determine the ecological validity of neuropsychological tests can therefore differ widely, for example: the number of complaints reported by the patient or the caregiver; the level of difficulty with unplanned activities; the level of burden; or the quality of life.

When using daily life functioning as an outcome measure, the identification of the underlying cognitive deficits is important. When a person cannot find his watch, this may be due to a memory problem, but also to an attention problem. Besides the accurate identification of the deficits, it is also important to note that a low score on a cognitive test does not necessarily indicate problems in daily life. Hence, a person with aphasia might score badly on a verbal cognitive test, but might perform well on non-verbal everyday cognitive activities [17]. With regard to the outcome measure used, the association between a neuropsychological test and daily life functioning based on a self-report might differ from the association between a neuropsychological test and daily life functioning based on direct observation of a patient performing a daily life activity.

This chapter will focus on the ecological validity of neuropsychological tests with regard to daily life functioning, as neuropsychologists are frequently asked to predict daily life functioning on the basis of neurological tests [3]. The main aspects will now be discussed separately below.

Neuropsychological tests

Many studies have investigated the ecological validity of neuropsychological tests. Table 2 shows an overview of the associations found between cognition and daily life functioning. Although some studies found high associations between cognition and daily life functioning [18-20], most studies found that the ecological validity of neuropsychological tests was low [21-25] to moderate [26-29].

In general, the strongest associations between neuropsychological tests and daily life functioning are noted when the outcome measure corresponded to the cognitive domain assessed by the neuropsychological tests [7]. An example of this is Johnson's Episodic Memory Scale (EMS) [30]. This involved comparing the EMS with the Functional Assessment Rating Scale (FRS) [31], an informant-based questionnaire which measures, among other things, everyday functional memory. However, some exceptions do exist: although executive functioning, which includes attention, strategy use and planning, is a cognitive domain of particular importance in daily life functioning [3], the current set of commonly used executive tests failed to significantly predict everyday executive functioning [21]. In this case, verisimilitude turned out to be somewhat superior in comparison to veridicality [7].

Table 2. Association between cognitive functioning and daily life functioning

Study	Cognitive measures	Measures for daily life functioning	Association
Low association			
Wood & Liossi [24]	Hayling A, B, C Test Brixton Test Zoo Map (BADS) Key Search (BADS)	Clinical interview to identify executive deficits in everyday life DEX	Only significant association was between Hayling C and DEX: r = -0.26, p < 0.05
Chaytor et al. [21]	TMT B COWAT SCWT WSCT	DEX BAFQ	Significant associations between r = 0.33 and r = 0.38, p < 0.05 and p < 0.01
Odhuba et al. [22]	Hayling Brixton	DEX ICHII CIQ	Significant associations range from r = 0.23 to r = 0.48 (p < 0.001, or p < 0.01, or p < 0.05)
Silverberg et al. [23]	BCET LNS (WAIS-III) TMT B Stroop JOLO GPT-D Recognition of CVLT	DRS	The only significant associations were between LNS and DRS (r = -0.31, p < 0.01), between TMT B and DRS (r = 0.42, p < 0.01), and between Stroop and DRS (r = 0.42, p < 0.01)
Zwecker et al. [25]	LOTCA MMSE FIM cognitive Subscale	FIM efficacy FIM efficiency MFRS	Significant associations range from r = 0.25 to r = 0.34 (p < 0.05, or p < 0.01)

Table 2. (continued)

Study	Cognitive measures	Measures for daily life functioning	Association
Moderate association			
Bouwens et al. [26]	MMSE CAMCOG	AMPS process score	r = 0.54, p < 0.001 r = 0.58, p < 0.001
Doble et al. [27]	MMSE	AMPS process score	r = 0.64, p < 0.002
Gillen & Gernert-Dott [28]	TMT WMS-R BNT	ILS	r ranges from r = -0.12 (p = 0.649) to r = 0.64 (p = 0.02)
Kizony & Katz [29]	COGNISTAT LOTCA CMT RCF SC	AMPS process score	Significant associations range from r = 0.34 to r = 0.59 (p < 0.05 or p < 0.01)
High association			
LaBuda & Lichtenberg [18]	MDRS	SAILS	r = 0.75, p < 0.05
Landi et al. 2002 [19]	CPS of MDS-PAC	IADL of MDS-PAC	Demented people appeared to have 64% decreased probability of recovery after a rehabilitation programme relative to patients with normal cognitive performance. OR = 0.36 (95% CI = 0.14-0.92)

Table 2. (continued)

Study	Cognitive measures	Measures for daily life functioning	Association
High association			
Sands et al. [20]	SPMSQ	BADL IADL Mobility	Patients with impaired cognitive performance were more likely to be admitted to a nursing home for the first time by 90 days after discharge. OR = 2.8 (95% CI = 1.8-4.5) for patients with mildly impaired cognitive performance and OR = 6.7 (95% CI = 4.5-9.8) for patients with moderate to severely impaired cognitive performance.

AMPS: Assessment of Motor and Process Skills [36]; BADL: basic activities of daily living; BADS: Behavioural Assessment of the Dysexecutive Syndrome [13]; BAFQ: Brock Adaptive Functioning Questionnaire [57]; BCET: Biber Cognitive Estimation Test [58]; BNT: Boston Naming Test [59]; CAMCOG: cognitive component (CAMCOG) of the revised Cambridge Examination for Mental Disorders of the Elderly (CAMDEX) [60]; CI: confidence interval; CIQ: Community Integration Questionnaire [61]; CMT: Contextual Memory Test [62]; COGNISTAT: Neurobehavioural Cognitive Status Examination [63]; COWAT: Controlled Oral Word Association Test [64]; CPS: Cognitive Performance Scale [65]; CVLT: California Verbal Learning Test [66]; DEX: Dysexecutive Questionnaire [13]; DRS: Disability Rating Scale [67]; FIM: Functional Independence Measure [68]; GPT-D: Grooved Pegboard Test, dominant hand [69]; IADL: Instrumental Activities of Daily Living; ICHII: Iowa Collateral Head Injury Interview [70]; ILS: Independent Living Scales [71]; JOLO: Judgement of Line Orientation [72]; LNS: Letter-Number Sequencing; LOTCA: Loewenstein Occupational Therapy Cognitive Assessment [73]; MDRS: Mattis Dementia Rating Scale [74]; MDS-PAC: Minimum Data Set for Post-Acute Care [65]; MRFS: Montebello Rehabilitation Factor Score [75]; MMSE: Mini-Mental State Examination [76]; OR: odds ratio; RCF: Rey Complex Figure [77]; SAILS: Structured assessment of independent living skills [78]; SC: Star Cancellation [79]; SCWT: Stroop Color and Word Test [80]; SPMSQ: Short Portable Mental Status Questionnaire [81]; TMT: Trail Making Test [82]; WIAS-III: Wechsler Adult Intelligence Scale, third edition [83]; WCST: Wisconsin Card Sorting Test [84]; WMS-R: Wechsler Memory Scale-R [85].

Outcome measures

The ecological validity of neuropsychological tests does not only depend on the neuropsychological test itself, but also on the outcome measure to which the test is compared. Recent studies have described different kinds of outcome measures e.g. self-reports, checklists, informant-based questionnaires, clinicians' ratings, IADL-scales, observations of simulated activities and questionnaires. Every outcome measure has some degree of error and is not always structured and objective [6, 7, 21]. Another problem with daily life functioning is that this is a broad concept with several activities. So far, an instrument that grasps all these aspects is lacking.

Informant-based ratings and clinicians' ratings revealed stronger associations between neuropsychological tests and daily life functioning than self-reports did [7]. Furthermore, tests had stronger relations to daily functioning when informant-based ratings were used that measured complex cognitive processes instead of specific skills. Rating-scales present problems because patients may lack insight into their problems [32], or because informants and clinicians can be somewhat insensitive when difficulties that were underestimated by patients later improve [33, 34].

Although self-reports and informant-based ratings of daily life functioning are practical and the most commonly used methods for the assessment of everyday functioning, these instruments are merely derivatives of daily life functioning. Self-based or informant-based questionnaires only tell us what a patient can do instead of what a patient actually does [35]. Bouwens et al. [26] showed that direct observation of the performance of activities assessed with the Assessment of Motor and Process Skills (AMPS) can provide additional relevant information about patients' functional ability problems, and may be helpful for a better understanding of these problems. Direct or performance-based assessments such as direct observation of simulated activities in naturalistic settings are, however, less frequently used as outcome measure for the ecological validity of neuropsychological tests in adults with cognitive disorders.

Simulated activities in clinical settings are still artificial and do not accurately reflect everyday functioning. A solution for the artificial aspects of simulated activities is direct observation of relevant and familiar activities in the patient's own environment. As described by Doble et al. [35] direct observation assesses what patients will actually do (i.e. performance) and not what patients can do (i.e. competence) since they are asked to perform relevant activities that they are used to performing in daily life. Occupational therapists frequently use direct observation instruments, such as the Assessment of Motor and Process Skills (AMPS) [36], or the Tailored Activity Program (TAP) [37] to assess whether patients can live independently in their own environment. However, no studies have been conducted to investigate the ecological validity of neuropsychological testing by observing the patients in their own environment.

Problems when investigating the ecological validity of neuropsychological tests

Recent studies describe several factors that have an influence on the association between neuropsychological tests and daily life functioning [6, 7, 21, 38-44]. These factors make it difficult to investigate the ecological validity of neuropsychological tests. Table 3 shows

an overview of the influencing factors described in the literature. As this table makes clear, the ecological validity of neuropsychological tests can be influenced by personal and environmental factors, as well as by factors that are due to the theoretical constructs of the test itself, by practical aspects of the test situation and by differences between test situations and everyday situations.

Stuss et al. [45] pointed to the influence of the artificial testing situation. In a test situation the therapist becomes the frontal lobes of the patient. This means that the therapist decides instead of the patient whether or not a task needs to be done, or selects a task from a number of competing possibilities. Moreover, affective arousal is carefully controlled for and, thus, an important variable for decision-making and behavioural control in everyday life is effectively omitted from the assessment process.

It is likely that different populations have difficulties with the same test but for different reasons [6, 7, 21]. Chaytor & Schmitter-Edgecombe [7] stress three problems with regard to differences between populations: 1) heterogeneity of the study populations makes it difficult to draw definite conclusions about the ecological validity; 2) people without brain dysfunction may have different ways of performing tasks, which rules out the possibility to make generalizations about the cognitive skills required for a task; 3) when the environment requires higher cognitive skills that are not assessed by the neuropsychological test, and the patient has problems with these skills, these could not be predicted from the test scores.

Although many ecologically oriented instruments (verisimilitude) tend to be more face valid than traditional tests (veridicality), only a subset have been validated empirically in terms of their ability to capture relevant everyday cognitive skills [43].

Clinical implications

Ecologically oriented instruments are of particular relevance for neuropsychologists who consider rehabilitation-related referral questions [43]. These instruments have less relevance when the referral question is concerned with diagnosing cognitive deficits, since they may fail to detect brain damage if the person is able to perform everyday tasks [7]. As long as only a subset of ecologically oriented instruments (verisimilitude) has been validated empirically, these instruments might be used complementary to traditional non-ecologically valid neuropsychological tests and not as a substitution of those tests.

When the purpose of the neuropsychological assessment is to predict everyday cognitive abilities, there is a need for testing hypotheses in the everyday situation to identify the underlying deficits of problematic daily life functioning. The strength of the results depends on these hypotheses.

Most neuropsychological tests are not developed for making predictions about patients' everyday cognitive abilities, but when they are used for that purpose, they cover only a small part of functions needed in daily life. Bouwens et al. [26] found that only 30% of the daily life functioning of psychogeriatric patients with cognitive disorders could be explained by considering their cognitive status. Direct observation of patients performing relevant daily life activities in their own environment is therefore recommended as a complementary tool to neuropsychological testing. Direct observation has the following advantages: it provides relevant and objective information about everyday functioning [41]; it allows patients to

Table 3. Factors influencing the ecological validity of neuropsychological tests

Personal factors
Emotion
Mood
Behaviour
Motivation
Personality
Physical/motor functioning
Health problems
Premorbid functioning
Educational level/IQ
Medical diagnosis
Illness severity
Time since injury
Environmental factors
Nature of testing environment
Family support
Environmental demands
Other factors
Incomplete agreement on what construct the test measures
Small sample of behaviour observed during testing
Inability to use compensatory strategies during test situation
Person completing the outcome measure

use compensatory strategies which they are not allowed to use during neuropsychological testing; it allows proper testing of executive functions since patients are not told when to begin, what to do exactly and how to be successful; and it allows the testing of hypotheses as well as of the effect of corrective feedback. The latter is a helpful way to estimate whether patients will benefit from rehabilitation [46].

In those cases when direct observation is not possible, ecological validity can be improved by using instruments designed with verisimilitude in mind. These tests are both standardized and ecologically relevant. When ecologically oriented tests are not available, the use of a broad range of sources will help neuropsychologists to draw accurate conclusions. Examples of these sources are informant-based questionnaires or IADL-rating scales. Moreover, the collection of qualitative data through carefully planned adaptations of the testing environment in order to assess discrepancies between test performance in optimal conditions and the performance of daily life activities in more natural conditions, also provides relevant data for the planning of treatment [7, 39, 47]. Neuropsychologists' handling of referrals concerned with everyday cognitive abilities will be even more ecologically valid if they take these influencing factors into account.

Recommendations for future research

As neuropsychologists are increasingly asked to make referrals on the basis of their predictions of patients' everyday cognitive abilities [3], more attention should be paid to the ecological validity of the tests they base their decisions on. Traditional tests could be improved or new instruments should be developed with ecological validity in mind (verisimilitude). Besides the improvement or development of tests, neuropsychologists should be encouraged to actually use ecologically valid measures [43]. Moreover, research is needed to investigate the factors that influence the association between neuropsychological tests and daily life functioning.

Direct observation of patients performing relevant and important instrumental activities of daily life in their own environment is recommended as an outcome measure for future research on the ecological validity of neuropsychological tests. In contrast to most of the other outcome measures available (e.g. informant-based questionnaires), direct observation is not a derivative of daily life. This will result in an important improvement of the ongoing studies on ecological validity and improve the ecological validity of neuropsychological tests. A problem with direct observation is, however, its standardization: direct observation is difficult to quantify and the relevance of daily life activities differs between patients.

Conclusions

Recent studies investigating the ecological validity of neuropsychological testing found a moderate association between neuropsychological tests and daily life functioning of people with cognitive disorders. This moderate association implies that the daily life functioning of people with brain dysfunction cannot be predicted solely on the basis of neuropsychological tests. All the outcome measures used to assess everyday functioning have their drawbacks and several factors may influence the association, making it difficult to investigate its ecological validity. The factors that may potentially harm the association between neuropsychological tests and daily life functioning should be taken into account. In view of the changing role of neuropsychologists as they are increasingly asked to predict the everyday consequences of brain dysfunction, it is clear that ecological validity is a crucial issue. It is therefore important to recognise that using supporting evidence from daily life situations can make the process of predicting daily life abilities easier and more valid. When neuropsychologists are confronted with rehabilitation-related referral questions, direct observation of patients in their own environment may provide useful information for the planning of treatment and future care as well as for the accurate prediction of daily life functioning.

Chapters 5 and Chapter 6 of this thesis will focus on further investigating the ecological validity of neuropsychological tests and the use of direct observation as an outcome measure for daily life functioning. The following research questions are to be answered in these chapters:

1. To what extent can neuropsychologists predict the daily life functioning of people with acquired brain injury solely on the basis of neuropsychological data? (Chapter 5)

2. Does the integration of a neuropsychological assessment with direct observation of the daily life functioning of people suffering from acquired brain injury have added value for treatment and future care? (Chapter 6)

References

1. Evans JJ. Basic concepts and principles of neuropsychological assessment. In: Halligan PW, Kischka U, Marshall JC, editors. Handbook of clinical neuropsychology. Oxford: Oxford University Press 2003:15-47.
2. Lezak MD. The neuropsychological examination: interpretation. In: Lezak MD, editor. Neuropsychological assessment. New York: Oxford University Press 1995.
3. Kibby MY, Schmitter-Edgecombe M, Long CJ. Ecological validity of neuropsychological tests: focus on the California Verbal Learning Test and the Wisconsin Card Sorting Test. Arch Clin Neuropsychol 1998;13(6):523-534.
4. Brewer M. Research Design and Issues of Validity. In: Reis H, Judd C, editors. Handbook of Research Methods in Social and Personality Psychology. Cambridge: Cambridge University Press 2000.
5. Shadish W, Cook T, Campbell D. Experimental and quasi-experimental designs for generalized causal inference. Boston: Houghton Mifflin 2002.
6. Farias ST, Harrell E, Neumann C, Houtz A. The relationship between neuropsychological performance and daily functioning in individuals with Alzheimer's disease: ecological validity of neuropsychological tests. Arch Clin Neuropsychol 2003;18(6):655-672.
7. Chaytor N, Schmitter-Edgecombe M. The ecological validity of neuropsychological tests: a review of the literature on everyday cognitive skills. Neuropsychol Rev 2003; 13(4):181-197.
8. Heaton RK, Pendleton MG. Use of Neuropsychological tests to predict adult patients' everyday functioning. J Consult Clin Psychol 1981;49(6):807-821.
9. Sbordone RJ, Long CJ, editors. Ecological validity of neuropsychological testing. 2nd edition. Boca Raton: St. Lucie Press 1998.
10. Franzen MD, Arnett PA. The validity of neuropsychological assessment procedures. In: Reese HW, Franzen MD, editors. Biological and neuropsychological mechanisms: Life-Span developmental psychology. Mahwah, NJ: Erlbaum 1997:51-69.
11. Heinrichs RW. Current and emergent applications of neuropsychological assessment: problems with validity and utility. Prof Psychol Res Pract 1990;21:171-176.
12. Franzen MD, Wilhelm KL. Conceptual foundations of ecological validity in neuropsychology. In: Sbordone RJ, Long CJ, editors. Ecological validity of neuropsychological testing. Delray Beach, FL: GR Press/St. Lucie Press 1996:91-112.
13. Wilson BA, Alderman N, Burgess PW, Emslie H, Evans JJ. Behavioural Assessment of the Dysexecutive Syndrome (BADS): Thames Valley Test Company 1996.
14. Robertson IH, Ward T, Ridgeway V, Nimmo-Smith I. The structure of normal human attention: the test of everyday attention. J Int Neuropsychol Soc 1996;2:525-534.
15. Wilson BA, Cockburn J, Baddeley AD. The Rivermead Behavioural Memory Test. Flempton, England: Thames Valley Test Company 1985.
16. Williams JM. Everyday cognition and the ecological validity of intellectual and neuropsychological tests. In: Williams JM, Long CJ, editors. Cognitive approaches to neuropsychology. New York: Plenum 1988:123-141.

17. Bouwens SF, van Heugten CM, Verhey FR. Association between cognition and daily life functioning in dementia subtypes. Int J Geriatr Psychiatry 2009 (in press).

18. LaBuda J, Lichtenberg P. The role of cognition, depression, and awareness of deficit in predicting geriatric rehabilitation patients' IADL performance. Clin Neuropsychol 1999;13(3):258-267.

19. Landi F, Bernabei R, Russo A, Zuccala G, Onder G, Carosella L, et al. Predictors of rehabilitation outcomes in frail patients treated in a geriatric hospital. J Am Geriatr Soc 2002;50(4):679-684.

20. Sands LP, Yaffe K, Covinsky K, Chren MM, Counsell S, Palmer R, et al. Cognitive screening predicts magnitude of functional recovery from admission to 3 months after discharge in hospitalized elders. J Gerontol A Biol Sci Med Sci 2003;58(1):37-45.

21. Chaytor N, Schmitter-Edgecombe M, Burr R. Improving the ecological validity of executive functioning assessment. Arch Clin Neuropsychol 2006;21(3):217-227.

22. Odhuba RA, van den Broek MD, Johns LC. Ecological validity of measures of executive functioning. Br J Clin Psychol 2005;44(Pt 2):269-278.

23. Silverberg ND, Hanks RA, McKay C. Cognitive estimation in traumatic brain injury. J Int Neuropsychol Soc 2007;13(5):898-902.

24. Wood RL, Liossi C. The ecological validity of executive tests in a severely brain injured sample. Arch Clin Neuropsychol 2006;21(5):429-437.

25. Zwecker M, Levenkrohn S, Fleisig Y, Zeilig G, Ohry A, Adunsky A. Mini-Mental State Examination, cognitive FIM instrument, and the Loewenstein Occupational Therapy Cognitive Assessment: relation to functional outcome of stroke patients. Arch Phys Med Rehabil 2002;83(3):342-345.

26. Bouwens SF, van Heugten CM, Aalten P, Wolfs CA, Baarends EM, van Menxel DA, et al. Relationship between measures of dementia severity and observation of daily life functioning as measured with the Assessment of Motor and Process Skills (AMPS). Dement Geriatr Cogn Disord 2008;25(1):81-87.

27. Doble SE, Fisk JD, MacPherson KM, Fisher AG, Rockwood K. Measuring functional competence in older persons with Alzheimer's disease. Int Psychogeriatr 1997;9(1): 25-38.

28. Gillen RW, Gernert-Dott P. Relationship of independent living scales to a functional outcome measure and neuropsychological test. Arch Clin Neuropsychol 2000;15:663.

29. Kizony R, Katz N. Relationships between cognitive abilities and the process scale and skills of the assessment of motor and process skills (AMPS) in patients with stroke. Occupation, Participation and Health 2002;22(2):82-92.

30. Johnson JL. Episodic memory deficits in Alzheimer's disease: a behaviorally anchored scale. Arch Clin Neuropsychol 1994;9(4):337-346.

31. Ward J, Dow M. Functional Assessment Rating Scale (FARS). In: Coughlin KM, editor. behavioral outcomes & guidelines sourcebook. New York: Faulkner & Gray 1999: 461-462.

32. Malec JF. Goal Attainment Scaling in Rehabilitation. Neuropsychol Rehabil 1999;9(3/4):253-275.

33. Nygard L, Winblad B. Measuring long term effects and changes in the daily activities of people with dementia. J Nutr Health Aging 2006;10(2):137-138.

34. Sohlberg MM, Mateer CA, Penkman L, Glang A, Todis B. Awareness intervention: who needs it? J Head Trauma Rehabil 1998;13(5):62-78.

35. Doble SE, Dalhousie U. Measuring competence in the performance of activities of daily living: the forgotten half of Alzheimer's disease. Dissertation Abstracts International Section B The Sciences and Engineering 2001;62(2-B):762.

36. Fisher A. Assessment of motor and process skills. 3th edition. Fort Collins, CO: Three star press 1999.

37. Gitlin LN, Winter L, Burke J, Chernett N, Dennis MP, Hauck WW. Tailored activities to manage neuropsychiatric behaviors in persons with dementia and reduce caregiver burden: a randomized pilot study. Am J Geriatr Psychiatry 2008;16(3):229-239.

38. Anderson P. Assessment and development of executive function (EF) during childhood. Child Neuropsychol 2002;8(2):71-82.

39. Bennett TL. Neuropsychological evaluation in rehabilitation planning and evaluation of functional skills. Arch Clin Neuropsychol 2001;16(3):237-253.

40. Burgess PW, Alderman N, Forbes C, Costello A, Coates LM, Dawson DR, et al. The case for the development and use of "ecologically valid" measures of executive function in experimental and clinical neuropsychology. J Int Neuropsychol Soc 2006;12(2):194-209.

41. Chaytor N, Temkin N, Machamer J, Dikmen S. The ecological validity of neuropsychological assessment and the role of depressive symptoms in moderate to severe traumatic brain injury. J Int Neuropsychol Soc 2007;13(3):377-385.

42. Coulthard E, Husain M. Bringing cognitive testing into the real world. J Neurol Neurosurg Psychiatry 2008;79(4):363.

43. Rabin LA, Burton LA, Barr WB. Utilization rates of ecologically oriented instruments among clinical neuropsychologists. Clin Neuropsychol 2007;21(5):727-743.

44. Wood RL, Rutterford NA. Relationships between measured cognitive ability and reported psychosocial activity after bilateral frontal lobe injury: an 18-year follow-up. Neuropsychol Rehabil 2004;14(3):329-350.

45. Stuss DT, Levine B, Alexander MP, Hong J, Palumbo C, Hamer L, et al. Wisconsin Card Sorting Test performance in patients with focal frontal and posterior brain damage: effects of lesion location and test structure on separable cognitive processes. Neuropsychologia 2000;38:388-402.

46. Manchester D, Priestley N, Jackson H. The assessment of executive functions: coming out of the office. Brain Inj 2004;18(11):1067-1081.

47. Sunderland A, Walker CM, Walker MF. Action errors and dressing disability after stroke: an ecological approach to neuropsychological assessment and intervention. Neuropsychol Rehabil 2006;16(6):666-683.

48. Brunswik E. Perception and the representative design of psychological experiments. 2nd edition. Berkeley: University of California Press 1956.

49. Tupper D, Cicerone KD. Introduction to the neuropsychology of everyday life. In: Tupper D, Cicerone K, editors. The Neuropsychology of everyday life: Assessment and basic competencies. Boston: Kluwer Academic 1990:3-18.

50. Lezak MD, Howleson DB, Loring DW, Hannay HJ. Neuropsychological assessment. 4th edition. New York: Oxford University Press 2004.

51. Ginsberg JP, Kibby MY, Long CJ. Ecological validity of neuropsychological data as indicated by interrelationships with vocational data. Poster session at the Annual Meeting of the National Academy of Neuropsychology; 1995; San Francisco, CA; 1995.

52. Higginson CI, Arnett PA, Voss WD. The ecological validity of clinical tests of memory and attention in multiple sclerosis. Arch Clin Neuropsychol 2000;15(3):185-204.

53. Ready RE, Stierman L, Paulsen JS. Ecological validity of neuropsychological and personality measures of executive functions. Clin Neuropsychol 2001;15(3):314-323.

54. Nadolne MJ, Stringer AY. Ecologic validity in neuropsychological assessment: prediction of wayfinding. J Int Neuropsychol Soc 2001;7(6):675-682.

55. Loring DW, Lezak MD, Howleson DB. Neuropsychological assessment. Oxford: Oxford University Press 2004.

56. Kvavilashvili L, Ellis J. Ecological validity and the real-life/laboratory controversy in memory research: a critical and historical review. History & Philosophy of Psychology 2004;6:59-80.

57. Dywan J, Segalowitz S. Self and family ratings of adaptive behavior after traumatic brain injury: psychometric scores and frontally generated ERPs. J Head Trauma Rehabil 1996;11:79-95.

58. Bullard SE, Fein D, Gleeson MK, Tischer N, Mapou RL, Kaplan E. The Biber Cognitive Estimation Test. Arch Clin Neuropsychol 2004;19(6):835-846.

59. Kaplan EF, Goodglass H, Weintraub S. The Boston Naming Test. Philadelphia: Lea & Febiger 1978.

60. Roth M, Huppert FA, Mountjoy CQ, Tym E. The revised Cambridge examination for mental disorders of the elderly. Second Edition Cambridge: Cambridge University Press 1999.

61. Willer BS, Rosenthal M, Kreutzer J, Gordon W, Rempel R. Assessment of community integration following rehabilitation for traumatic brain injury. J Head Trauma Rehabil 1993;8:75-87.

62. Toglia JP. Contextual memory test manual. Tucson, Arizona: Therapy Skill Builders 1993.

63. COGNISTAT. The neurobehavioral cognitive status examination manual. Fairfax, CA: The Northern California Neurobehavioral Group 1995.

64. Benton AL, Hamsher K, Sivan AB. Multilingual aphasia examination 3rd edition. Iowa City, IA: AJA Associates 1983.

65. Landi F, Tua E, Onder G, Carrara B, Sgadari A, Rinaldi C, et al. Minimum data set for home care: a valid instrument to assess frail older people living in the community. Med Care 2000;38(12):1184-1190.

66. Delis DC, Kramer JH, Kaplan E, Ober BA. California Verbal Learning Test-Second Edition: Adult version. San Antonio, TX: The Psychological Corporation 2000.

67. Rappaport M, Hall KM, Hopkins K, Belleza T, Cope DN. Disability rating scale for severe head trauma: coma to community. Arch Phys Med Rehabil 1982;63(3):118-123.

68. Guide for the Use of the Functional Independence Measure. Uniform Data System for Medical Rehabilitation. Bufallo: State University of New York at Buffallo; 1987.

69. Reitan R, Wolfson D. The Halstead-Reitan Neuropsychological Test Battery: Theory and clinical interpretation. Tucson, AZ: Neuropsychology Press 1985.

70. Martzke JS, Swan CS, Varney NR. Pots-traumatic anosmia and orbital frontal damage: neuropsychological and neuropsychiatric correlates. Neuropsychology 1991;5(3):213-225.

71. Loeb PA. Independent Living Scales (ILS): Pearson Education Incorporation 1996.

72. Benton A, Hamsher K, Varney NR, Spreen O. Contributions to neuropsychological assessment. New York: Oxford University Press 1983.

73. Itzkovich M, Elazar B, Averbuch S, Katz N. LOTCA manual. 2nd edition. Pequannock, NJ: Maddack Inc. 2000.

74. Mattis S. Dementia Rating Scale: Professional Manual. Odessa, Florida: Psychological Assessment Resources 1988.

75. Drubach DA, Kelly MP, Taragano FE. The Montebello Rehabilitation Factor Score. J Neurol Rehabil 1994;8:92-96.

76. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975;12(3):189-198.

77. Rey A. L'examinen psychologique dans les cas d'encephalopathie traumatique. Archives de Psychologie 1941;28:286-340.

78. Mahurin RK, DeBettignies BH, Pirozzolo FJ. Structured assessment of independent living skills: preliminary report of a performance measure of functional abilities in dementia. J Gerontol 1991;46(2):58-66.

79. Wilson B, Cockburn J, Halligan P. Behavioral Inattention Test manual. London, England: Thames Valley Test Company 1987.

80. Stroop JR. Studies of interference in serial verbal reactions. J Exp Psychol 1935;18:643-662.

81. Pfeiffer E. A short portable mental status questionnaire for the assessment of organic brain deficit in elderly patients. J Am Geriatr Soc 1975;23(10):433-441.

82. Reitan R. Trail Making test: Manual for administration, scoring, and interpretation. Bloomington: Indiana University 1956.

83. Wechsler D. Wechsler Adult Intelligence Scale-III. San Antonio, TX: The Psychological Corporation 1997.

84. Berg EA. A simple objective technique for measuring flexibility in thinking. J General Psychology 1948;39:15-22.

85. Wechsler D. WMS-III. Administration and scoring manual. San Antonio, TX: Harcourt Brace Jovanovich 1997.

Chapter 5

The ecological validity of neuropsychological assessment with regard to daily life functioning after acquired brain injury: a different approach

Abstract

Background
Neuropsychologists are frequently asked to predict the daily life functioning of people with brain dysfunction, such as acquired brain injury.

Objective
This study investigates the ecological validity of neuropsychological assessment by exploring to what extent experienced neuropsychologists can predict the daily life functioning of people with acquired brain injury solely on the basis of neuropsychological data.

Methods
Eight neuropsychologists were asked to make predictions about the daily life functioning of brain-injured patients by using neuropsychological test data and data about patients' emotional state, behaviour and personality. The level of agreement between neuropsychologists was calculated and their predictions were compared to the data provided by videotaped observation of the participants performing one or more relevant, familiar and important daily life activities in their own environment.

Results
The level of agreement between neuropsychologists about expected problems during the performance of daily life activities varied considerably. The accuracy rate of the predictions was 60%. A sufficient accuracy rate was attained in only one category, the overall performance of an activity (i.e. success or fail).

Conclusion
The ecological validity of neuropsychological tests is enhanced by taking the neuropsychologist's decision-making process into account. However, a substantial part of daily life functioning remains unexplained by a neuropsychological assessment alone. Direct observation of a patient performing daily life activities provides important and unique information and is therefore recommended in addition to the neuropsychological test assessment.

Introduction

One billion people worldwide are affected by neurological disorders [1] that can lead to deficits in many areas of functioning [2-4]. Neuropsychological assessment is concerned with identifying the cognitive, emotional, psychosocial, and behavioural consequences of brain dysfunction due to acquired brain injury or dementia [5, 6]. Neuropsychological assessment is conducted for different purposes, such as: to support the diagnostic process; to help with management, care and planning; to evaluate the effectiveness of a treatment; and to examine the practical consequences of cognitive impairments for daily life functioning [5, 6]. The purpose of a neuropsychological assessment further depends on the clinical setting. In people with dementia neuropsychological assessment is mainly used to support the diagnostic process, while in people with brain injury it is mainly used to plan rehabilitation treatment and set goals.

The International Classification of Functioning, Disability, and Health (ICF) [7] distinguishes different components related to functioning and disability: body functions and structure (impairment); activities (limitation) and participation (restriction); and environmental factors (context). Impairments (e.g. memory loss) can lead to limitations (e.g. forgetting names), and these limitations may lead to restrictions (e.g. not being able to work as a secretary). However, a deficit on a neuropsychological test (impairment) does not necessarily indicate a deficit in daily life functioning (restriction), as this also depends on environmental factors (context) (e.g. forgetting names is not necessarily a problem for people having solitary professions such as forester or furniture maker). This ambiguous relationship between body function and structure, activities and participation and environmental factors should be taken into account when predicting daily life functioning solely on the basis of a neuropsychological assessment.

The core tool of a neuropsychological assessment consists of standardized tests. These tests are the most effective tools for quantifying impairments [5]. When the neuropsychological assessment aims to examine the consequences of cognitive impairments for daily life functioning, it is important to know to what extent these tests predict practical consequences. Numerous attempts have been undertaken to study the extent to which neuropsychological tests reflect the cognitive abilities necessary for daily life functioning in people with cognitive problems [8-15].

Chaytor & Schmitter-Edgecombe [16] reviewed the so-called ecological validity of neuropsychological tests for everyday cognitive skills and described several shortcomings of these tests. One conclusion was that it is erroneously assumed that the outcome measures used accurately reflect patients' everyday skills. In fact, the outcome measures used are mostly a derivative of everyday functioning. Furthermore, they mentioned that instruments for assessing daily life functioning always contain some degree of error. For example, self-reports and informant-based questionnaires can be biased, since both the patient and the caregiver can over or underestimate everyday functioning. The ideal way to assess daily life functioning is direct observation of patients performing familiar activities in their own environment. This conclusion is in line with Evans [5] who states that since some areas of cognitive functioning

are relatively poorly assessed by standardized tests, observational data about the patient carrying out practical or functional tasks are useful.

However, most studies investigating the ecological validity of neuropsychological assessment measured daily life functioning by using self-reports or informant-based questionnaires instead of direct observation [10, 17, 18]. Some studies did observe daily life functioning in a simulated naturalistic environment [19, 20], but none of the studies used direct observation of patients in their own environment.

The aim of this study is to investigate the ecological validity of neuropsychological assessment by exploring the extent to which experienced neuropsychologists can predict daily life functioning in people with acquired brain injury solely on the basis of neuropsychological data.

Methods

Although the neuropsychological test forms the main ingredient of the neuropsychological assessment, we are aware that the neuropsychological assessment is based on more than cognitive tests alone; it also contains information about other relevant aspects like emotional state, behaviour and personality [6]. Therefore, besides providing the neuropsychologists with a set of neuropsychological test results, we also provided them with observational data gathered during the test assessment, as well as with data on mood and an overview of the level of activities of the individual patient.

Design
This cross-sectional study contained neuropsychological data and observational data about brain-injured patients performing one or more problematic daily life activities in their own environment. Eight neuropsychologists were asked to make predictions about the patient's capacity to perform the precise daily life activity as observed in the patient's environment on the basis of neuropsychological data, as well as on the basis of data about patient's emotional, behavioural and personality aspects. The neuropsychologists' predictions were compared to the data provided by videotaped observation of the participants performing one or more daily life activities in their own environment. The Ethics Committee approved the procedure.

Patient data
Data came from brain-injured patients who were referred to a cognitive rehabilitation programme by neurologists, general practitioners and rehabilitation physicians. The programmes were provided by rehabilitation centre Blixembosch in Eindhoven and rehabilitation centre Hoensbroeck in Hoensbroek between September 2006 and December 2007.

The criteria for referral to the programme are the following: the patient has sustained an acquired brain injury at least 3 months earlier; the patient is older than 18; the patient experiences cognitive, emotional, and/or behavioural problems that interfere with daily functioning for which information, advice and treatment are necessary in order to assure

adaptation to these problems; the patient has sufficient insight into his or her own problems to benefit from low frequency treatment; communication in daily life is not hindered (the patient can read and write); the patient's social behaviour is adequate for functioning in a group programme; there are no treatment goals in the sensori-motor area of functioning (any more). Patients with primary psychological or psychiatric disorders, dementia, or a whiplash trauma are excluded from the programme.

Neuropsychologists

The neuropsychologists who were asked to participate in this study all work at the Maastricht University Hospital in Maastricht, The Netherlands. All had ample experience in the clinical neuropsychological diagnosis of people with cognitive disorders.

Measures

Neuropsychological assessment

Neuropsychological tests

The test assessment consists of well-accepted and validated neuropsychological tests of which adequate norms are available. Short-term memory was measured with the Digit Span and Reverse Digit Span of the Wechsler Adult Intelligence Scale III (WAIS-III) [21]. Episodic memory was tested with the Auditory Verbal Learning Test (AVLT) [22] and the Complex Figure Test (CFT) [23, 24]. Attention was measured with the D2 [25], part A of the Trail Making Test (TMT) [26] and card 1 of the Stroop Colour Word Test (SCWT) [27]. Executive functions were measured with card 3 of the SCWT, the action plan test, the rule shift test, the zoo map test and the six element test of the Behavioural Assessment of Dysexecutive Syndrome (BADS) [28] and part B of the TMT. Language was assessed with the naming part of the Aachen Aphasia Test (AAT) [29], verbal phonemic fluency with the digits 'D', 'A' and 'T', and verbal semantic fluency of the Groninger Intelligence Test (GIT) [30]. General intellectual functioning was assessed with the Matrices reasoning of the WAIS-III [21].

To enhance the validity of the neuropsychological test scores, observational data gathered during the test assessment as well as data about patients' mood and neuropsychiatric symptoms were incorporated in this study, as described below.

Observational data during test assessment

Aspects that were observed during the neuropsychological assessment were: general appearance, sensory and motor functions, social skills, language, mood, concentration, motivation, attitude, and work method. The neuropsychological test assistant assessed these aspects in a structured manner for all participants by filling in an observation form during the test assessment.

Mood
The Hospital Anxiety and Depression Scale (HADS) [31] was used for measuring mood. It is a self-screening questionnaire for depression and anxiety. It consists of 14 questions, seven for anxiety and seven for depression. The range for both anxiety and depression is 0-21. Cut-off for each domain is 8 or higher, indicating the presence of depressive and or anxiety symptoms.

Neuropsychiatric symptoms
The Neuropsychiatric Inventory (NPI) [32] was used for measuring the presence of neuropsychiatric symptoms. The NPI is a retrospective (up to 1 month) informant-based rating scale developed to assess psychopathology in patients with neurodegenerative diseases. It is a semi-structured interview assessing 12 neuropsychiatric symptoms: delusions, hallucinations, depression, anxiety, agitation, euphoria, disinhibition, irritability, apathy, aberrant motor behaviour, night-time behaviour disturbances, and appetite/eating abnormalities.

Daily life functioning
Instrumental activities of daily life (IADL) interview
To identify relevant, important and problematic daily life activities of brain-injured patients according to the patient and/or caregiver, we used a semi-structured interview about instrumental activities of daily life (IADL). The interview is based on the International Classification of Functioning, Disability, and Health (ICF) [7] and existing instruments for measuring IADL, such as the Lawton IADL Scale [33] and the Bristol Activities of Daily Living Scale [34]. The IADL interview used in this study consists of 87 IADL activities.

Videotapes of everyday functioning
Videotapes of participants performing one or more problematic daily life activities in their own environment were used to objectively observe and identify problems encountered in daily life. These videotapes indicate the quality of the performance and were made by the observer, who was also the researcher (SB, first author). The participants were acquainted with the observer from former appointments. During observation, the participants were instructed to perform one or more relevant and important activities the way they would do normally. The choice of activities was based on the IADL interview with the patient and the caregiver. Table 1 gives an overview of the activities observed during this study. The activities had to be relevant and important to the patient and a time limit of one hour was set for observing the activities in the patient's own environment. The participants were told that they were allowed to do anything in order to perform the activity successfully (for instance use strategies). The observer was 'a fly on the wall', but in cases where the interviews indicated that the participants only had problems when they were distracted, the observer started a conversation or asked questions to distract them. In all other cases, the observer did not intervene as long as she was not asked to do so and as long as the situation was safe.

Table 1. Observed daily life activities

Activity	Sub activity
Preparing a lunch	Making a cup of coffee
	Making a cup of tea
	Making a sandwich
	Baking/boiling an egg
Computer use	Browsing the internet for information
	Sending an email
	Creating and saving a text document in Microsoft Word
Cell phone use	Making a phone call
	Sending a text message
	Adding a contact person to the contact list
Television use	Using CEEFAX
	Programming a television programme with a digital television recorder

Procedure

Each neuropsychologist received the neuropsychological assessment data and they were given a list of relevant and problematic activities based on the IADL interview. The neuropsychologists were told that the activities on the list were relevant and important for the patient. They were asked to write down whether, and to what extent, they expect problems during the performance of these daily life activities. If difficulties were expected, the neuropsychologists were asked to describe in concrete and observable terms what kind of problems the participant would encounter. To assist the neuropsychologists in formulating concrete and observable predictions, examples were given such as 'person stops executing task and focuses on irrelevant things', or, 'person fails to notice mistakes and keeps repeating unsuccessful actions'.

The neuropsychologists were also asked to answer structured questions such as: 'Do you expect there to be an emotional reaction?'; 'Do you expect behaviour to change?'; 'Does the participant need help?'; 'To what extent does the participant need help?'; and 'Will the participant succeed in performing the activity?'. Afterwards the predictions were divided into the following categories: memory, attention, speed, executive functioning, language, abstract reasoning, mood, environment and other. The structured questions covered the categories emotion, behaviour, role of the observer, level of help and overall performance (i.e. success or fail).

In order to determine the level of agreement between the neuropsychologists, the neuropsychological assessment data and the lists of 16 observed activities from 10 participants were provided to the neuropsychologists in such a way that each neuropsychologist received data about 6 observations and each observation was evaluated by three different neuropsychologists.

The observer, who was not part of the team of neuropsychologists and only answered procedural questions, compared the neuropsychologists' predictions with the videotapes to check if the predictions of the neuropsychologists were correct. After the predictions were scored by the observer, she showed the videotapes to each neuropsychologist separately to discuss the following points: why predictions were correct or incorrect; which aspect of making predictions presented difficulties and which did not; and what kind of information necessary for making reliable predictions was missing.

Statistical analyses

The AC_1 statistic as proposed by Gwet [35] was used to calculate the level of agreement between neuropsychologists about the expected problems participants would encounter during the performance of the observed daily life activities. The AC_1 statistic is a robust extension of Cohen's kappa for assessing the interrater reliability of more than two raters simultaneously [36]. The level of agreement was moderate when AC_1 is between 0.40 and 0.59, substantial between 0.60 and 0.79, and outstanding when 0.80 or higher was scored [37, 38].

To test the accuracy of predictions, total numbers and percentages were calculated for the number of correct and incorrect predictions in total and per category. Predictions were considered accurate when the rate was 80% or higher. Qualitative descriptions were made about the results of showing the videotapes to the neuropsychologists after they made the predictions. Statistical analyses were performed with the Statistical Package for Social Sciences (version 16.0) and Statistical Analyses System (version 9.1) with an alpha level set at 0.05 for all analyses.

Results

Level of agreement

Table 2 shows the level of agreement between neuropsychologists with regard to their predictions as to which category the participant might encounter problems in during the performance of the observed daily life activities. Level of agreement for language was outstanding and significant. Since there was no variability in the scores for abstract reasoning, no p-value could be calculated. Level of agreement for the category 'other' could not be calculated since the variability in content of the predictions was too high, or the predictions could not be checked with the videotape (i.e. 'if the cat jumps on the couch, the participant is distracted'). Level of agreement was substantial and significant for mood, emotion and overall performance. It was moderate and significant for attention and behaviour. Memory, environment and level of help had significant but low levels of agreement. Executive functioning, speed, and the role of the observer all showed low and non-significant levels of agreement.

Table 2. Level of agreement between neuropsychologists

Category	Level of agreement (AC_1)
Predictions	
Memory	0.27*
Attention	0.54*
Speed	0.08
Executive functioning	0.08
Language	0.85*
Mood	0.64*
Environment	0.25*
Structured questions	
Emotion	0.64*
Behaviour	0.56*
Role of observer	0.00
Level of help	0.33*
Overall performance	0.74*

* $p < 0.05$

Accuracy of predictions

Eight neuropsychologists made 214 predictions about the quality of the performance of daily life activities and the problems they expected patients to encounter, based on the neuropsychological tests and additional data. Table 3 shows the accuracy rates of the predictions for the total number of predictions and per category. Nearly 60 percent of these predictions were correct. Memory was the category with the lowest rate of accurate predictions, and language had the highest rate of accurate predictions (50% and 75% respectively). Table 4 presents the accuracy rates of predictions for the structured questions. Level of help had the lowest rate of accurate predictions and overall performance the highest (13% and 83% respectively).

Table 3. Accuracy of predictions

	n	Correct n (%)
Memory	10	5 (50)
Attention	30	19 (63)
Speed	30	18 (60)
Executive functioning	38	27 (71)
Language	4	3 (75)
Mood	38	21 (55)
Environment	10	6 (60)
Other	54	28 (52)
Overall predictions	214	127 (59)

Table 4. Accuracy of predictions for the structured questions (n = 48)

	Correct n (%)
Emotion	36 (75)
Behaviour	36 (75)
Role of observer	21 (44)
Level of help	6 (13)
Overall performance	40 (83)

Qualitative analysis post hoc

After seeing the videotapes the neuropsychologists explained why they thought predictions were correct or incorrect; which aspect of making predictions presented difficulties and which did not; and what kind of information necessary for making reliable predictions was missing. The most frequent comment was the absence of the raw data of the tests. The neuropsychologists only received standardized scores of the neuropsychological tests. Insight into how participants achieve their scores – based upon the raw data - guides the neuropsychologist in making predictions about daily life functioning.

Some neuropsychologists were less experienced with the NPI than others. This made it difficult to base expectations on this measure. Others mentioned that an explanation for their incorrect predictions for some cases was that they had paid more attention to other sources of information than the test scores, while the latter turned out to be the best source of information in these cases.

Two participants had no cognitive deficits according to the neuropsychological

assessment. Two out of six neuropsychologists who had to predict the daily life functioning of these two patients did however expect cognitive problems during performance.

Neuropsychologists further mentioned that they found it difficult to make predictions about a person whom they have never seen, e.g. the appearance of a patient or the way the patient talks is deemed to be valuable data in making predictions. After watching the videotapes, neuropsychologists mentioned that the way a household is organized provides relevant information for formulating a treatment plan. The lack of information about patients' appearance and environment makes it difficult to predict daily life functioning.

Discussion

This study is the first attempt to investigate the ecological validity of neuropsychological assessment by taking a different approach and exploring the extent to which experienced neuropsychologists can make predictions about daily life functioning of people with acquired brain injury solely on the basis of neuropsychological data. The level of agreement between neuropsychologists about expected daily life problems varied considerably. The overall rate of accuracy of the predictions was 59%. Since there are no reference data we chose 80% as a cut-off point for sufficient accuracy. This means that the overall accuracy of the predictions in the present study was insufficient.

The low levels of agreement between neuropsychologists are remarkable. We expected to find that if they received the same information about a participant neuropsychologists would come to comparable predictions concerning participants' capacities to function in daily life. These findings point to the conclusion that the decision-making process of neuropsychologists is not as straightforward as it may seem. Most neuropsychologists correctly predicted a patient's overall performance, but how participants achieved this performance proved to be still unclear. Since we found that predictions made about a patient differed in category and varied in number, we must conclude that neuropsychologists might not pay attention to different sources of information to the same extent. Unravelling the specific contributions of the different sources of information (e.g. neuropsychological tests, mood, neuropsychiatric symptoms) in making predictions about the daily life functioning of people with acquired brain injury is still a challenge.

Memory and executive functioning, which includes attention, strategy use and planning, are cognitive domains of particular importance in daily life functioning [12]. It is therefore remarkable that in this study memory is predicted with the least sufficient level of accuracy (50%). Executive functioning is, however, predicted with almost sufficient level of accuracy (71%).

A further remarkable finding is that some neuropsychologists paid more attention to observational data during the test assessment and to the behavioural data of the NPI, than to the neuropsychological tests. Since these tests are the core tool of neuropsychologists, we expected that neuropsychologists to pay more attention to them.

All participants in this study were in the chronic phase of their disease. They had to cope with both evident and subtle difficulties. Especially the subtle difficulties are hard to

catch with neuropsychological tests alone. Bouwens et al. [39] showed that cognitive tests could predict only 30% of daily life functioning. They recommended also paying attention to emotional, neuropsychiatric and environmental functioning. The present study showed that when neuropsychologists take these additional sources into account, their ability to predict daily life functioning increases by almost 30%. Theoretically this means that 40% of daily life functioning still cannot be explained on the basis of a neuropsychological assessment. Direct observation may possibly provide more insight into these subtle problems. In addition, all neuropsychologists who participated in this study agreed that observing how people behave in their own environment provides valuable information for further treatment.

Strengths and limitations

The strength of this study lies in how the investigation of the ecological validity of neuropsychological assessment was operationalized. First of all, the reference was the direct observation of daily life activities in the patients' own environment. This approach increases the ecological validity [13]. Secondly, using neuropsychologists to predict daily life functioning on the basis of neuropsychological data is innovative. To our knowledge, no other study has investigated ecological validity in this way.

The participants were brain-injured patients with a relatively high level of functioning. Therefore, our sample is selective, making it difficult to extrapolate the results to other groups, such as brain-injured patients with a lower level of functioning. On the other hand, however, the use of this selective group enables us to examine the sensitivity of neuropsychological assessment when it comes to predicting the daily life functioning of a cognitively less impaired group. Predicting the outcome is more difficult when patients have less severe cognitive problems.

This exploratory study also has some limitations. Due to the small number of predictions in some categories, we could not perform analyses to correct for chance. These results must therefore be interpreted with caution. The neuropsychologists who participated in this study, all came from the same department. Therefore, it is possible that the neuropsychologists have the same way of working and that this biases the results. Although they all had ample experience, we could not control whether the level of experience with the neuropsychological data was the same for all neuropsychologists. Furthermore, the neuropsychologists did not work in a rehabilitation setting; rehabilitation was part of their work, but all the neuropsychologists worked with all kinds of populations and referrals.

A general disadvantage of direct observation is the complexity of the interaction between the patient and the environment, possibly making it unable to detect the impact of specific neuropsychological deficits [40]. Since this problem is inherent to direct observation, we could not control for it.

Recommendations for future research

Future research should use a larger sample of both patients and neuropsychologists. To test for interrater reliability the neuropsychologists should be recruited from different departments and from different regions and they should have experience with some form of rehabilitation referrals. Neuropsychologists should be provided with raw test data when asked to make

predictions. These raw data provide insight into how a patient attains a certain test score, and supplement the observational data gathered during the neuropsychological assessment.

Predictions will be more reliable as well if the neuropsychologists are given the opportunity to see the patient in advance and have a conversation with the patient. Future research should focus on repeated observations of the same activities in the patient. This will control for chance. Moreover, the decision-making process used by neuropsychologists when they are confronted with referral questions that involve predictions about daily life functioning should be examined in further detail. One way to do this is to focus in greater detail on the different sources of information neuropsychologists use for making predictions.

Clinical implications

The ideal experiment would have been a participant observation, a form of observation in which the observer is part of the setting and the patients do not know that they are being observed. This form of observation is mostly seen in inpatient settings. However, for our study, participant observation was not feasible. We opted for non-participant observation in the participant's environment. The presence of the observer could have influenced the performance of the participant, but since all participants knew the observer from former appointments, we think the influence of the observer is limited.

The use of a camera is debatable, but there is evidence that the presence of a camera does not affect functioning [41]. Although the use of the camera has some (technical) drawbacks, the advantages outweigh the disadvantages: video recording captures non-verbal and verbal interactions simultaneously so that these can be analyzed in detail, and tapes can be played back [41]. Furthermore, observation at home may change roles. In the clinic, the patient and the caregiver are guests and the clinician is the host. Therefore the patient and the caregiver will behave differently when they are in their own environment where the observer will be the guest and they are the hosts [42]. This aspect will make observation at home a valuable source besides neuropsychological tests.

Direct observation is not always necessary, but certainly indicated when patients live alone or when the caregiver cannot give a clear view of how the patient acts.

Conclusion

The ecological validity of a neuropsychological assessment is enhanced by taking the neuropsychologist's decision-making process into account. It is still the case, however, that when neuropsychologists are asked to predict the daily life functioning of people with acquired injury solely on the basis of neuropsychological data, they do so with insufficient accuracy. Our study was the first attempt to make the decision-making process of neuropsychologists explicit. We accomplished this by confronting them with referral questions that involved the prediction of daily life functioning. Neuropsychologists seem to make use of different sources of information when asked to predict the daily life functioning of the same person. This lack of clarity about the way different sources of information contribute to the predictions influences the investigation of the ecological validity of neuropsychological assessment. To improve the clinical implications of ecologically valid approaches, the decision-making process requires further study.

Direct observation of patients performing daily life activities provides important, unique and necessary information, which cannot be gained from a neuropsychological assessment. Information about how patients behave in their own environment will provide neuropsychologists with insights that are important when it comes to planning adequate treatment or rehabilitation.

References

1. World Health Organization. Neurological disorders: Public health challenges. Geneva; 2007.

2. Ettenhofer ML, Abeles N. The significance of mild traumatic brain injury to cognition and self-reported symptoms in long-term recovery from injury. J Clin Exp Neuropsychol 2008;11:1-10.

3. Greenwood R. The consequences of brain injury: Classification assessment of outcome. Neuropsychol Rehabil 1999;9(3-4):231-240.

4. Sohlberg MM, Mateer CA. Cognitive rehabilitation: An integrative neuropsychological approach. New York: The Guilford Press 2001.

5. Evans JJ. Basic concepts and principles of neuropsychological assessment. In: Halligan PW, Kischka U, Marshall JC, editors. Handbook of clinical neuropsychology. Oxford: Oxford University Press 2003:15-47.

6. Lezak MD, Howleson DB, Loring DW, Hannay HJ. Neuropsychological assessment. 4th edition. New York: Oxford University Press 2004.

7. World Health Organization. International Classification of Functioning, Disability and Health. Geneva; 2001.

8. Alderman N, Burgess PW, Knight C, Henman C. Ecological validity of a simplified version of the multiple errands shopping test. J Int Neuropsychol Soc 2003;9(1): 31-44.

9. Bennett TL. Neuropsychological evaluation in rehabilitation planning and evaluation of functional skills. Arch Clin Neuropsychol 2001;16(3):237-253.

10. Burgess PW, Alderman N, Evans J, Emslie H, Wilson BA. The ecological validity of tests of executive function. J Int Neuropsychol Soc 1998;4(6):547-558.

11. Burgess PW, Alderman N, Forbes C, Costello A, Coates LM, Dawson DR, et al. The case for the development and use of "ecologically valid" measures of executive function in experimental and clinical neuropsychology. J Int Neuropsychol Soc 2006;12(2): 194-209.

12. Kibby MY, Schmitter-Edgecombe M, Long CJ. Ecological validity of neuropsychological tests: focus on the California Verbal Learning Test and the Wisconsin Card Sorting Test. Arch Clin Neuropsychol 1998;13(6):523-534.

13. Manchester D, Priestley N, Jackson H. The assessment of executive functions: coming out of the office. Brain Inj 2004;18(11):1067-1081.

14. Wilson BA. Ecological validity of neuropsychological assessment: do neuropsychological indexes predict performances in everyday activities? Appl Prev Psychol 1993;2: 209-215.

15. Wood RL, Liossi C. The ecological validity of executive tests in a severely brain injured sample. Arch Clin Neuropsychol 2006;21(5):429-437.

16. Chaytor N, Schmitter-Edgecombe M. The ecological validity of neuropsychological tests: a review of the literature on everyday cognitive skills. Neuropsychol Rev 2003; 13(4):181-197.

17. Hanks RA, Millis SR, Ricker JH, Giacino JT, Nakese-Richardson R, Frol AB, et al. The predictive validity of a brief inpatient neuropsychologic battery for persons with traumatic brain injury. Arch Phys Med Rehabil 2008;89(5):950-957.

18. Hanks RA, Rapport LJ, Millis SR, Deshpande SA. Measures of executive functioning as predictors of functional ability and social integration in a rehabilitation sample. Arch Phys Med Rehabil 1999;80(9):1030-1037.

19. Chevignard M, Taillefer C, Picq C, Pradat-Diehl P. Ecological assessment of executive functions in a patient with acquired brain injury. Ann Readapt Med Phys 2008; 51(2):74-83.

20. Chevignard MP, Taillefer C, Picq C, Poncet F, Noulhiane M, Pradat-Diehl P. Ecological assessment of the dysexecutive syndrome using execution of a cooking task. Neuropsychol Rehabil 2008;18(4):461-485.

21. Wechsler D. WAIS-III, Nederlandstalige bewerking, technische handleiding. Lisse: Swets Test Publishers 1997.

22. Rey A. L'examen clinique en psychologie [The clinical examination in psychology]. Paris: Presses Universitaires de France 1958.

23. Osterrieth PA. Le test de copie d'une figure complex: contribution a l'etude de la perception et de la memoire. Archives de Psychologie 1944;30:286-356.

24. Rey A. L'examinen psychologique dans les cas d'encephalopathie traumatique. Archives de Psychologie 1941;28:286-340.

25. Brickenkamp R. Aufmerksamkeits-Belastungs-Test (test d2) [The d2 test of attention]. Gottingen: Hogrefe 1962.

26. Reitan R. Trail Making test: Manual for administration, scoring, and interpretation. Bloomington: Indiana University 1956.

27. Stroop JR. Studies of interference in serial verbal reactions. J Exp Psychol 1935;18: 643-662.

28. Wilson BA, Alderman N, Burgess PW, Emslie H, Evans JJ. Behavioural Assessment of the Dysexecutive Syndrome (BADS): Thames Valley Test Company 1996.

29. Graetz P, de Bleser R, Willmes K. Akense Afasietest, Nederlandse versie. Lisse: Swets & Zeitlinger 1992.

30. Luteijn F, van der Ploeg FAE. Manual Groninger Intelligence Test. Lisse: Swets and Zeitlinger 1983.

31. Snaith RP. The Hospital Anxiety and Depression Scale. Health Qual Life Outcomes 2003;1(1):29.

32. Cummings JL, Mega M, Gray K, Rosenberg-Thompson S, Carusi DA, Gornbein J. The Neuropsychiatric Inventory: comprehensive assessment of psychopathology in dementia. Neurology 1994;44(12):2308-2314.

33. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. Gerontologist 1969;9(3):179-186.

34. Bucks RS, Ashworth DL, Wilcock GK, Siegfried K. Assessment of activities of daily living in dementia: development of the Bristol Activities of Daily Living Scale. Age Ageing 1996;25(2):113-120.

35. Gwet K. Handbook of inter-rater reliability: STATAXIS Publishing Company 2001.

36. King JE. Software solutions for obtaining a Kappa-type statistic for use with multiple raters. Southwest Educational Research Association. Dallas, Texas 2004.

37. Landis JR, Koch GG. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. Biometrics 1977; 33(2):363-374.

38. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33(1):159-174.

39. Bouwens SF, van Heugten CM, Aalten P, Wolfs CA, Baarends EM, van Menxel DA, et al. Relationship between measures of dementia severity and observation of daily life functioning as measured with the Assessment of Motor and Process Skills (AMPS). Dement Geriatr Cogn Disord 2008;25(1):81-87.

40. Sunderland A, Walker CM, Walker MF. Action errors and dressing disability after stroke: an ecological approach to neuropsychological assessment and intervention. Neuropsychol Rehabil 2006;16(6):666-683.

41. Caldwell K, Atwal A. Non-participant observation: using video tapes to collect data in nursing research. Issues in research 2005;13(2):42-54.

42. von Koch L, Wottrich AW, Holmqvist LW. Rehabilitation in the home versus the hospital: the importance of context. Disabil Rehabil 1998;20(10):367-372.

Chapter 6

The application of an integrated approach
combining neuropsychological assessment and
observational data in clinical practice:
data from PROFINTEG

SFM Bouwens, CM van Heugten, E Salmon, S Majerus, M Poncelet & FRJ Verhey

Abstract

Patients with brain dysfunction often experience restrictions in daily functioning, either socially or within their profession. Neuropsychological assessment is usually used to identify the cognitive deficits that underlie this interference with daily life activities. However, the extent to which neuropsychological tests can be used to predict daily life functioning is moderate. To improve predictions, an approach is required that takes a closer look at the discrepancies between neuropsychological testing and daily life functioning, in order to identify the factors that lead to interference in a particular individual situation. By identifying the specific individual difficulties, rehabilitation can be improved and tailored to the individual patient's needs.

In this chapter we describe an international cooperative project called PROFessional INTEGration (PROFINTEG) which attempts to enhance the ecological validity of assessments by integrating neuropsychological data and observational data when assessing patients with brain dysfunction.

Introduction

Many patients with brain dysfunction experience restrictions in their daily life functioning, either socially or within their profession [1, 2]. Neuropsychological assessment is the usual method used to identify the cognitive deficits that underlie this interference with daily life activities [3, 4]. In previous chapters, we demonstrated that the use of neuropsychological tests and cognitive assessment instruments to predict daily life functioning is, at best, only moderately successful [5-7]. This is due to problems related to the ecological validity of neuropsychological assessments [5]. Consequently, many patients are faced with sustained problems upon returning to their social or work situation, despite successful cognitive rehabilitation.

One way to improve the predictive validity of neuropsychological tests is to design tests with higher ecological validity, i.e. tests that measure aspects of cognitive functioning which more directly parallel the cognitive tasks in daily life, such as the Rivermead Behavioural Memory Test (RBMT) [6] and the Test of Everyday Attention (TEA) [7]. An alternative approach is to have a closer look at the discrepancies between neuropsychological testing and daily life functioning in order to identify the factors that lead to the interference in the particular individual situation. When the specific individual difficulties are known, rehabilitation can be improved and tailored to the individual patient.

Although several standardized observational instruments exist, such as the frequently used Assessment of Motor and Process Skills (AMPS) [8], virtually no instrument makes use of the neuropsychological framework, thereby neglecting the added value of the neuropsychological approach to provide a better understanding in terms of cognitive functioning.

In this chapter we describe an international cooperative project called PROFINTEG which attempts to enhance the ecological validity of neuropsychological assessments by integrating neuropsychological data and observational data when assessing patients with brain dysfunction.

Development of the PROFINTEG assessment procedure

General aims of PROFINTEG

The general aim of PROFINTEG was to develop an assessment procedure that could clarify difficulties that patients with brain dysfunction experience while performing a particular instrumental activity of daily life (IADL) in terms of underlying cognitive, psychological and behavioural deficits. A three-level model was used for this purpose, consisting of: 1) a battery of neuropsychological tests and other psychometric instruments assessing mood and behavioural problems; 2) a (newly developed) IADL scale based on a patient and informant interview; and 3) a (newly developed) instrument providing a detailed description of a limited number of difficulties encountered by the individual patient in concrete situations. The components of this model are schematically shown in Figure 1.

Figure 1. Association between levels

Level 1:
• Neuropsychological test assessment
• HADS
• NPI

Level 2: IADL-interview

Level 3: Direct observation of IADL
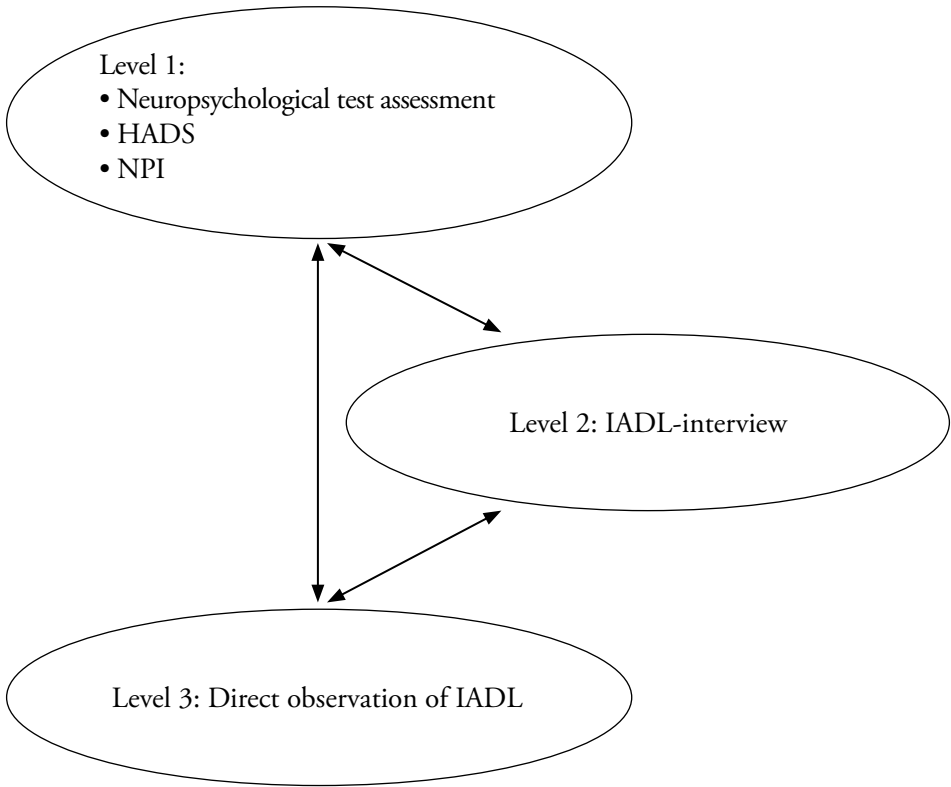
HADS: Hospital Anxiety and Depression Scale; NPI: Neuropsychiatric Inventory; ADL: Instrumental Activities of Daily Living

The integration of the three levels is demonstrated by: 1) selecting one or more activities identified via the IADL interview with both the patient and the caregiver which are particularly difficult but important for the patient/caregiver (level 2); 2) decomposing the different steps of the activity in order to identify the steps which are carried out incorrectly and which hinder the overall execution of the activity (level 3); 3) generating a cognitive hypothesis about the difficulties encountered at each of these steps, based on the patient's cognitive profile (level 1); 4) testing these cognitive hypotheses by rearranging the environment or by helping the patient in such a way that his or her cognitive impairments do not hinder the execution of the activity anymore. The specific aims of the assessment procedure are described in Box 1.

Box 1. Specific aims of the assessment procedure

1. To provide a detailed description of the difficulties encountered by an individual patient when carrying out IADL in his or her own environment

2. To provide a detailed analysis of the cognitive processes involved in these concrete situations

3. To provide a measurement of the patient's available cognitive resources, based on existing neuropsychological instruments

4. To provide a measurement of patient's psychological resources (motivation) and deficits (depression, anxiety, personality disorders), based on existing assessment tools

5. To provide a measurement of the patient's behavioural disorders, based on existing assessment scales

6. To provide an integration of the information gathered in points 1-5, in order to better understand the neuropsychological difficulties in everyday life of the individual patient with brain dysfunction; this understanding will then improve the implementation of intervention strategies

Procedure

The PROFINTEG project was financed by a European INTEREG programme, and unites the following three centres located in the Maas-Rhine Euregion area in the Netherlands and Belgium: 1) Centre de la Memoire, University Hospital of Liege, Belgium; 2) University Hospital Maastricht, the Netherlands; 3) Centre Come-Back, Eupen, Belgium. All of these centres experienced problems with the integration of people with brain dysfunction and consequently they decided to combine their expertise and know-how in the PROFINTEG project. The project team consisted of researchers, neuropsychologists, a psychiatrist, a neurologist, and occupational therapists who met regularly. The project started in November 2004 and ended in October 2008.

On the basis of available data from the literature, the project team agreed on a minimal set of neuropsychological tests in order to identify the cognitive status of the patients with brain dysfunction (Level 1). A literature search was then performed to identify instruments for evaluating people with brain dysfunction who experience problems with IADL functioning. This inventory formed the basis of a newly developed anamnestic instrument to identify IADL problems, since the existing tools did not cover all relevant IADL activities. This interview can be administered to the patient and caregiver separately to evaluate

the capacities of the patient during the performance of various steps of IADL (Level 2). The interview was evaluated in all centres by testing its inter-rater reliability. Since no instrument existed with which to identify the underlying neuropsychological aspects of IADL activities, the project team also developed a new assessment instrument with which to identify the specific cognitive capacities of patients, as well as their mood and the influence of the environment during the performance of a relevant and important activity in their own environment (Level 3). All the centres tested the applicability of the assessment procedure.

Level 1: Neuropsychological assessment

The neuropsychological assessment battery covered short-term memory, episodic memory, basal speed, attention, executive functioning, language and general intellectual functioning. Overall neuropsychological functioning is considered to be impaired when at least one cognitive domain is affected. Apart from the above-mentioned cognitive domains, observational data are collected during test assessment. The Hospital Anxiety and Depression Scale (HADS) [9] is used for measuring mood. The Neuropsychiatric Inventory (NPI) [10] is used for measuring the presence of neuropsychiatric symptoms.

Level 2: Instrumental activities of daily life (IADL) interview

As none of the existing instruments covered all relevant daily life activities, we developed a new IADL instrument, i.e. the IADL interview, which is a semi-structured interview (Appendix 1). The purpose of this interview is to identify which daily life activities are relevant, important and problematic for brain injured patients according to the patient and/or caregiver. The interview is based on the International Classification of Functioning, Disability, and Health (ICF) [11] and on existing instruments for measuring IADL such as the IADL of Lawton [12] and the Bristol Activities of Daily Living Scale [13]. The IADL interview consists of 87 IADL activities and is administered to the patient and caregiver separately. This interview distinguishes various steps of an activity and identifies the level of help a patient needs during the performance of an activity; identifies the importance of the activity for the patient; and assess the burden for the caregiver when assisting the patient with the activity.

Level 3: Direct observation of daily life activities

Patients are observed in their own environment while performing problematic daily life activities, as assessed with the help of the IADL interview. These observations are used to identify and score the underlying cognitive deficits. During the observation at home, the patients are instructed to perform one or more relevant and important activities in the way they normally do. The choice of activities is based on the IADL interview with the patient and the caregiver. The activities need to be relevant for the patient. The patients are told that they are allowed to do anything in order to perform the activity successfully (i.e. using strategies). The observer does not intervene as long as he or she is not asked to do so and as long as the situation is safe. The project team developed 87 grids with which to decompose the activities. In these grids, the sequence of steps during the performance of the activity can be recorded, as well as the level of independence, additional comments and deficit hypotheses (Appendix 2).

Applicability of the PROFINTEG assessment procedure

The aim of this section is to describe the feasibility and applicability of the PROFINTEG assessment procedure. We will mainly focus on the way the IADL activities have been decomposed into their components. The ecological validity of neuropsychological assessment (Level 1) is described in Chapter 5 of this thesis and the validation of Level 2 is described in a report [14].

The sensitivity, specificity and accuracy of the procedure were tested by comparing the neuropsychological assessment (Level 1) and direct observation of daily life activities (level 3) in 48 eligible patients with acquired brain injury who were recruited from the participating centres. The sensitivity measures the proportion of actual positives that are correctly identified as such and the specificity measures the proportion of negatives that are correctly identified as such. Accuracy is the degree of closeness of a measured or calculated quantity to its actual (true) value. With regard to daily life activities, the sensitivity, specificity and accuracy were calculated for the overall neuropsychological assessment and per cognitive domain separately. Both the sensitivity and specificity were high for language (75% and 84% respectively); sensitivity was high for overall neuropsychological functioning (77%); specificity was high for memory, attention and abstract reasoning (71%, 79%, 98% respectively); and accuracy was high for language and abstract reasoning (83% and 92% respectively).

These results show that neuropsychological assessment does not contain sufficient information when the neuropsychologist is asked to examine the everyday consequences of neuropsychological deficits. This was also the conclusion of a study conducted by Bouwens et al. [15]. In this experimental study, the authors investigated the extent to which neuropsychologists could predict daily life functioning in people with acquired brain injury solely on the basis of neuropsychological data.

Qualitative descriptions are presented in Box 2 and Box 3 of a first attempt to decompose the daily life activities of two patients with dementia into components in order to gain a better understanding of the underlying neuropsychological aspects. The patients were recruited from the participating centres. These cases illustrate that the decomposition of daily life activities is relevant for clinicians such as neuropsychologists, who have knowledge of cognitive functions. Observing patients in their own environment while they are performing relevant daily life activities provides unique insight into the neuropsychological deficits causing daily life problems (e.g. level of distraction, attentional capacities and memory). Moreover, during the observation specific cognitive hypotheses can be tested. The cognitive hypotheses about problematic daily life activities (Level 2) have to be based on the results of the neuropsychological assessment (Level 1), as well as on the type of errors made during the observation and the comments of the patient.

The decomposition of each activity into its components is so precise that it will tell the examiner where the problem is and what should be done to relieve it (e.g. cognitive intervention or adapting the environment). We learned that combining the cognitive interpretation of the performance of the activity with the knowledge about the patient's available cognitive resources based on existing neuropsychological instruments (Level 1) will help neuropsychologists to make a more accurate prediction with regard to everyday functioning.

Box 2. Case study 1 showing the usefulness of the assessment procedure

Case 1
- Patient data: 59-year-old female with Alzheimer's disease, MMSE: 22/30
- Aim of the rehabilitation: decreasing problems during meal preparation
- Home visit: identifying the situation at home when patient has to cook a recipe from a diet cookbook (her husband has health problems; therefore she has to make special meals). She forgets each specific step of the recipe. Therefore, she has to consult the recipe over and over again. She also forgets that she recently performed an action and starts with that action again. She needs to reread sentences and gets distracted when her dog walks nearby or when someone rings the doorbell. In the latter cases, she interrupts her activities and asks her husband for help. She does not start with cooking potatoes simultaneously with the other activities but starts with this activity when she reads it (at the end of the recipe)
- Solutions: 1) developing adapted recipes in which the ingredients are listed in a logical order. Small steps are written down in a logical order. A cursor is used to point to the step that needs to be done (this support her working memory problems and lack of planning skills). The other steps are covered with a piece of paper to avoid distraction; 2) use of time schedule in recipe; and 3) advise to husband to minimise external distractions
- Transfer: after training at the rehabilitation centre the patient is able to use the adapted recipes in her own home. Forgetfulness and errors in the performance decreased, as well as asking for help and interruptions of the activity

Box 3. Case study 2 showing the usefulness of the assessment procedure

Case 2
- Patient data: 74-year-old male with vascular dementia and degenerative components, MMSE: 22/30
- Aim of the rehabilitation: decreasing apathy via solving crosswords. This activity was done in the past with great enthusiasm, but nowadays not performed anymore
- Home visit: to understand the situation during the performance of the activity, a simulation was done at home. The clinician noticed language problems, executive problems and memory problems during solving the crossword. The patient needs to be encouraged to start with filling out the crossword. External stimuli easily distract him. During the filling of the crossword he is sensitive for interferences of the complexity of the scheme. He fills in the wrong words or words that are too short or too long
- Solutions: 1) the proposal is to fill in a Swedish puzzle instead of a crossword; 2) advise for his wife to create an optimal environment in which her husband can be encouraged to fill in a puzzle (e.g. prepare the material visibly on a table, wait till he is calm and avoid distraction)
- Transfer: Application of the solutions led to proper performance of the activity at home

Use of the assessment procedure in daily practice and the future direction of the PROFINTEG project

The choice of the specific neuropsychological tests for Level 1 was based on the availability of neuropsychological tests in the participating centres and it had to fulfil the criterion that the total battery did not take too long to administer. Therefore, it is possible that other tests show different results with regard to sensitivity, specificity and accuracy. However, the relevance of neuropsychological assessment for stating cognitive hypotheses about difficulties with IADL is clear. With regard to clinical practice, the choice of the test battery will depend on the availability of tests in each country or clinical setting. We think that it is important that the battery covers at least the following cognitive domains: memory, executive functioning, abstract reasoning, attention and visual perception [16, 17].

The added value of Level 2 of the PROFINTEG project is that the IADL interview is almost complete with regard to relevant everyday activities. Moreover, it is possible to add activities to the list.

The main strength of the PROFINTEG assessment procedure is Level 3; the decomposition of activities provides insight into the underlying neuropsychological deficits, and the observation of patients in their own home provides relevant information for rehabilitation. The use of the assessment procedure depends on the organization of the health

care system in each region. Since many clinical settings have multidisciplinary teams in which observations in patients' homes already take place, the neuropsychological decomposition might be integrated within these observations.

The materials for the decomposition of the observed activities are proposed as aids for guiding a structured and detailed assessment of specific IADL activities. The material should not be considered as final; it can (and should) be adapted by the clinician, as different patients will carry out a given activity in different ways. The latter is especially important in the case of professional activities.

The aim of the PROFINTEG project in the long term will be to make it possible to start the process of social adaptation and to achieve a more efficient professional re-integration. Furthermore, the project will function as a Euregional centre of competence with regard to the ecological evaluation of patients with brain dysfunction.

We are currently working on a paper that investigates the validity of the IADL interview. Moreover, in the future a website will be made available to provide clinicians with information about the use of the assessment procedure.

References

1. Deb S, Burns J. Neuropsychiatric consequences of traumatic brain injury: a comparison between two age groups. Brain Inj 2007;21(3):301-307.

2. Milders M, Fuchs S, Crawford JR. Neuropsychological impairments and changes in emotional and social behaviour following severe traumatic brain injury. J Clin Exp Neuropsychol 2003;25(2):157-172.

3. Evans JJ. Basic concepts and principles of neuropsychological assessment. In: Halligan PW, Kischka U, Marshall JC, editors. Handbook of clinical neuropsychology. Oxford: Oxford University Press 2003:15-47.

4. Lezak MD, Howleson DB, Loring DW, Hannay HJ. Neuropsychological assessment. 4th edition. New York: Oxford University Press 2004.

5. Chaytor N, Schmitter-Edgecombe M. The ecological validity of neuropsychological tests: a review of the literature on everyday cognitive skills. Neuropsychol Rev 2003; 13(4):181-197.

6. Wilson BA, Cockburn J, Baddeley AD. The Rivermead Behavioural Memory Test. Flempton, England: Thames Valley Test Company 1985.

7. Robertson IH, Ward T, Ridgeway V, Nimmo-Smith I. The structure of normal human attention: the test of everyday attention. Journal of the International Neuropsychological Society 1996;2:525-534.

8. Fisher A. Assessment of motor and process skills. 3th edition. Fort Collins, CO: Three star press 1999.

9. Snaith RP. The Hospital Anxiety And Depression Scale. Health Qual Life Outcomes 2003;1(1):29.

10. Cummings JL, Mega M, Gray K, Rosenberg-Thompson S, Carusi DA, Gornbein J. The Neuropsychiatric Inventory: comprehensive assessment of psychopathology in dementia. Neurology 1994;44(12):2308-2314.

11. World Health Organization. International Classification of Functioning, Disability and Health. Geneva; 2001.

12. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. Gerontologist 1969;9(3):179-186.

13. Bucks RS, Ashworth DL, Wilcock GK, Siegfried K. Assessment of activities of daily living in dementia: development of the Bristol Activities of Daily Living Scale. Age Ageing 1996;25(2):113-120.

14. PROFINTEG. A project in collaboration with Centre de la Memoire, Service de Neurologie, CHU Liege, Projekt Come-Back, Behindertanstatten Eupen VoE, Instituut Hersenen en Gedrag, Universiteit Maastricht. Final report and manual; 2008.

15. Bouwens SFM, van Heugten CM, Ponds R, Verhey F. The ecological validity of neuropsychological assessment with regard to daily life functioning after acquired brain injury: a different approach. Submitted for publication.

16. Kibby MY, Schmitter-Edgecombe M, Long CJ. Ecological validity of neuropsychological tests: focus on the California Verbal Learning Test and the Wisconsin Card Sorting Test. Arch Clin Neuropsychol 1998;13(6):523-534.

17. Nys GM, van Zandvoort MJ, de Kort PL, van der Worp HB, Jansen BP, Algra A, et al. The prognostic value of domain-specific cognitive abilities in acute first-ever stroke. Neurology 2005;64(5):821-827.

Appendix 1. Instrumental Activities of Daily Life (IADL) interview

1a	food preparation	15a1	and-line: calling	24c1	watching a videotape
1b	heating/defrosting	15a2	land-line: being called	24c2	watching a DVD
1c	preparation of bread/cake	15a3	land-line: leaving a message	25a	computer: general functions
2a1	making coffee by using a coffee machine	15a4	and-line: listening to voicemail	25b1	computer: word processing
2a2	making coffee by using a senseo	15b1	cell phone: calling	25b2	computer: making graphs and tables using Excel
2a3	making coffee by using a coffee can	15b2	cell phone: being called	25c	computer: scanning digital photos
2a4	making coffee by using a percolator	15b3	cell phone: leaving a message	25d	computer games
2a5	making coffee by using a filter	15b4	cell phone: leaving a message	25e	sending and receiving mails
2a6	making instant coffee	15b5	send SMS	25e2	surfing on the internet
2b1	making tea by using a teabag	15b6	read SMS	26	reading
2b2	making tea by using loose tea	16a	public transfer	27	other activities
3a	laying the table	16b	biking	28	putting away things like keys/ losing things like keys
3b	cleaning the table	16c	walking	29	other
4a	doing the dishes	17a	initiating a meeting		
4b	using a dishwasher	17b	receiving an invitation for a meeting and planning it afterwards		
5a	preparing medication	18	administration		
5b1	taking medication by using a pill-box	19	finances		
5b2	taking medication without using a pill-box	20a	daily use of money		

Appendix 1. (continued)

6a	dusting	20b	credit card use
6b1	vacuuming	21	energy management*
6b2	wiping	22a1	individual game: crossword**
6c	mopping the floor	22a2	individual game: puzzle**
6d	leaning the bathroom	22a3	individual game: alone**
6e	washing windows	22a4	individual game: other**
7a	separate garbage	22b1	collective game: scrabble ***
7b	disposing of garbage	22b3	collective game: other***
8	washing clothes	23a	drawing/painting
9	ironing	23b1	analog photos: automatic
10a	making a shopping list	23b2	analog photos: by hand
10b	doing groceries	23b3	digital photos
11	taking care of pets	23c	other artistic activity
12	watering plants	24a1	operating radio
13a	mowing the lawn	24a2	operating CD player
13b	maintaining vegetable garden	24a3	operating LP
13c	maintaining garden	24a4	operating cassette deck
14a	knitting	24b1	watching TV
14b	ewing	24b2	recording TV

*energy management: turning of a light when leaving the room, closing windows etc.

Appendix 2 . Observation of daily activity

CLEAR THE TABLE Different steps:	Level of independence: from 0 to 4	Sequence of steps: OK or not	Observations: Errors, provided help, comments	Complicating factors			DEFICIT HYPOTHESIS:					
				EMOTIONAL, AFFECTIVE, AND BEHAVIOURAL FACTORS	ENVIRONMENT	SENSORY AND MOTOR FUNCTIONS	WORKING MEMORY	LONG TERM MEMORY	GNOSTIC FACTORS AND PRAXIS	LANGUAGE	ATTENTIONAL FUNCTIONS	EXECUTIVE FUNCTIONS
(1) Clear those things from the table that need to be rinsed a) Successively clear those things from the table that need to be rinsed in a logical order so they can be piled up in the kitchen	0 1 2 3 4											
b) Put the things that fit in the dishwasher, in the dishwasher	0 1 2 3 4											
(2) Clear the other things from the table a) Successively clear the other things from the table and store them in the appropriate place	0 1 2 3 4											

Appendix 2. (continued)

b) Clean out the table-cover and store the mats c) Put the table-cover and mats in the laundry or clean them if they are not suitable for the washing machine	0 1 2 3 4				
(3) Clean the table	0 1 2 3 4				

Level of independence: 0: autonomous; 1: general help (verbal cue, gestures, in case of blockade encourage to continue with the activity; signal an error in a subtle manner, like: Stop! No! You're doing something wrong); 2: concrete help (verbal cue, gestures, point exactly to the error e.g.: Aren't you forgetting something? Isn't that written down somewhere? Point to the nature of the error, without exactly saying what the error is); 3: detailed help (explain in detail and/or show how it works); 4: Take over the activity.

Part III

Individual goal-setting

Chapter 7

Review of Goal Attainment Scaling as a useful outcome measure in psychogeriatric patients with cognitive disorders

Abstract

Background
Since evidence-based interventions are the standard, there is an urgent need for more information concerning individual ways of measuring clinically relevant outcomes of interventions in cognitive disorders such as dementia. Goal Attainment Scaling (GAS) seems to offer a meaningful outcome measure.

Aim
To examine the applicability of GAS in psychogeriatric patients with cognitive disorders.

Methods
A systematic review was performed on the available literature on the clinimetric aspects and the feasibility of GAS when used as an outcome measure for psychogeriatric patients with cognitive disorders. Eight databases were screened. Two authors independently reviewed all the data. Ten studies were included.

Results
Mixed results were found for responsiveness, content validity, inter-rater reliability and construct/convergent validity. The involvement of patient and/or caregiver in the goal-setting procedure is possible and multiple domains can be implemented. The possibility to set at least three realistic goals per patient in less than 30 minutes is unclear and the involvement of a blinded assessor needs is not well established.

Conclusion
GAS proved to be useful on important aspects of an outcome measure for psychogeriatric patients with cognitive disorders. Since other relevant aspects showed mixed results and the number of studies investigating the use of GAS in psychogeriatric patients with cognitive disorders is small, the evidence is not strong enough yet to state that GAS is an applicable outcome measure in this population.

Introduction

Neurodegenerative disorders, such as dementia, cause serious cognitive impairments which lead to interference with daily life functioning [1]. There is an ongoing debate about the best way to measure clinically relevant outcomes of the treatment of dementia [2-4] such as with cholinesterase inhibitors [5, 6]. Outcome measures should not only be responsive, reliable and valid, but also be tailored to the personal goals and needs of patients and their caregivers in relation to their daily life. However, current standardized outcome measures in the field of cognitive disorders do not always take account of these personal aspects. For instance, the widely used ADAS-Cog [7], which is a cognitive scale used in Alzheimer's disease (AD), measures effects on cognition but does not take into account the individual situation of the patient [8].

Goal Attainment Scaling (GAS) is a global outcome measure that does take account of these individual needs. GAS is a measurement method first introduced by Kiresuk and Sherman [9] for programme evaluation in mental health centres [10]. GAS allows for both individualization of patient goals according to the needs of each patient, and standardization of measurement by using a summary formula that calculates the extent to which a patient's goals are met [11]. GAS reflects actual improvement in a patient's functional ability [12] and can be adapted to any level or domain of the International Classification of Functioning, Disability, and Health (ICF) [13], thereby covering aspects of life as a whole. Due to the increasing emphasis on a more client-centered approach, GAS received renewed attention, and has recently been used in patient groups suffering from various disorders, including dementia [14-16].

Despite the increasing use of GAS, its applicability is unclear in the case of psychogeriatric patients with cognitive disorders. Dementia is a multi-domain disorder, characterised by interference with daily life activities, which differs for each individual, as well as by a progressive nature. Hence, to be a useful measure GAS must be responsive to both these aspects. In instances where the clinical relevance of interventions such as treatment with cholinesterase inhibitors is unclear [17], GAS could be used to reveal clinical information relevant for the patient, caregiver and clinician. However, as cognitively impaired patients often lack insight into their capabilities [10], it is to be expected that constructing GAS follow-up guides with these patients could be problematic and less feasible.

The aim here is to examine the applicability of GAS in people with cognitive disorders by reviewing the available literature on the clinimetric aspects and the feasibility of GAS when used as an outcome measure for psychogeriatric patients.

Description of Goal Attainment Scaling

The GAS method consists of a six-step process. To select the goals that are relevant to the individual patient in step 1 the patient and/or caregiver is either interviewed by a clinician to identify problem areas and to determine goals for those areas in which intervention is planned, or team members set goals themselves after interviewing the patient and/or caregiver about problem areas. To ensure reliable calculation with the standard formula, a minimum of three goals is recommended [18]. Step 2 is optional. It provides the opportunity to weight the goals from the most important to the least important with a number between 10 and 1. If there is no hierarchy in the importance of the goals according to the patient and/or therapist, all

weights are set at 1. In Step 3 a follow-up time period before the start of the evaluation is determined, at which point the clinician will measure the level of goal attainment. In step 4 the expected outcome, i.e. the goal, is set. Outcome levels are specified in observable terms in such a way that laymen are able to score the follow-up guides (Table 1). Levels do not overlap. The expected outcome is scored with '0'. Step 5 consists of completion of the other scale levels, using possible outcomes including much less than expected (-2), somewhat less than expected (-1), somewhat better than expected (+1) and much better than expected (+2). In step 6, GAS baseline levels are noted on the follow-up guide. If a clinically relevant deterioration is plausible, the description of the current status can be scored as '-1', with the worsened state scored as '-2'. If the problem is at its worst, the current level of functioning is scored as '-2'. Patients receive an intervention and at the defined follow-up time (e.g. end of intervention), the patient is scored at the attained level. The overall GAS score at baseline and follow-up is calculated with a standard formula (Appendix 1). When goals are weighted, this weight is inserted in this formula. The hypothetical mean GAS score at follow-up is 50 (SD = 10); consistently high or low follow-up scores indicate that goals were too easy (scores > 50) or too difficult to attain (scores < 50) [19]. A GAS score at follow-up of 50 indicates that all predefined goals met the expected outcome at follow-up. Table 1 provides an example of a goal attainment scaling follow-up guide.

Table 1. An example of a goal attainment scaling follow-up guide for three goals of a patient with Alzheimer's disease

	Goal Areas		
	Goal 1	Goal 2	Goal 3
	Dealing with impaired memory	Social contacts	Coping with aggression
Much less than expected (- 2)	Poor short term memory, no use of memory aids	* No interest in social contacts. Sees nobody during the day except caregiver	Becomes aggressive in all situations of misunderstanding
Somewhat less than expected (-1)	* Uses memory aids when another person cues to do so	Ambivalent to visit others. Caregiver needs to be present	* Becomes aggressive in almost all situations of misunderstanding
Program goal (0)	√ Able to use memory aids with external cue like watch	√ Enthusiastic to visit others. Caregivers needs to be present	Learned how to stay calm, but does not succeed in all situations
Somewhat better than expected (+1)	Able to use memory aids spontaneously	Visits other on her own when asked to do so	√ Rarely becomes aggressive
Much better than expected (+2)	No need for memory aids	Initiates visiting others on her own	Never becomes aggressive anymore

* Baseline level; √ Follow-up level

Figure 1. Flowchart of the literature search and study selection

```
┌─────────────────────────────────────────────────────────────┐
│                 Search of electronic databases               │
│       ('free text words' search and controlled vocabulary    │
│       words): CINAHL, PsychINFO, Pubmed, Medline. Cochrane   │
│              Library, EMBASE, Rehabdata and Amed             │
└─────────────────────────────────────────────────────────────┘
                              │
                              ▼
┌─────────────────────────────────────────────────────────────┐
│                    Identification of studies:                │
│   CINAHL = 292,  PsychINFO = 86,  Pubmed = 84, Medline = 69, │
│   Cochrane Library = 135, EMBASE=68, Rehabdata=59 and Amed   │
│                      = 30. Total = 823                        │
└─────────────────────────────────────────────────────────────┘
                              │
                              ▼
┌──────────────────────┐        ┌──────────────────────────────┐
│  Review of 823 titles│ ─────▶ │  Rejection of 637 titles     │
│                      │        │  based on subject matter,    │
│                      │        │  and 64 duplicates           │
└──────────────────────┘        └──────────────────────────────┘
                              │
                              ▼
┌──────────────────────┐        ┌──────────────────────────────┐
│  Review of 122       │ ─────▶ │  Rejection of 112 titles:    │
│  abstract/full text  │        │  Different patient population│
│                      │        │  (21), no psychometric       │
│                      │        │  aspects (55), no full report│
│                      │        │  (5), review (28), and not   │
│                      │        │  published in English (3)    │
└──────────────────────┘        └──────────────────────────────┘
                              │
                              ▼
┌─────────────────────────────────────────────────────────────┐
│                  Inclusion of **10 studies**                 │
└─────────────────────────────────────────────────────────────┘
```

Methods

Literature search and study selection

Figure 1 shows the process of literature search and study selection. Studies on 'GAS' and 'cognitive disorders' and 'clinimetric aspects' were identified by searching the following computerized databases until September 2007: CINAHL, PsychINFO, Pubmed, Medline, Cochrane Library, EMBASE, Rehabdata and Amed. Both controlled vocabulary words and free text words in the title, abstract or keywords were combined in our searches. Details of the search are shown in Appendix 2. The initial search resulted in 823 studies. Two reviewers (SB and CvH) independently assessed all studies for inclusion based on the title. In case of doubt, the first author obtained the abstracts or the full texts, if available and both reviewers decided on inclusion. References of these studies were hand-searched to find additional potentially relevant studies. International requests were submitted when a study was not available in the Netherlands. After excluding 637 references due to non-relevant subject matter

and 64 duplicates, 122 studies remained. These studies were further selected for this review based on the criteria that 1) the study included -at least partially- patients with cognitive disorders (dementia or mild cognitive disorders); 2) the study described clinimetric aspects of GAS; 3) the study used GAS as developed by Kiresuk and Sherman [9], or deviations from this original approach if the aforementioned authors were credited; 4) the study was a full report of a group study describing original data (i.e. not a case study, review, editorial); 5) the study was published in English. Careful reading of the 122 remaining studies, resulted in the rejection of 112 studies because of different patient population (21), lack of description of clinimetric aspects (55), no full report (5), review (28) or non-English publication (3). Finally, a total of 10 studies were included in this review.

All included studies were judged on study characteristics (design, intervention, setting and sample); participant characteristics (mean age, severity level of disease) [20]; clinimetric aspects (responsiveness, reliability, validity); and feasibility (time required for setting the goals, number of goals, the extent to which goals were realistic, involvement of patient and/or caregiver in setting the goals, involvement of a blinded assessor in assessing the follow-up GAS scores and domains that were included in the goals).

Evaluation of the clinimetric aspects

Various definitions of responsiveness exist and responsiveness can be used for different purposes [21]. For our review, we defined responsiveness as the extent to which a measure is capable of detecting change when actual change occurred [22]. Furthermore, a variety of methods exist for calculating responsiveness [21]. In our study, responsiveness was interpreted as positive in the following instances: if the paired sample t-test for analyzing the difference in assessment and the follow-up GAS scores was significant ($p < 0.05$); or if the Norman's method involving repeated measures analysis of variance (ANOVA) [23] was > 0.5; or if the Cohen's d estimate of effect size (ES) proposed by Kazis et al. [24] was > 0.5; or if the standardized response mean (SRM) [25] was > 0.5; or if the relative efficiency (RE) [26] was > 1. Responsiveness was interpreted as negative if the changes were non-significant [27]. Effect size is frequently used to measure the magnitude of an effect of an intervention, but it can be used for any other numerical comparison or contrast [28]. We used the effect size as an operationalisation of the sensitivity to change of GAS as outcome measures compared to secondary measures [29]. Reliability was good when Pearson's correlation coefficients (r), or intraclass correlation coefficients (ICC) were > 0.8, moderate between 0.7 and 0.8, and low when < 0.7 [30]. Content validity was considered good when an expert panel reviewed the selected domains for GAS as good, or when domains correlated with position statements. Construct/convergent validity was considered to be good when the correlation coefficient (Pearson or Spearman) between GAS and secondary outcome measures was > 0.7, moderate if the coefficient was between 0.5-0.7, and low if the coefficient was < 0.5 [31].

Evaluation of feasibility

The time required for setting the goals was considered good when it took a maximum of 30 minutes to construct a GAS follow-up guide of at least three goals. The number of goals was considered sufficient when at least three goals were set. Goals were considered realistic when

the hypothetical mean GAS score at follow-up was 50 (SD = 10, range 40-60). Feasibility was considered good when the patient or the caregiver was involved in setting the goals. Feasibility was considered good when a blinded assessor has been involved in assessing the follow-up GAS scores. Domains that were included in the goals had to cover at least two of the following domains: cognition, instrumental activities of daily life (IADL)/self-care, mood, behaviour or mobility.

Results

Ten studies met the inclusion criteria, of which nine studies were from Rockwood and co-workers [19, 32-40]. Two independent raters checked the titles and abstracts. The third rater was not consulted, as there were no discrepancies between rater one and two. Table 2 presents the main characteristics and the clinimetric aspects of GAS of the selected studies.

Study characteristics
Four of the ten studies were samples with dementia patients [32-34, 37]. In the study of Gordon et al. [19], 77% of the sample with nursing-home patients had dementia. The characteristics of the remaining 33% in this study were not described in further detail. Rockwood et al. [36], 47% had dementia and the other patients in this study all showed at least a moderate degree of functional impairment. Stolee et al. [38] provided data of a sample in which 13% had a primary diagnosis of dementia, the remaining patients had stroke, depression, or physical problems like hip fracture or congestive hearth failure. The remaining three studies [35, 39, 40] did not provide details about numbers and specific diagnosis of the patients with cognitive disorders. The mean age of the participants was 79 years, range 51-96 (N = 846, 551 female). Most studies had a prospective, descriptive study design (N = 6). Other designs used were randomized controlled trials (RCT) (N = 3). There was one retrospective, descriptive study. All studies used GAS as a primary outcome measure for an intervention. Seven studies described some form of specialized care and three studies were medical trials. Sample sizes ranged from 10 to 170 participants.

Clinimetric aspects
Table 3 shows a comparison of the clinimetric aspects of GAS.

Responsiveness
Responsiveness was examined in eight studies. Table 2 shows the details for every study. Six of these eight studies showed a medium to large ES (0.6-4.9), SRM (0.2-1.7) or RE (0.5-100) [19, 32, 35, 36, 39, 40]; or a significant Norman's ANOVA (0.3 to 0.8) [32, 35, 39]; or a paired sample t-test for analyzing the difference in assessment and the follow-up GAS scores (t = 2.9, p = .02) [32]. One study had mixed results, i.e. small RE but large ES [37], and one study showed small SRMs [33].

Table 2. Study and participant characteristics and clinimetric aspects of GAS used in psychogeriatric patients

Study	Design	Intervention	Sample and setting	Mean age (SD)	Responsiveness	Reliability	Content validity	Construct/ convergent validity	Major findings
Gordon et al. [19]	Prospective descriptive study	Specialized geriatric medicine consultation	53 nursing-home patients of which 41 had dementia	81 years ± 8	ES = 1.3 RE = 53.7	NA	Inferred from its use in other geriatric settings	Spearman rank correlations of change scores BI: r = 0.2 HABAM: r = 0.2 CIRS: r = -0.2 Axis 8 of BCRS: r = -0.2	GAS was the most responsive measure. GAS did not correlate well with other measures
Hartman et al. [32]	Prospective descriptive study	Program of occupational therapy and therapeutic recreation	10 male persons with Alzheimer's disease or other form of dementia	76 years	Paired t-test = 2.9, df = 9, p = 0.02 ES by Kazis = 2.3 ANOVA by Norman = 0.4	NA	NA	NA	GAS was responsive
Rockwood et al. [36]	Prospective descriptive study	Specialized geriatric inpatient wards	45 consecutive elderly patients of which 21 had dementia	81 years ± 8	ES = 4.9 RE = 4.5	Inter-rater reliability (ICC = 0.9)	NA	Pearson correlations of change scores BI: r = 0.6 FIM: r = 0.5 MMSE: r = 0.0 KADL: r = 0.5 PSMS: r = - 0.5 SQLI: r = 0.4	GAS appears to be somewhat more responsive to change than the other outcome measures and correlated moderate with the BI

Table 2. (continued)

Study	Design	Intervention	Sample and setting	Mean age (SD)	Responsiveness	Reliability	Content validity	Construct/ convergent validity	Major findings
Rockwood et al. [37]	RCT	Linopirdine, Aviva, DuPont Pharma	15 patients with Alzheimer's Disease of at least mild severity	73 years	ES = 0.6 RE = 0.5	NA	Expert panel	Pearson correlations of change scores ADAS-Cog: r = -0.5 GDS: r = -0.6 CGI: r = -0.9 MMSE: r = 0.0 PSMS: r = -0.3 IADL: r = -0.4	Results on responsiveness and construct/ convergent validity are mixed and content validity is unclear
Rockwood et al. [34]	Prospective study	A 12 month phase IV trial with Donepezil	100 community-dwelling patients with mild to moderate AD	76 years ± 8 Range 51-96	NA	NA	NA	MMSE ADAS-Cog PSMS IADL FAQ CDS CES-D CIBIC-plus Patient/caregiver: Spearman r = 0.0 to 0.6; Clinician: Spearman r = 0.0 to 0.8	With the exception of the CIBIC-plus, most correlations between other measures and GAS were low to moderate. These correlations increased slightly over time
Rockwood et al. [35]	RCT	Inter-disciplinary Mobile Geriatric Assessment Team (MGAT)	165 rural, community-dwelling, geriatric frail elderly patients	82 years ± 7	ES = 2.9 to 4.4, Cohen's d SRM = 0.8 to 1.1 RE = 29.9 to >100 Norman's responsiveness statistics = 0.3 to 0.6	NA	NA	NA	GAS was more responsive than any of the standard tests

Table 2. (continued)

Study	Design	Intervention	Sample and setting	Mean age (SD)	Responsiveness	Reliability	Content validity	Construct/convergent validity	Major findings
Rockwood et al. [33]	RCT	Galantamine	130 patients with mild to moderate Alzheimer's Disease	78 years ± 8 Range 51-94	SRM patient-caregiver = 0.4/0.2 for SRM clinician = 0.4/0.4	NA	NA	NA	Responsiveness of GAS was low for both clinicians and patients/caregivers
Stolee et al. [38]	Prospective descriptive study	Geriatric Assessment Unit (GAU)	15 patients who were consecutively admitted to a geriatric service of which 2 were demented	79 years (10) Range 65-94	NA	Inter-rater reliability (ICC = 0.9 for change, 0.9 for discharge, r = 0.9 for change, 0.9 for discharge)	Position statements	Pearson correlations of change scores BI: r = 0.9 GCOR: r = 0.8	GAS showed high content and construct/ convergent validity and good inter-rater reliability
Stolee et al. [39]	Prospective descriptive study	Geriatric rehabilitation unit	170 consecutively admitted geriatric elderly patients	81 years ± 7 Range 61-96	ES = 3.5 SRM = 1.7 RE = 3.1 ANOVA (Norman) = 0.8	Inter-rater reliability (ICC = 0.9 and 0.9)	Expert panel	Pearson correlations of change and follow-up scores Change scores BI: r = 0.6 OARS IADL: r = 0.5 MMSE: r = 0.2 NHP (overall): r = 0.0 Follow-up scores BI: r = 0.7 OARS IADL: r = 0.5 MMSE: r = 0.3 GCOR: r = 0.7 NHP (overall): r = -0.2	GAS was responsive and showed high inter-rater reliability. Content validity and construct/ convergent validity is unclear

Table 2. (continued)

Study	Design	Intervention	Sample and setting	Mean age (SD)	Responsiveness	Reliability	Content validity	Construct/convergent validity	Major findings
Yip et al. [40]	Retrospective descriptive study	Geriatric Assessment and Rehabilitation Unit (GARU)	143 geriatric patients	77 years ± 8	SRM = 1.6 RE = 3.2	NA	Position statements	SMMSE: Spearman r = 0.1 BI KADL OARS IADL Spearman r between 0.4* and 0.5*	GAS was responsive and showed good content validity. Correlations with other measures were low

*p < 0.05

Abbreviations of outcome measures in alphabetical order: AD: Alzheimer's disease; ADAS-Cog: Alzheimer's Disease Assessment Scale-Cognitive Section [7]; BCRS: Brief Cognitive Rating Scale [53]; BI: Barthel Index [54]; CDS: Cornell Depression Scale [55]; CES-D: Centre for Epidemiologic Studies Depression Scale [56]; CGI: Clinical Global Impression; CIBIC-plus: Clinician's Interview-Based Impression of Change-Plus caregiver input [49]; CIRS: Cumulative Illness Rating Scale [57]; FAQ: Functional Activities Questionnaire [58]; FIM: Functional Independence Measure [59]; GAS: Goal Attainment Scaling [9]; GCOR: Global Clinical Outcome Rating; GDS: Global Deterioration Scale [60]; HABAM: Hierarchical Assessment of Balance and Mobility [61]; IADL: Instrumental Activities of Daily Living [62]; KADL: Katz Activities of Daily Living Index [63]; MMSE: Mini-Mental State Examination [64]; NHP: Nottingham Health Profile [65]; OARS IADL: IADL subscale of the Older Americans Resources and Services [66]; PSMS: Physical Self-Maintenance Scale [62]; SMMSE: Standardized Mini-Mental State Examination [67]; SQLI: Spitzer Quality of Life Index [68].

Other abbreviations in alphabetical order: ES: effect size; NA: not an aim of the study; RCT: Randomized Controlled Trial; RE: relative efficiency; SRM: standardized response mean.

Table 3. Comparison of the clinimetric aspects of Goal Attainment Scaling (GAS) used in psychogeriatric patient

Study	Responsiveness[a]	Reliability[b]	Content validity[c]	Construct/convergent validity[d]
Gordon et al. [19]	+	NA	?	-
Hartman et al. [32]	+	NA	NA	NA
Rockwood et al. [36]	+	+	NA	+/-
Rockwood et al. [37]	+/-	NA	+	+/- , +
Rockwood et al. [34]	NA	NA	NA	- , +/- , +
Rockwood et al. [35]	+	NA	NA	NA
Rockwood et al. [33]	-	NA	NA	NA
Stolee et al. [38]	NA	+	+	+
Stolee et al. [39]	+	+	+	- , +/- , +
Yip et al. [40]	+	NA	+	-

[a] Changes between baseline and follow-up of GAS: - = NS, +/- = mixed results, + = S (p < 0.05, effect size (ES) > 0.5, standardized response mean (SRM) > 0.5, relative efficiency > 1), ? = unclear, NA = not an aim of the study

[b] Pearson correlation coefficient or intraclass correlation coefficient (ICC): - = < 0.7, +/- = between 0.7 and 0.79, + = > 0.8, ? = unclear , NA = not an aim of the study

[c] An expert's panel reviewed the selected domains for GAS: - = reviewed as insufficient, + = reviewed as sufficient, ? = unclear, NA = not an aim of the study

[d] Correlation coefficient (Pearson or Spearman) between GAS and secondary outcome measures: - = low correlation (< 0.5), +/- = moderate correlation (between 0.5-0.69), + = high correlation (> 0.69), ? = unclear, NA = not an aim of the study

Reliability
Only three studies examined the inter-rater reliability (IRR). The ICC was used by Rockwood et al. [36], Stolee et al. [38], and Stolee et al. [39], and was 0.9. Stolee et al. [38] used also the Pearson's correlation coefficient (0.9).

Content validity
Five studies examined the content validity. Stolee et al. [38] and Yip et al. [40] assessed content validity by comparing identified goal areas with a list of assessment areas derived from position statements describing the essential components of geriatric assessment [41-43]. In Stolee et al. [38], two geriatricians independently identified 82% (71 out of 87) of the goal areas; the remainder were set by consensus. Both geriatricians identified the same chief goals in 93% (14 out of 15) of the patients and were able to agree on the relative importance of goals. In Yip et al. [40] all recommended domains were addressed. Stolee et al. [39] and Rockwood et al. [37] examined content validity through content analysis of identified goal areas. In Stolee et al. [39] clinicians from the geriatric rehabilitation unit reviewed the categorizations and suggested modifications. In Rockwood et al. [37] a panel with expertise in anti-dementia drug trials, consisting of 3 neurologists, 2 geriatricians, an epidemiologist and a biostatistician reviewed the methodology and suggested modifications. Gordon et al. [19] inferred content validity from the use of GAS in other geriatric settings, but it was unclear if the domains found in their study were related to the other settings as no results were shown.

Construct/convergent validity
Seven studies examined the construct/convergent validity of GAS using several secondary outcome measures (Table 2). Spearman correlations were calculated by Gordon et al. [19], Rockwood et al. [34], Yip et al. [40] and ranged from 0.0-0.8. Pearson's correlation coefficients were calculated by Rockwood et al. [36, 37], and Stolee et al. [38, 39] and ranged from 0.0-0.9.

Feasibility
Table 4 outlines the details about the feasibility of GAS of the selected studies and Table 5 shows the comparison of the details about the feasibility of GAS.

Time required for setting the goals
Five studies described the time required for setting the goals. Gordon et al. [19], Hartman et al. [32], Stolee et al. [38] and Yip et al. [40] described they were able to set goals in less than 30 minutes. The study by Rockwood et al. [37] was the only one in which it was not possible to set goals in less than half an hour.

Number of goals
All ten studies reported the number of goals set in their study. The exact number of goals for each study varied from 1 to 9. Further details are shown in Table 4. In the study of Rockwood et al. [34], physicians set fewer goals than patients/caregivers (3 and 9 respectively). The same pattern is shown in Rockwood et al. [33]. Rockwood et al. [35] reported that the care-

Table 4. Feasibility of GAS used in psychogeriatric patients

Study	Time required for setting goals	Number of goals Mean (SD) Range	Realistic goals* Mean (SD)	Person(s) involved in goal setting	Person(s) involved in assessing the follow-up GAS scores	Main domains
Gordon et al. [19]	Less than 10 minutes	1.7 1-5	46 (6.9)	Family physician Two Geriatricians Nurse	Not reported	Medical problems Behaviour Ethics
Hartman et al. [32]	An average of 20 minutes	1.4 1-2	53.4 (16.0)	Occupational therapist Therapeutic recreational specialist Family members	Not reported	IADL Therapeutic recreation
Rockwood et al. [36]	Not reported	5	51 (6)	Team members	Blinded primary care nurse	Not reported
Rockwood et al. [37]	2 hours over 3 initial patients visits	3.7 2-6	36.4 (15.6)	Physicians Patient Caregiver	Not reported	Self care Behaviour Cognition Leisure activities
Rockwood et al. [34]	Not reported	Physician 3 (1) Patient/caregiver 9 (3)	Not reported	Physician Nurse Patient Caregiver Field researcher	Clinical team Field researchers	Cognition Functioning Behaviour Leisure Social interaction

Table 4. (continued)

Study	Time required for setting goals	Number of goals Mean (SD) Range	Realistic goals* Mean (SD)	Person(s) involved in goal setting	Person(s) involved in assessing the follow-up GAS scores	Main domains
Rockwood et al. [35]	Not reported	Usual care 5.2 (1.9) Intervention 5.7 (2.1)	40.7 (5.3) (control) 48.0 (7.1) (intervention)	Geriatric nurse assessor Team members Geriatrician Patient	Blinded independent physician	Self care Mobility Incontinence
Rockwood et al. [33]	Not reported	Clinicians set fewer goals than patients Median = 3 Maximum = 6	No exact numbers mentioned	Physician Patient Caregiver Health professional	Physician Health professional	Cognition Functioning Leisure Social activities
Stolee et al. [38]	Approximately 30 minutes without experience; with experience 15-20 minutes	5.8 5-8	Geriatrician: 48.6 (8.5) Primary care nurse: 46.6 (7.5)	Two geriatricians	Geriatrician Primary care nurse	Mobility Internal medicine Future care

GAS: Goal Attainment Scaling; IADL: Instrumental Activities of Daily Living

*A follow up GAS score between 40-60 indicates that all predefined goals met the expected outcome at follow up

Table 5. Comparison of the feasibility of Goal Attainment Scaling (GAS) used in psychogeriatric patients

Study	Time required for setting goals[a]	Number of goals[b]	Realistic goals[c]	Patient and/or caregiver involved in goal setting[d]	Blinded assessor involved in assessing the follow-up GAS scores[e]	Main domains[f]
Gordon et al. [19]	+	-	+	-	NR	-
Hartman et al. [32]	+	-	+	+	NR	-
Rockwood et al. [36]	NR	+	+	?	+	NR
Rockwood et al. [37]	_	+	-	+	NR	+
Rockwood et al. [34]	NR	+	NR	+	_	+
Rockwood et al. [35]	NR	+	+/-	+	+	+
Rockwood et al. [33]	NR	+	NR	+	_	+
Stolee et al. [38]	+	+	+	?	-	-
Stolee et al. [39]	NR	+	+	?	NR	+
Yip et al. [40]	+	+	+	?	-	+

[a] Time required setting the goals: - = >30 minutes, + = ≤ 30 minutes, NR = Not reported

[b] The number of goals: - = < 3

[c] Realistic goals: - = follow up GAS score < 40 or >60, + = follow up GAS score 40 ≥ x ≤60, +/- = mixed results, NR = Not reported

[d] Patient and caregiver involved in setting the goals: - = absence of patient or caregiver, + = presence of patient or caregiver, ? = role of patient or caregiver unclear

[e] Blinded assessor involved in assessing the follow up GAS score: - = absence of blinded assessor, + = presence of blinded assessor

[f] Main domains: - = ≤ 1 domain out of 5*, + = 2 or more domains out of 5*, NR = Not reported
*Cognition, IADL, mood, behaviour or mobility.

as-usual group and the intervention group set a similar number of goals (5 and 6 respectively). Two out of ten studies had a mean number of goals set below 3. Eight studies showed a mean number of goals set above 3.

To what extent were goals realistic?

Whether goals were realistic was examined in eight studies. Six studies had GAS follow-up scores between 40 and 60 [19, 32, 36, 38-40]. One study showed mixed results [35] and one study had GAS follow-up scores below 40, indicating that the goals set were too difficult [37]. No study had a GAS follow-up score above 60, which would have indicated that the goals were too easy.

Patient and caregiver involved in setting the goals

All ten studies described who were involved in the goal-setting process. Seven studies involved the patient and/or the caregiver in the goal-setting process [32-34, 37, 39, 40]. Two studies did not involve the patient and/or the caregiver because the clinician set the goals [19, 36]. In Stolee [38] two geriatricians independently assessed patients comprehensively. Then goals were set for each patient by consensus between the two geriatricians. However, it is not clear to what extent patients were involved in the goal-setting process here.

A blinded assessor involved in assessing the follow-up GAS scores

Six studies reported who assessed the follow-up GAS scores. Two studies used a blinded assessor [35, 36]. Four studies described that the persons who assessed the follow-up GAS scores were the same persons who set the goals. Therefore these assessors were not blinded [33, 34, 38, 40].

Domains that were included in the goals

Nine studies described the domains that were covered by the goals. The number of domains varied from 2 to 12. Six studies reported two or more of the following five domains: cognition, IADL/self care, mood, behaviour or mobility [33-35, 37, 40, 44]. Three studies set goals in less than two predefined domains [19, 32, 38].

Discussion

We reviewed the literature on the clinimetric aspects and the feasibility of GAS when used as an outcome measure for psychogeriatric patients with cognitive disorders in order to study the applicability of GAS in this population. The literature search resulted in the identification of ten studies of which most were performed by Rockwood and co-workers.

Although one study [19] argued that the presence of dementia did not make it feasible to involve the patients in setting their own goals, our review shows that the involvement of the patient with dementia and/or caregiver actually is feasible. Moreover, GAS can cover multiple domains relevant to the psychogeriatric population such as cognition, daily life activities and mood.

Six of the ten studies did not provide full details of their sample. We could not rule out that the conclusions in these studies are based on non-cognitively impaired patients. Therefore, we were not able to make firm conclusions about those clinimetric aspects and feasibility when the majority of studies investigating these aspects were those with heterogeneous samples. This was the case for all the clinimetric aspects, time required for setting goals, realistic goals and the involvement of a blinded assessor in assessing the follow-up GAS score. Reliability was positive in all studies that investigated this clinimetric aspect, but since these were just three studies and all had a heterogeneous sample, results must be interpreted with caution.

Strengths and limitations of the studies in the review

Nine of the ten studies reviewed were performed by one centre (Rockwood and co-workers). Very few studies outside this research group investigated GAS in psychogeriatric patients with cognitive disorders [45]. None of them described clinimetric aspects and feasibility of GAS. The absence of independent studies raises questions about the use of GAS in cognitively impaired psychogeriatric patients and may be a reason why this promising approach is not studied and used more widespread in this population. Another explanation for the fact that we did not find studies from other research groups could be publication bias: the latter might have found negative results with regard to the use of GAS in psychogeriatric patients with cognitive disorders which they were not willing to publish.

With regard to the methodology of the reviewed studies we noted the following. First of all, the intraclass correlation coefficient (ICC) for measuring inter-rater reliability (IRR) is preferred over Pearson's correlation when the sample size is small (< 15). As Pearson's correlation makes no assumptions about rater means, a t-test of the significance of the correlation is necessary to see if inter-rater means differ [46]. All three studies assessing IRR used ICC. However, the range of sample size varied between N = 15 and N = 170, so ICC was not the preferred statistic. Secondly, large sample sizes are generally preferred in order to find reliable results [31]. However, with a highly responsive instrument - as has turned out to be the case with GAS - smaller sample sizes are sufficient to demonstrate the efficacy of an intervention [36]. Therefore, the small sample sizes of some of the reviewed studies do not form a limitation.

Recommendations for future use of GAS

We found that mixed results were obtained when evaluating the validity of GAS. This can be due to the fact that the secondary outcome measures used in the studies are fixed in terms of the scoring items and cannot vary across individuals, while GAS can. Therefore, correlations between GAS and standardized tests are generally expected to be low [47]. Rockwood et al. [36] mentioned that there is no gold standard (criterion) for measuring clinically important change; Donnelly & Carswell [48] underscore the idiosyncratic nature of GAS, both making it hard to validate it. Although several authors have pointed to the fact that validation of GAS may be problematic because of its highly personalized approach, the claim that it is a highly ecologically valid instrument still needs to be proven. Using a blinded assessor for measuring the follow-up GAS scores can enhance cross validity; this also ensures an unbiased outcome of interventions. In addition, comparative measures, such as the clinical global impression

(CGI) and the Clinician's Interview-Based Impression of Change-Plus caregiver input (CIBIC-plus) [49] are also sensitive to clinical change and can thus be used to validate GAS. Although these latter instruments do not take into account the specific needs of the patient and/or caregiver, their aim is to measure clinical change relevant after an intervention.

We have selected studies based on the criterion that the original GAS was used and only small deviations were accepted. Two studies [34, 40] described small adaptations of which the adaptation made by Rockwood et al. [34] to set the baseline level on '0' instead of '-2' or '-1' is recommended. According to them, this modification incorporates more levels of decline, which seems preferable, given that deterioration is likely in patients with dementia. When shifting the baseline status to '0', the GAS formula must be adapted as well, since the score of 50 in the GAS formula is not the hypothetical mean for the follow-up score anymore. Yip et al. [40] developed a standard menu of goal areas, developed by the geriatric assessment and rehabilitation unit team, from which relevant goals could be selected. However, the expected outcome was a non-numerical scoring which is qualitative and therefore difficult to interpret.

As most studies reviewed were performed by one centre (Rockwood and co-workers), a more widespread use by other independent research groups may produce more data on the applicability of GAS.

Clinical implications

As dementia is characterised by a complex multidimensional nature, it is important that outcome measures reflect this multidimensionality. Our review showed that GAS is an outcome measure that can cover several relevant domains like cognition, behaviour and activities of daily life. Therefore, it is a relevant method for measuring outcomes in the psychogeriatric field. Outcome measures also need to take the progressive nature of dementia into account; as GAS can be adapted at any time, it is useful in dementia.

Furthermore, depending on the aim for which GAS is used, both process and outcome goals can be set. However, for intervention research it is important not to adapt the GAS follow-up guide because a baseline and predefined follow-up measure are needed.

Part of the definition of dementia is that cognitive impairments lead to interference with daily life activities. Traditional measures for dementia often do not reflect this interference [50], GAS, however, does. Outcome measures seldom focus on individual needs relevant for the individual patient and the caregiver, focusing instead on general concepts, like mood and fatigue, with norms based on groups. As we showed in our review, many studies actively involved patients and/or caregivers in the goal-setting process, thereby ensuring that relevant goals were selected. In other words, the fact that GAS takes individual preferences into account is a safeguard for clinical relevance of this outcome measure. Patients with dementia may lack insight into their problems, in which case a caregiver can be asked to help with setting the goals. Caregiver bias can be present here, but since dementia also affects the caregiver [51], caregiver's goals are highly relevant as well. Our conclusion is that caregivers should not only be asked to help setting goals for their partners with dementia, but should also set goals that are relevant for themselves.

GAS has a special position among regular outcome measures in the field of cognitive disorders and it is a useful complementary instrument [52] providing information that cannot be obtained by traditional standard outcome measures like the ADAS-cog. Although it can be complicated to involve a whole multidisciplinary team in setting goals for each patient, this approach provides a lot of relevant information that will remain hidden when a mono-disciplinary approach is used. We realise that the GAS constructing time of 30 minutes is relatively long compared to a quick MMSE, but the extra information provided by GAS justifies the time involved in training and constructing the GAS follow-up guides.

Since GAS is developed to fit the individual needs of the patient and is constructed separately for each patient, it bridges possible language and culture barriers. As there is no uniform description of how to use GAS in a cognitively impaired population, training and practice to set goals in this particular population is recommended and practice guidelines for standardizing GAS in a cognitively impaired population would be desirable.

Conclusion

GAS proved to be useful on important aspects of an outcome measure for psychogeriatric patients with cognitive disorders. Since other relevant aspects showed mixed results and the number of studies investigating the use of GAS in psychogeriatric patients with cognitive disorders is small, the evidence is not strong enough yet to state that GAS is an applicable outcome measure in this population. We do think that GAS has potential value and although the use of GAS requires training, which is time-consuming, we believe that this investment is worthwhile because GAS is a unique example of an instrument able to reflect the multidimensionality of dementia and other psychogeriatric conditions, including interference with daily life activities, for both patient and caregiver.

References

1. American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision. Washington, DC, American Psychiatric Association 2000.
2. Beswick AD, Rees K, Dieppe P, Ayis S, Gooberman-Hill R, Horwood J, et al. Complex interventions to improve physical function and maintain independent living in elderly people: a systematic review and meta-analysis. Lancet 2008;371(9614):725-735.
3. Boothroyd RA, Banks SM. Assessing outcomes in individually-tailored interventions. Lancet 2006;367(9513):801-802.
4. Katona C, Livingston G, Cooper C, Ames D, Brodaty H, Chiu E. International Psychogeriatric Association consensus statement on defining and measuring treatment benefits in dementia. Int Psychogeriatr 2007;19(3):345-354.
5. Gauthier S. Cholinesterase inhibitors in late-stage Alzheimer's disease. Lancet Neurol 2006;5(6):468-469.
6. Hogan DB. Donepezil for severe Alzheimer's disease. Lancet 2006;367(9516):1031-1032.
7. Rosen WG, Mohs RC, Davis KL. A new rating scale for Alzheimer's disease. Am J Psychiatry 1984;141(11):1356-1364.
8. Harrison JE. Measuring cognitive change in Alzheimer's disease clinical drug trials. J Nutr Health Aging 2007;11(4):327-329.
9. Kiresuk T, Sherman RE. Goal Attainment Scaling: a general method for evaluating comprehensive community mental health programs. Community Ment Health J 1968;4(6):443-453.
10. Malec JF. Goal Attainment Scaling in Rehabilitation. Neuropsychol Rehabil 1999;9(3/4):253-275.
11. Joyce BM, Rockwood KJ, Mate-Kole CC. Use of goal attainment scaling in brain injury in a rehabilitation hospital. Am J Phys Med Rehabil 1994;73(1):10-14.
12. Fisher K. Assessing clinically meaningful change following a programme for managing chronic pain. Clin Rehabil 2008;22(3):252-259.
13. World Health Organization. ICIDH-2: International classification of functioning, disability, and health-Final Draft. Madrid: WHO 2001.
14. Bravo G, Dubois MF, Roy PM. Using goal attainment scaling to improve the quality of long-term care: a group-randomized trial. Int J Qual Health Care 2005;17(6):511-519.
15. Rockwood K, Fay S, Jarrett P, Asp E. Effect of galantamine on verbal repetition in AD: a secondary analysis of the VISTA trial. Neurology 2007;68(14):1116-1121.
16. Yau DCN, Lam MCC, Lung MKW, Chow AYH, Yung NSY, Chui TY, et al. Outcome-focused nursing practice in geriatric day-rehabilitation service: facilitating elderly patients to achieve therapeutic goals. The Hong Kong Nursing Journal 2002;38(4):33-38.
17. AD2000 Collaborative Group. Long-term donepezil treatment in 565 patients with Alzheimer's disease (AD2000): randomised double blind trial. Lancet 2004;363:2105-2115.

18. Malec JF, Smigielski JS, DePompolo RW. Goal attainment scaling and outcome measurement in postacute brain injury rehabilitation. Arch Phys Med Rehabil 1991;72(2):138-143.

19. Gordon JE, Powell C, Rockwood K. Goal attainment scaling as a measure of clinically important change in nursing-home patients. Age Ageing 1999;28(3):275-281.

20. Van Tulder MW, Assendelft WJ, Koes BW, Bouter LM. Method guidelines for systematic reviews in the Cochrane Collaboration Back Review Group for Spinal Disorders. Spine 1997;22(20):2323-2330.

21. Terwee CB, Dekker FW, Wiersinga WM, Prummel MF, Bossuyt PM. On assessing responsiveness of health-related quality of life instruments: guidelines for instrument evaluation. Qual Life Res 2003;12(4):349-362.

22. Badia X, Podzamczer D, Casado A, Lopez-Lavid C, Garcia M. Evaluating changes in health status in HIV-infected patients: Medical Outcomes Study-HIV and Multidimensional Quality of Life-HIV quality of life questionnaires. Spanish MOS-HIV and MQOL-HIV Validation Group. Aids 2000;14(10):1439-1447.

23. Norman GR. Issues in the use of change scores in randomized trials. J Clin Epidemiol 1989;42:1097-1105.

24. Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. Med Care 1989;27:S178-S189.

25. Liang MH, Fossel AH, Larson MG. Comparisons of five health status instruments for orthopedic evaluation. Med Care 1990;28:632-642.

26. Liang MH, Larson MG, Cullen KE. Comparative measurement efficiency and sensitivity of five health status instruments for arthritis research. Arthritis Rheum 1985;28:542-547.

27. Cohen J. Statistical Power Analysis for the Behavioral Sciences. Hillsdale, NJ: Lawrence Erlbaum 1988.

28. Faraone SV. Understanding Effect Size: How It's Measured and What It Means. Medscape Psychiatry & Mental Health 2008.

29. Angst F, Goldhahn J, Drerup S, Aeschlimann A, Schwyzer HK, Simmen BR. Responsiveness of six outcome assessment instruments in total shoulder arthroplasty. Arthritis Rheum 2008;59(3):391-398.

30. Heiman GW. Research Methods in Psychology. Boston: Houghton Mifflin Company 1999.

31. Hinkle DE, Wiersma W, Jurs SG. Applied statistics for the behavioral sciences. New York: Houghton Mifflin Company 1998.

32. Hartman D, Borrie MJ, Davison E, Stolee P. Use of Goal Attainment Scaling in a Dementia Special Care Unit. Am J Alzheimers Dis Other Demen 1997;12(3):111-116.

33. Rockwood K, Fay S, Song X, MacKnight C, Gorman M. Attainment of treatment goals by people with Alzheimer's disease receiving galantamine: a randomized controlled trial. CMAJ 2006;174(8):1099-1105.

34. Rockwood K, Graham JE, Fay S. Goal setting and attainment in Alzheimer's disease patients treated with donepezil. J Neurol Neurosurg Psychiatry 2002;73(5):500-507.

35. Rockwood K, Howlett S, Stadnyk K, Carver D, Powell C, Stolee P. Responsiveness of goal attainment scaling in a randomized controlled trial of comprehensive geriatric assessment. J Clin Epidemiol 2003;56(8):736-743.

36. Rockwood K, Stolee P, Fox RA. Use of goal attainment scaling in measuring clinically important change in the frail elderly. J Clin Epidemiol 1993;46(10):1113-1118.

37. Rockwood K, Stolee P, Howard K, Mallery L. Use of Goal Attainment Scaling to measure treatment effects in an anti-dementia drug trial. Neuroepidemiology 1996;15(6):330-338.

38. Stolee P, Rockwood K, Fox RA, Streiner DL. The use of goal attainment scaling in a geriatric care setting. J Am Geriatr Soc 1992;40(6):574-578.

39. Stolee P, Stadnyk K, Myers AM, Rockwood K. An individualized approach to outcome measurement in geriatric rehabilitation. J Gerontol A Biol Sci Med Sci 1999;54(12): M641-647.

40. Yip AM, Gorman MC, Stadnyk K, Mills WG, MacPherson KM, Rockwood K. A standardized menu for Goal Attainment Scaling in the care of frail elders. Gerontologist 1998;38(6):735-742.

41. Fox RA, Clarfield AM, Hogan DB. Competencies required for the practice of geriatric medicine as a consultant physician CMAJ 1989;141:1045-1048.

42. Rubenstein LV, Calkins DR, Greenfield S, Jette AM, Meenan RF, Nevins MA, et al. Health status assessment for elderly patients. Report of the Society of General Internal Medicine Task Force on Helath Assessment. J Am Geriatr Soc 1989;37:562-569.

43. Rubenstein LZ. Comprehensive geriatric assessment Ann Intern Med 1988;108:724-727.

44. Stolee P, Zaza C, Pedlar A, Myers AM. Clinical experience with Goal Attainment Scaling in geriatric care. J Aging Health 1999;11(1):96-124.

45. Asp E, Cloutier F, Fay S, Cook C, Robertson ML, Fisk J, et al. Verbal repetition in patients with Alzheimer's disease who receive donepezil. Int J Geriatr Psychiatry 2006; 21(5):426-431.

46. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 1979;86:420-428.

47. Ottenbacher KJ, Cusick A. Discriminative versus evaluative assessment: some observations on goal attainment scaling. Am J Occup Ther 1993;47:349-354.

48. Donnelly C, Carswell A. Individualized outcome measures: a review of the literature. Can J Occup Ther 2002;69(2):84-94.

49. Knopman DS, Knapp MJ, Gracon SI, Davis CS. The Clinician Interview-Based Impression (CIBI): a clinician's global change rating scale in Alzheimer's disease. Neurology 1994;44(12):2315-2321.

50. Bouwens SFM, van Heugten CM, Aalten P, Wolfs CAG, Baarends EM, van Menxel DAJ, et al. Relationship between measures of dementia severity and observation of daily life functioning as measured with the Assessment of Motor and Process Skills (AMPS). Dement Geriatr Cogn Disord 2008;25:81-87.

51. De Vugt ME, Stevens F, Aalten P, Lousberg R, Jaspers N, Winkens I, et al. Behavioural disturbances in dementia patients and quality of the marital relationship. Int J Geriatr Psychiatry 2003;18(2):149-154.

52. Ponsford J, Olver J, Nelms R, Curran C, Ponsford M. Outcome measurement in an inpatient and outpatient traumatic brain injury rehabilitation programme. Neuropsychol Rehabil 1999;9(3/4):517-534.

53. Reisberg B, Ferris SH. Brief Cognitive Rating Scale (BCRS). Psychopharmacol Bull 1988;24(4):629-636.

54. Mahoney FI, Barthel DW. Functional Evaluation: the Barthel Index. Md State Med J 1965;14:61-65.

55. Alexopoulos GS, Abrams RC, Young RC, Shamoian CA. Cornell Scale for Depression in Dementia. Biol Psychiatry 1988;23(3):271-284.

56. Radloff LS. The CES-D scale: a self-report depression scale for research in the general population. Applied Psychological Measurement 1977;1:385-401.

57. Parmelee PA, Thuras PD, Katz IR, Lawton MP. Validation of the Cumulative Illness Rating Scale in a geriatric residential population. J Am Geriatr Soc 1995;43(2):130-137.

58. Pfeffer RI, Kurosaki TT, Harrach CH, Chance JM, Filos S. Measurement of functional activities in older adults in the community. J Gerontol 1982;37:232-329.

59. Guide for the Use of the Functional Independence Measure. Uniform Data System for Medical Rehabilitation. Bufallo: State University of New York at Buffallo; 1987.

60. Reisberg B, Ferris SH, de Leon MJ, Crook T. The Global Deterioration Scale for assessment of primary degenerative dementia. Am J Psychiatry 1982;139(9):1136-1139.

61. MacKnight C, Rockwood K. A Hierarchical Assessment of Balance and Mobility. Age Ageing 1995;24(2):126-130.

62. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. Gerontologist 1969;9(3):179-186.

63. Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW. Studies of Illness in the Aged. the Index of ADL: a Standardized Measure of Biological and Psychosocial Function. J Am Med Assoc 1963;185:914-919.

64. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975;12(3):189-198.

65. Hunt SM, McEwan J, McKenna SP. Measuring Health Status. London, UK: Croom Helm 1986.

66. George LK, Fillenbaum GG. OARS methodology. A decade of experience in geriatric assessment. J Am Geriatr Soc 1985;33:607-615.

67. Molloy DW, Alemayehu E, Roberts R. Reliability of a Standardized Mini-Mental State Examination compared with the traditional Mini-Mental State Examination. Am J Psychiatry 1991;148(1):102-105.

68. Spitzer WO, Dobson AJ, Hall J, Chesterman E, Levi J, Shepherd R, et al. Measuring the quality of life of cancer patients: a concise QL-index for use by physicians. J Chronic Dis 1981;34(12):585-597.

Appendix 1. Formula to calculate the Goal Attainment Scaling (GAS) scores

$$\text{GAS score} = \frac{50 + 10\sum(w_i x_i)}{\sqrt{\left(0.7\sum w_i^2 + 0.3\left(\sum w_i\right)^2\right)}}$$

where xi = the attainment level and wi = the weight assigned to the goal area. Because the expected outcome/programme goal is assigned the score of 0, if all goals are met, the outcome score would be 50.

Appendix 2. Details of literature search

Controlled vocabulary words (combinations of these terms)

Category	Terms
Goal attainment scaling	goal attainment scaling
Cognitive disorders	Alzheimer's disease, delirium, dementia, amnestic cognitive disorders, dementia, multi infarct dementia, vascular dementia, cognitive rehabilitation, cognition disorders
Clinimetric aspects	psychometrics, psychometry, interrater reliability, reliability and validity, reliability, intrarater reliability, test-retest reliability, validity, face validity, qualitative validity, content validity, consensual validity, predictive validity, criterion related validity, concurrent validity, external validity, discriminant validity, internal validity, construct validity, sensitivity to change, sensitivity and specificity, reproducibility of results

Free text words (combinations of these terms and in combination with controlled vocabulary words)

Category	Terms
Goal attainment scaling	"goal attaining scale", "goal attainment scale", "goal attainment scaling", "goal attainment score", "goal attainment scoring", "goal attainment scores"
Cognitive disorders	cognit*, "cognitive rehabilitation", "cognition disorder", dement*, "Alzheimer's disease", "Alzheimer, dementia"
Clinimetric aspects	psychometrics, clinimetrics, "reproducibility of results", reliability, "interrater reliability", "intrarater reliability", "test-retest reliability", validity, "face validity", "content validity", "predictive validity", "criterion related validity", "concurrent validity", "external validity", "discriminant validity", "internal validity", "construct validity", responsiveness, sensitivity, "sensitivity to change", specificity, feasibility

Chapter 8

Effectiveness of a post-acute outpatient
community re-entry cognitive rehabilitation
programme for patients with acquired brain injury

Abstract

Objective
To examine the effectiveness of a post-acute outpatient community re-entry cognitive rehabilitation programme for patients with acquired brain injury and their relatives.

Design
Prospective cohort study.

Participants
Twenty-seven patients with acquired brain injury (30% stroke; 52% male) with a mean age of 49.5 (SD 9.2) years. Time since injury was 1.9 years (SD 2.0). And 25 relatives with a mean age of 48.8 (SD 8.8) years.

Intervention
A group programme with seven participants per group, 15 weekly sessions of 2.5 hours and a booster session six weeks after the last session. Group sessions consisted of cognitive strategy training, social skills training, and psycho-education. Patients also received homework. Relatives were invited twice.

Methods
Repeated measurements: prior to treatment (baseline, T0); directly after treatment (T1, 21 weeks); and at follow-up (T2, 45 weeks). Primary outcome measures: cognitive failures (CFQ); quality of life (SA-SIP); and individualized goals (GAS).

Results
There were no significant differences in CFQ; the SA-SIP physical functioning deteriorated significantly at first (T1), but improved at follow-up (T2). At T2 the level of social activities improved and patients reported less fatigue ($p < 0.05$). Patients did improve significantly on individual goals ($p < 0.05$) between T0 and T1. The level of attainment did not change between T1 and T2. Goals were mostly set in the cognitive and behavioural domains.

Conclusions
The programme had a positive effect on the individual goals set by the patients. It seems that although patients attain their individual goals, this does not result in a higher participation level or a better quality of life.

Introduction

Neuropsychological disorders are common after acquired brain injury. Cognitive, emotional and behavioural deficits fall within the scope of neuropsychological disorders and are highly prevalent after brain injury [1-11]. For a substantial proportion of patients, return to work becomes only partially possible, or absolutely impossible [12-16]; the same is applicable for leisure activities [3]. Personality changes often account for family disruption [17, 18].

Few patients are well prepared to manage the challenges of daily life and social reintegration after their brain injury. Most patients with acquired brain injury who do not experience obvious physical consequences are not screened for cognitive and emotional deficits. Yet, even patients who are discharged to their home environment after neurorehabilitation without follow-up visits experience a range of psychosocial consequences in daily life that affect adjustment, long-term recovery and quality of life [19]. Most of these problems become obvious when patients have to return to pre-injury daily routines and try to live an independent life within the full scope of societal participation [20].

Despite a wide array of research on neuropsychological issues concerning brain injured adults, few programmes exist that address the long-term needs of brain injury survivors once they are discharged from post-acute programmes [21, 22]. Most cognitive rehabilitation programmes are offered to patients in the sub-acute phase after brain injury and are specifically designed to address only one cognitive disorder, such as attention or memory [23, 24]. In contrast, in comprehensive (holistic) cognitive rehabilitation programmes, all aspects of cognitive and emotional consequences after brain injury are incorporated. These programmes are intensive (lasting for a couple of weeks, mostly four to five days a week), costly and seem to be effective [25-30]. However, for some patients both kinds of programmes are rather too specific or too intensive for the problems they encounter.

We are referring to patients who function independently in their daily activities and for whom returning to work was (partially) possible, but who still experience difficulties. These patients run the risk of burnout, depression or social isolation. To prevent secondary problems, these patients need practical individualised strategies to compensate for cognitive deficits and to learn specific social skills to help them manage their daily life and enhance self-efficacy. Although restoration of function can occur even several years post injury, teaching compensatory strategies is nowadays considered standard treatment that is effective long after the brain injury occurred [31, 32].

Until now there have been few treatment possibilities for this specific group of patients. Some studies included patients with mild deficits and found that psychotherapy and cognitive remediation had an effect on reduction of emotional distress. However, these treatments were mostly focused on coping with the emotional consequences of brain injury [19, 33]. Other studies found that psycho-education to inform patients with mild brain injury about the consequences is effective [34-37]. So it is still unclear what kind of treatment is the best.

In their review, Malec & Basford [38] classified post-acute brain injury rehabilitation programmes in several categories. One of these is the outpatient community re-entry programme. These programmes typically focus on circumscribed rehabilitative treatment and vocational and social reintegration. In this study we evaluate the effectiveness of such

an outpatient community re-entry programme. The purpose of this programme is five-fold: 1) to help patients and caregivers to gain insight into the consequences of the brain damage; 2) to offer strategies to handle cognitive deficits in their daily lives; 3) to learn social skills in order to live with the consequences; 4) to learn how to control emotional reactions; and 5) to enhance self-efficacy. In the present study the effectiveness of the programme is determined for both the patients and their relatives directly after treatment and in the long term.

Methods

Study design

A prospective cohort study was conducted with repeated measurements, i.e. before the start of the outpatient cognitive rehabilitation programme (T0: baseline); directly after treatment (T1: 21 weeks); and 6 months after treatment (T2: 45 weeks).

Participants

All patients who had been referred to the outpatient cognitive rehabilitation programme between September 2006 and December 2007 and who met the inclusion and exclusion criteria for participation in the programme were included in the study. Brain injured patients were referred to the treatment programme by neurologists, general practitioners and rehabilitation physicians. Patients were selected for the programme by means of an interview with a psychologist and a rehabilitation physician.

The criteria for referral to the treatment programme were the following: patients had sustained an acquired brain injury at least three months earlier; patients were older than 18; patients experienced cognitive, emotional, and/or behavioural problems that interfered with daily functioning and for which information, advice and treatment were necessary to assure adaptation to these problems; patients had sufficient insight into their individual problems in order to benefit from a low-frequency treatment intensity; communication in daily life was not hindered (patients could read and write); patients' social behaviour was adequate enough for them to function in a group programme; there were no treatment goals in the sensori-motor area of functioning (any more). Hence, patients with primary psychological or psychiatric disorders, dementia, or a whiplash trauma were excluded from the programme. The fact that a patient had previously undergone (cognitive) rehabilitation was not an exclusion criterion, as mostly these earlier interventions were focused on specific cognitive deficits. All participants signed informed consent.

Inclusion criteria for participation in the research study for relatives (the informal caregivers) of patients referred to the treatment programme were: age older than 18 years, sufficient knowledge of the Dutch language in order to read and understand the questionnaires, and informed consent. Caregivers were excluded if they had a brain injury or primary psychiatric disorders.

Intervention

The aims of the outpatient cognitive rehabilitation programme are described in the introduction. With these aims, the impairments as such are not treated, but the patients and their caregivers learn to live with, manage, compensate for and accept these consequences in order to maximize participation in society.

The treatment programme was offered as a group intervention, with a maximum of seven participants per group. Two cognitive therapists conducted the group sessions. The programme consisted of 15 weekly sessions of 2,5 hours, including a half-an-hour coffee break. The first hour of each session consisted of cognitive training aimed at teaching patients strategies for attention, memory and problem-solving. Each patient formulated individual strategies. The second hour of each session consisted of social skills training. In this section the patients learned how to ask for help, or how to describe the problems encountered due to brain damage. Patients learned to become more assertive, without being too intrusive. In addition, during all sessions acceptance of the consequences and social support were discussed. Each week patients needed to perform homework. In this homework patients practised the new strategies, sometimes only in one pre-described situation, depending on the capacities of the patients. During the programme a relative of the patient was invited, and psycho-education was the main content of these sessions. In addition, the patients had two individual sessions with a psychologist (apart from the intake procedure). Six weeks after the 15 sessions, a last session was organized in which the patients discussed with each other what they had achieved.

Measurements

Neuropsychological functioning

The cognitive status of participants was assessed prior to treatment with well-accepted and validated neuropsychological tests for which norms were available (T0). Short-term memory was measured with the Digit Span and Reverse Digit Span of the Wechsler Adult Intelligence Scale III (WAIS-III) [39]. Episodic memory was tested with the Auditory Verbal Learning Test (AVLT) [40], and the Complex Figure Test (CFT) [41, 42]. Attention was measured with the D2 [43, 44], part A of the Trail Making Test (TMT) [45], and card 1 of the Stroop Colour Word Test (SCWT) [46]. Executive functions were measured with card 3 of the SCWT, the action plan test, the rule shift test, the zoo map test and the six element test of the Behavioural Assessment of Dysexecutive Syndrome (BADS) [47], and part B of the TMT. Language was assessed with the naming part of the Aachen Aphasia Test (AAT) [48], verbal phonemic fluency with the digits 'D', 'A' and 'T', and verbal semantic fluency of the Groninger Intelligence Test (GIT) [49] (animals). General intellectual functioning was assessed with the Matrices reasoning of the WAIS-III [39]. A cognitive domain was impaired if one or more scores of the neuropsychological tests for that domain were below cut-off (standard deviation ≥ 2 or percentile ≤ 5 or decile < 2). As the main purpose of rehabilitation is not to improve test performance, but to target the functional consequences of the impairment, evaluation should not be focused on cognitive deficits [32, 50]. As such we did not re-assess the neuropsychological tests as we expect them to be the same or show only minimal change.

Primary outcome measures for patients
Stroke Adapted Sickness Impact Profile (SA-SIP)
The SA-SIP is a questionnaire, which assesses quality of life. The Stroke Adapted version is shorter than the original version, only 30 of the original 136 items are included. Scores range from 0 to 100 (higher score denotes worse functioning or lower quality of life). The scale is divided in two subscales: physical and psychosocial functioning. The reliability and validity of the adapted SA-SIP are good [51].

Cognitive Failure Questionnaire (CFQ)
In this self-report questionnaire patients rate the amount of mistakes they make daily due to cognitive deficits, such as forgetting names, or problems with attending to objects or persons. The questionnaire consists of 25 items; patients can rate the frequency of their mistakes on a 5-point scale ('never' to 'very often'). The questionnaire is filled in by the patient [52]. The range of scoring is 0 to 100 (a higher score denotes more everyday cognitive failures). Ponds et al. [53] extended the CFQ with 4 items: 'higher frequency of', 'more hinder of', 'more worrying about', and 'more irritated about'. These extensions were also used in the current study.

Goal Attainment Scaling (GAS)
GAS is a global outcome measure for setting individualized goals and assessing the level of attainment of the goals [54]. GAS allows for both individualization of patient goals according to the needs of each patient, and standardization of measurement by using a summary formula that calculates the extent to which patients' goals are met [55]. GAS reflects actual improvement in a patients' functional ability [56]. The hypothetical mean GAS score at follow-up is 50 (SD = 10); consistently high or low follow-up scores indicate that goals were too easy (scores > 50), or too difficult to attain (scores < 50) [57]. A GAS score at follow-up of 50 indicates that all predefined goals met the expected outcome at follow-up. GAS was administered to 22 participants, because the final treatment group did not participate in this measurement due to practical issues.

Secondary outcome measures for patients
Frenchay Activities Index (FAI)
The FAI assesses the level of instrumental ADL and social activities of the patient. Originally the questionnaire was developed to assess impairments due to stroke [58]. The questionnaire consists of 15 items. The range of scoring is 0 to 45 (a higher score denotes better functioning).

Symptom Checklist 90 items (SCL-90)
This self-report questionnaire is multidimensional and assesses a broad range of complaints: anxiety, agoraphobia, depression, somatic complaints, insufficiency in thoughts, distrust, hostility, and sleeping problems. Patients rate the frequency of the complaint in the last week on a 5-point scale. Total score is 90-450 (a higher score denotes more complaints) [59].

Community Integration Questionnaire (CIQ)
In this self-report questionnaire 10 statements are offered measuring societal participation or community reintegration, such as: I feel committed to someone, I have enough leisure time. Patients indicate how much they agree with the statement (from 1 = full agreement to 5= no agreement at all). Total score ranges from 10 to 50. A higher score means better participation [60].

Fatigue Severity Scale (FSS)
This self-report questionnaire consists of 9 items, containing information about the severity of fatigue during daily situations in the past few days. The patient scores the amount of interference due to fatigue in daily life on a seven-point scale (1 = completely agree to 7 = completely disagree) [61]. A lower score denotes more fatigue.

Informant Questionnaire on Cognitive Decline (IQCODE)
This short questionnaire is designed to assess cognitive decline, compared to earlier functioning. A relative fills in this questionnaire about the patients functioning. The questionnaire consists of 16 items. The relative can rate cognitive decline on a five-point scale [62]. Scoring is from 1 to 5; 1 means same level of cognitive functioning as before the brain injury and 5 means much worse cognitive functioning than before.

Secondary outcome measures for relatives
Utrecht Coping List (UCL)
This questionnaire has been developed to describe different characteristics of coping style. In this questionnaire coping is defined as a personality trait, and it is presumed that persons have a preference for a (combination of) coping style in different situations. The questionnaire consists of 7 subscales: active approach, palliative reaction, avoidance, seeking social support, passive reaction pattern, expression of emotions, and reassuring thoughts. The total scale contains 47 items. Answers are possible on a four-point scale ranging from 1: never to 4: very often) [63].

Caregiver strain index (CSI)
The CSI has been developed to describe the burden caregivers encounter when caring for a sick relative. The 13 items of the CSI are based on 10 daily stressors in a typical situation. Caregivers rate with either one or zero. A score higher than 7 indicates a high level of burden [64].

Procedure
After referral to the treatment programme, patients and relatives were asked to participate in the study. After informed consent was obtained, baseline measurements concerning neuropsychological functioning were conducted in the rehabilitation centre prior to the start of the treatment. Demographic variables as well as medically relevant data (diagnosis, severity of injury, side of injury) were collected from the medical files. Education was divided into high (i.e. primary school + secondary school + higher education) versus low (i.e. primary school and or secondary vocational education).

The questionnaires for patients as well as relatives were sent to their home address and they were asked to fill in the questionnaires themselves, independently of each other. During a home visit by a member of the research team the questionnaires were checked to see if they were complete. In addition, the goals for the GAS procedure were set together with the patient. This procedure was conducted prior to the treatment (T0). In the week immediately after the treatment (T1: 21 weeks later) and 6 months after the end of the treatment (T2: in total 45 weeks after the start of treatment) the questionnaires were sent by mail and a member of the research team called the participants to score the level of attainment per goal by telephone. The Ethics Committee approved the study protocol.

Statistical analyses

Descriptive statistics were used to describe the demographic variables, neuropsychological functioning and daily functioning. Primary outcome measures of this study are the SA-SIP, the CFQ and GAS. Pre and post assessment were compared to each other (T0-T1 and T1-T2). For these analyses, paired Samples T tests were used. A repeated measures ANOVA (General linear model with repeated measures) was not performed because missing data at T1 and T2 would lead to a smaller sample in the overall analysis. As data on GAS were complete for 22 patients, a repeated measures ANOVA was conducted with the outcome score (GAS scores) as within-subject variable and the time of measurement (T0, T1 and T2) as between-subject variable in order to investigate the effect of treatment. Tukey post hoc tests were requested to evaluate whether differences in mean outcome scores for the different times of measurement were significant. Statistical analyses were performed with the Statistical Package for Social Sciences (version 16.0) with an alpha level set at 0.05 for all analyses.

Results

Participants

During the study period, five groups of patients were treated. Thirty-two patients received treatment, indicating that five patients refused participation in the study. Thus 27 patients participated in the study. Table 1 shows the patient characteristics.

Most participants were middle-aged, while more than half the patients had a low level of education. Most frequent cause of brain injury was stroke, followed by traumatic brain injury. Mean time since injury was almost 2 years. In both groups (patients and caregivers) the distribution of males versus females was equal and the caregivers had the same mean age as the patients.

Neuropsychological functioning is presented in Table 2. Deficits were found in all cognitive domains except language. Most deficits were found in the field of attention and memory, followed by executive functioning. Although aphasia was not an exclusion criterion for participation in the programme, none of the patients have language deficits.

Table 1. Patient characteristics

	Patient		Caregiver
Men (n(%))	14 (52)		13(52)
Age at assessment (mean, SD, range)	49.5 (9.2) 29-65		48.8 (8.8) 27 - 68
Time since injury in years (mean, SD, range)	1.9 (2.0) 0.4-7.4		
Education, low-high (n(%))	16 (59) - 11 (41)		
Cause of injury n(%)	Stroke	9 (33)	
	TBI	5 (18)	
	SAH	3 (11)	
	Brain tumour	4 (14)	
	Epilepsy	1 (4)	
	Viral infection	1 (4)	
	Sinus thrombosis	1 (4)	
	HIBI	1 (4)	
	OPS	1 (4)	
	Mixed type	1 (4)	

TBI: traumatic brain injury; SAH: subarachnoidal haemorrhage; HIBI: hypoxia induced brain injury; OPS: Organopsychosyndrome

Short-term effects of the programme

Table 3 presents the results of the primary and secondary outcome measures. The only significant change from baseline (T0) to directly after the programme (T1) was found for the physical subscale of the SA-SIP (p < 0.05). The scores of this scale increase between T0 and T1, indicating worse physical functioning. These scores decrease between T1 and T2, although not significantly. We also looked at individual changes on the other questionnaires. These were random: some patients improved on some questionnaires, whereas others deteriorated or did not change, there was no consistency and there were no statistical differences.

Long-term effects of the programme

We explored the differences between directly after the programme (T1) and 6 months after the end of the treatment (T2). We found that patients improved on social activities as measured with the FAI (p < 0.05) and reported feeling less fatigued (FSS, p < 0.05). No significant changes were found on the other scales.

Effect of the programme measured with GAS

GLM shows an effect of time on the GAS scores (F = 94.705, p < 0.05). A significant mean difference on the GAS scores exists between T0 and T1 and T0 and T2 (p < 0.05). The effect size was 4.7. No significant differences were found on the GAS scores between T1 and T2 (p > 0.05). Figure 1 shows the GAS scores for all participants over time.

Table 2. Neuropsychological functioning, including number of patients with a deficit

	Mean (SD)	Range	N deficit
Short term memory			
WAIS III Digit Span	13.8 (3.9)	8.0 - 22.0	2
Episodic memory			
AVLT	43.5 (9.8)	22.0 - 63.0	7
RCFT (delayed recall)	16.8 (6.2)	6.0 - 27.0	11
Attention			
D2	336.2 (94.5)	154.0 - 507.0	8
TMT A	43.4 (24.6)	22.0 - 150.0	11
Executive functioning			
TMT B	92.7 (51.0)	40.0 - 279.0	7
Stroop	116.3 (59.0)	63.0 - 357.0	11
Phonetic Fluency	29.8 (8.2)	15.0 - 57.0	7
Semantic Fluency	22.0 (5.2)	12.0 - 35.0	11
BADS			
Action Plan	4.5 (1.1)	1.0 - 5.0	6
Zoo Map	9.6 (3.7)	0.0 - 16.0	5
Rule Shift	1.5 (2.1)	0.0 - 7.0	5
Six Elements	5.1 (1.5)	2.0 - 6.0	4
Language			
AAT	115.4 (3.2)	108.0 - 120.0	0
Abstract reasoning			
WAIS III Matrices	17.1 (5.7)	5.0 - 25.0	9

WAIS III: Wechsler Adult Intelligence Scale, third edition; AVLT: Auditory Verbal Learning Test; RCFT: Rey Complex Figure Test; TMT: Trail Making Test, part A and B; BADS: Behavioural Assessment of the Dysexecutive Syndrome; AAT: Akense Afasie Test

Table 3. Mean scores on the questionnaires and differences between the assessments

Questionnaire	T0 Mean (SD)	T1 Mean (SD)	T2 Mean (SD)	T0-T1 p-value (N)	T1-T2 p-value (N)
Patient					
SA-SIP Total	24.9 (11.5)	24.5 (15.2)	23.9 (15.1)	1.00 (26)	0.90 (23)
SA-SIP Physical	8.8 (10.5)	12.6 (13.9)	9.9 (14.7)	**0.03* (26)**	0.31 (23)
SA-SIP Psychosocial	41.0 (18.2)	36.7 (23.1)	39.1 (22.9)	0.13 (26)	0.47 (23)
CFQ	48.8 (15.6)	49.3 (17.1)	49.4 (16.7)	0.43 (26)	0.72 (23)
FAI	27.5 (7.9)	28.9 (6.6)	31.9 (6.6)	0.11 (26)	**0.03* (23)**
SCL-90	156.5 (46.0)	147.4 (53.2)	145.2 (41.2)	0.05 (26)	0.43 (23)
CIQ	17.4 (4.2)	17.6 (4.0)	18.0 (3.2)	0.69 (26)	0.64 (23)
FSS	3.0 (1.5)	2.8 (1.5)	3.1 (1.5)	0.08 (26)	**0.02* (23)**
Relative					
IQCODE	3.6 (0.4)	3.6 (0.5)	3.6 (0.5)	0.22 (24)	0.92 (21)
UCL Active reaction	19.8 (3.6)	18.8 (3.3)	18.8 (3.4)	0.06 (24)	0.69 (20)
UCL Palliative reaction	17.6 (4.1)	17.4 (3.8)	16.2 (3.8)	1.00 (24)	0.23 (20)
UCL Avoidance	15.0 (3.4)	14.8 (3.8)	14.0 (2.9)	0.76 (24)	0.93 (20)
UCL Seeking social support	12.5 (3.3)	12.3 (3.0)	11.8 (2.8)	0.86 (24)	0.49 (20)
UCL Passive reaction pattern	11.1 (3.6)	11.0 (3.7)	9.5 (1.5)	0.92 (24)	0.19 (20)
UCL Expression of emotions	6.0 (1.7)	5.9 (1.6)	5.7 (1.4)	0.81 (24)	0.73 (20)
UCL Reassuring thoughts	12.2 (2.3)	11.9 (2.0)	10.9 (2.1)	0.62 (24)	0.08 (20)
CSI	5.4 (3.0)	5.6 (3.0)	4.6 (2.8)	0.37 (24)	0.08 (21)

*p < 0.05; SA-SIP: Stroke Adapted Sickness Impact Profile; CFQ: Cognitive Failure Questionnaire; FAI: Frenchay Activities Index; SCL-90: Symptom Checklist; CIQ: Community Integration Questionnaire; FSS: Fatigue Severity Scale; IQCODE: Informant Questionnaire on Cognitive Decline; UCL: Utrecht Coping List; CSI: Caregiver Strain Index

[a] T0-T1: Effect from before to directly after programme

[b] T1-T2: Effect from directly after programme to 6 months later

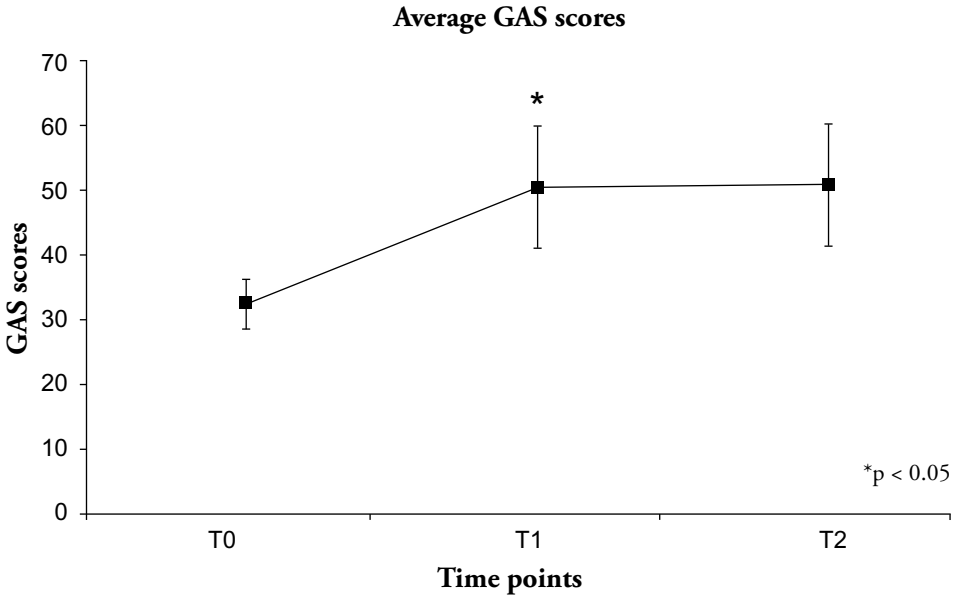**Figure 1.** Change of Goal Attainment Scaling (GAS) scores over time

Average GAS scores

*$p < 0.05$

Table 4. Domains in which goals were set with Goal Attainment Scaling (GAS)

Cognition	n	Emotion	n	Behaviour	n	Other	n
Memory	14	Self-esteem	1	Impulsivity/ inhibition	3	Work related	1
Attention/ concentration	12	Fear	4	Aggression/ temper	5	Education	4
Planning/ organizing	5	Acceptance	8	Social skills	5	Physical	1
Structure	1	Mood	1	Fatigue	5	Administration	1
Strategy use	6	Burden	2	Take some rest	7		
Dual tasking	1			Restless	4		
Flexibility	2						
Total	**41**		**16**		**29**		**7**

Table 4 gives an overview of the distribution of the number of goals set in the general and specific domains. In total 93 goals were set. The mean (SD) number of goals was 4.2 (1.1) per participant (range 2-6). Most goals were set in the cognitive domain (e.g. memory and attention), followed by the behavioural domain (e.g. take some rest and fatigue).

Discussion

This is the first prospective report showing the effect of an outpatient community re-entry cognitive rehabilitation programme on a group of patients with acquired brain injury with subtle cognitive as well as emotional and social problems. We found that patients attained their pre-defined goals. This finding remained stable six months after the end of the programme. The only significant changes noted in the answers to the questionnaires were on the physical subscale of the SA-SIP (short-term deterioration), the FAI and the FSS (long-term improvements). None of the caregivers' questionnaires showed changed answers either in the short or the long term. It seems that although patients attain their individual goals, this does not result in a higher participation level or a better quality of life.

We did not find robust effects of the programme on the questionnaires. An explanation for this is that GAS focuses on what the participants want to attain themselves and this is probably of more importance than the domains covered by the other outcome measures used for this study. In other words, GAS is patient-based while the other measures were probably more intervention-based. The goal of the intervention is to improve daily life by focusing on coping with daily life problems. GAS turned out to be the most sensitive instrument with which to measure outcome of the intervention. Liu et al. [65] state that goal attainment is seldom reported in studies of rehabilitation effectiveness. While GAS has the advantage of individual flexibility and sensitivity to treatment, it has been criticised as being overly subjective [66]. In addition, during the measurement at the end of treatment, all patients reported that they were very content about the treatment programme as it gave them the opportunity to talk about and share their experiences. Peer support programmes have indeed been found to be a promising way to approach the enhancing of patients' and families' ability to cope [67].

Apparently, the questionnaires used in this study are not suitable for identifying the benefit of a cognitive rehabilitation programme such as described in this study. We chose these questionnaires because they incorporate a broad domain of the possible effects of such a rehabilitation programme. It is remarkable that most goals were attained in the cognitive domain, while cognitive failures did not improve significantly. A measure of self-efficacy or mastery might have been a more sensitive instrument for change in this group, since the focus is strongly on managing the presence of cognitive and emotional problems in daily life. However, as we started the study there were no reliable measurement instruments available that had been validated in a brain injury population. Recently a traumatic brain injury Self-Efficacy Questionnaire was developed, based on generic self-efficacy measures [68] which could be a suitable candidate for future studies.

We did not expect patients to show deterioration on some of the questionnaires. However, we know from other studies that a response shift could have occurred, in which patients become more aware of their problems due to the programme [69]. We did not find that any particular group of patients deteriorated or improved. Changes in answers on questionnaires were random; some patients showed improvement on some questionnaires, but deterioration on others.

Other studies did find improvement in the quality of life, participation or employability [19, 70, 71]. However, the intensity of these programmes was much higher compared to ours. Anson & Ponsford [33] studied the effect of a programme comparable to ours (coping skills group) and also found improvement on subjective measurements and coping. This again stresses the fact that the evaluation of a programme should exactly match the purposes of that programme, and not go beyond this purpose in terms of more general expectations like increasing quality of life or feeling happier. It could be that changes in these more general domains need more time.

The results found should be interpreted against the shortcomings of the study. The design of this study is a pre-post test design, with a follow-up assessment. This was an open trial in order to examine the effectiveness of a new cognitive rehabilitation programme, awaiting further confirmation by a randomized controlled trial at a later stage. With this design we get a first insight into the effects of this programme for patients with acquired brain injury. We can conclude that this programme has had a positive effect on the patients, so the next step would be to perform a randomized controlled trial to indicate the effective parts of the programme.

In conclusion, our results indicate that patients profit from the education, cognitive strategies and social skills learned in this post-acute outpatient community re-entry cognitive rehabilitation programme. The effects are mostly observed in individually defined goals. The effectiveness of this programme should be further evaluated in a large-scale randomised controlled trial.

References

1. Colantonio A, Ratcliff G, Chase S, Kelsey S, Escobar M, Vernich L. Long-term outcomes after moderate to severe traumatic brain injury. Disabil Rehabil 2004;26(5):253-261.
2. Dikmen SS, Bombardier CH, Machamer JE, Fann JR, Temkin NR. Natural history of depression in traumatic brain injury. Arch Phys Med Rehabil 2004;85(9):1457-1464.
3. Dikmen SS, Machamer JE, Powell JM, Temkin NR. Outcome 3 to 5 years after moderate to severe traumatic brain injury. Arch Phys Med Rehabil 2003;84(10):1449-1457.
4. Frencham KA, Fox AM, Maybery MT. Neuropsychological studies of mild traumatic brain injury: a meta-analytic review of research since 1995. J Clin Exp Neuropsychol 2005;27(3):334-351.
5. Kashluba S, Paniak C, Blake T, Reynolds S, Toller-Lobe G, Nagy J. A longitudinal, controlled study of patient complaints following treated mild traumatic brain injury. Arch Clin Neuropsychol 2004;19(6):805-816.
6. Kersel DA, Marsh NV, Havill JH, Sleigh JW. Neuropsychological functioning during the year following severe traumatic brain injury. Brain Inj 2001;15(4):283-296.
7. Van der Naalt J. Prediction of outcome in mild to moderate head injury: a review. J Clin Exp Neuropsychol 2001;23(6):837-851.
8. Rao V, Lyketsos C. Neuropsychiatric sequelae of traumatic brain injury. Psychosomatics 2000;41(2):95-103.
9. Rapoport MJ, Feinstein A. Outcome following traumatic brain injury in the elderly: a critical review. Brain Inj 2000;14(8):749-761.
10. Rasquin SM, Lodder J, Ponds RW, Winkens I, Jolles J, Verhey FR. Cognitive functioning after stroke: a one-year follow-up study. Dement Geriatr Cogn Disord 2004;18(2):138-144.
11. Turner-Stokes L, Hassan N. Depression after stroke: a review of the evidence base to inform the development of an integrated care pathway. Part 1: Diagnosis, frequency and impact. Clin Rehabil 2002;16(3):231-247.
12. Bootes K, Chapparo CJ. Cognitive and behavioural assessment of people with traumatic brain injury in the work place: occupational therapists' perceptions. Work 2002;19(3):255-268.
13. Cattelani R, Tanzi F, Lombardi F, Mazzucchi A. Competitive re-employment after severe traumatic brain injury: clinical, cognitive and behavioural predictive variables. Brain Inj 2002;16(1):51-64.
14. Drake AI, Gray N, Yoder S, Pramuka M, Llewellyn M. Factors predicting return to work following mild traumatic brain injury: a discriminant analysis. J Head Trauma Rehabil 2000;15(5):1103-1112.
15. Franulic A, Carbonell CG, Pinto P, Sepulveda I. Psychosocial adjustment and employment outcome 2, 5 and 10 years after TBI. Brain Inj 2004;18(2):119-129
16. Pepin M, Dumont C, Hopps S. Relationship between cognitive capabilities and social participation among people with traumatic brain injury. Brain Inj 2000;14(6):563-572.

17. Visser-Meily J, Post M, Schepers V, Ketelaar M, van Heugten C, Lindeman E. Spouses' satisfaction with caregiver support in stroke rehabilitation. Scand J Caring Sci 2005;19 (4):310-316.

18. Wyller TB, Thommessen B, Sodring KM, Sveen U, Pettersen AM, Bautz-Holter E, et al. Emotional well-being of close relatives to stroke survivors. Clin Rehabil 2003;17(4):410-417.

19. Tiersky LA, Anselmi V, Johnston MV, Kurtyka J, Roosen E, Schwartz T, et al. A trial of neuropsychologic rehabilitation in mild-spectrum traumatic brain injury. Arch Phys Med Rehabil 2005;86(8):1565-1574.

20. Visser-Meily J, van Heugten C, Schepers V, van den Bos G. There are also suitable treatments during the chronic phase following a stroke. Dutch. Ned Tijdschr Geneeskd 2007;15(151):2753-2757.

21. Fraas M, Balz M, Degrauw W. Meeting the long-term needs of adults with acquired brain injury through community-based programming. Brain Inj 2007;21(12): 1267-1281.

22. McAllister T. Evaluation of brain injury related behavioral disturbances in community mental health centers. Community Ment Health J 1997;33(4):341-364.

23. Carney N, Chesnut RM, Maynard H, Mann NC, Patterson P, Helfand M. Effect of cognitive rehabilitation on outcomes for persons with traumatic brain injury: a systematic review. J Head Trauma Rehabil 1999;14(3):277-307.

24. Cicerone KD, Dahlberg C, Malec JF, Langenbahn DM, Felicetti T, Kneipp S, et al. Evidence-based cognitive rehabilitation: updated review of the literature from 1998 through 2002. Arch Phys Med Rehabil 2005;86(8):1681-1692.

25. Chua KS, Ng YS, Yap SG, Bok CW. A brief review of traumatic brain injury rehabilitation. Ann Acad Med Singapore 2007;36(1):31-42.

26. Cicerone KD, Mott T, Azulay J, Friel JC. Community integration and satisfaction with functioning after intensive cognitive rehabilitation for traumatic brain injury. Arch Phys Med Rehabil 2004;85(6):943-950.

27. Malec JF. Impact of comprehensive day treatment on societal participation for persons with acquired brain injury. Arch Phys Med Rehabil 2001;82(7):885-895.

28. Ponsford J, Olver J, Ponsford M, Nelms R. Long-term adjustment of families following traumatic brain injury where comprehensive rehabilitation has been provided. Brain Inj 2003;17(6):453-468.

29. Prigatano GP. Principles of neuropsychological rehabilitation. Oxford: Oxford University Press 1999.

30. Salazar AM, Warden DL, Schwab K, Spector J, Braverman S, Walter J, et al. Cognitive rehabilitation for traumatic brain injury: a randomized trial. Defense and Veterans Head Injury Program (DVHIP) Study Group. JAMA 2000;283(23):3075-3081.

31. Wilson BA. Compensating for cognitive deficits following brain injury. Neuropsychol Rev 2000;10(4):233-243.

32. Wilson BA, Evans JJ, Keohane C. Cognitive rehabilitation: a goal-planning approach. J Head Trauma Rehabil 2002;17(6):542-555.

33. Anson K, Ponsford J. Coping and emotional adjustment following traumatic brain injury. J Head Trauma Rehabil 2006;21(3):248-259.

34. Dittmar C. Outpatient rehabilitation program for clients with persisting mild to moderate symptoms following traumatic brain injury. Appl Neuropsychol 1997;4(1):50-54.

35. Miller LJ, Mittenberg W. Brief cognitive behavioral interventions in mild traumatic brain injury. Appl Neuropsychol 1998;5(4):172-183.

36. Mittenberg W, Tremont G, Zielinski RE, Fichera S, Rayls KR. Cognitive-behavioral prevention of postconcussion syndrome. Arch Clin Neuropsychol 1996;11(2):139-145.

37. Paniak C, Toller-Lobe G, Reynolds S, Melnyk A, Nagy J. A randomized trial of two treatments for mild traumatic brain injury: 1 year follow-up. Brain Inj 2000;14(3):219-226.

38. Malec JF, Basford JS. Postacute brain injury rehabilitation. Arch Phys Med Rehabil 1996;77(2):198-207.

39. Wechsler D. Wechsler Adult Intelligence Scale-III. San Antonio, TX: The Psychological Corporation 1997.

40. Rey A. L'examen clinique en psychologie [The clinical examination in psychology]. Paris: Presses Universitaires de France 1958.

41. Osterrieth PA. Le test de copie d'une figure complex: contribution a l'etude de la perception et de la memoire. Archives de Psychologie 1944;30:286-356.

42. Rey A. L'examinen psychologique dans les cas d'encephalopathie traumatique. Archives de Psychologie 1941;28:286-340.

43. Brickenkamp R. Aufmerksamkeits-Belastungs-Test (test d2) [The d2 test of attention]. Gottingen: Hogrefe 1962.

44. Brickenkamp R, Zillmer E. The d2 Test of Attention. Seattle, Washington: Hogrefe & Huber Publishers 1998.

45. Reitan R. Trail Making test: Manual for administration, scoring, and interpretation. Bloomington: Indiana University 1956.

46. Stroop JR. Studies of interference in serial verbal reactions. J Exp Psychol 1935;18:643-662.

47. Wilson BA, Alderman N, Burgess PW, Emslie H, Evans JJ. Behavioural Assessment of the Dysexecutive Syndrome (BADS): Thames Valley Test Company 1996.

48. Graetz P, de Bleser R, Willmes K. Akense Afasietest, Nederlandse versie. Lisse: Swets & Zeitlinger 1992.

49. Luteijn F, van der Ploeg FAE. Manual Groninger Intelligence Test. Lisse: Swets and Zeitlinger 1983.

50. Ylvisaker M, Hanks R, Johnson-Greene D. Perspectives on rehabilitation of individuals with cognitive impairment after brain injury: rationale for reconsideration of theoretical paradigms. J Head Trauma Rehabil 2002;17(3):191-209.

51. Van Straten A, de Haan RJ, Limburg M, Schuling J, Bossuyt PM, van den Bos GA. A stroke-adapted 30-item version of the Sickness Impact Profile to assess quality of life (SA-SIP30). Stroke 1997;28(11):2155-2161.

52. Broadbent DE, Cooper PF, FitzGerald P, Parkes KR. The Cognitive Failures Questionnaire (CFQ) and its correlates. Br J Clin Psychol 1982;21(Pt 1):1-16.

53. Ponds RW, van Boxtel M, Jolles J. The Cognitive Failure Questionaire as a measure for subjective functioning. Tijdschrift voor Neuropsychologie 2006;1(2):37-42.

54. Kiresuk T, Sherman RE. Goal Attainment Scaling: a general method for evaluating comprehensive community mental health programs. Community Ment Health J 1968;4(6):443-453.

55. Joyce BM, Rockwood KJ, Mate-Kole CC. Use of goal attainment scaling in brain injury in a rehabilitation hospital. Am J Phys Med Rehabil 1994;73(1):10-14.

56. Fisher K. Assessing clinically meaningful change following a programme for managing chronic pain. Clin Rehabil 2008;22(3):252-259.

57. Gordon JE, Powell C, Rockwood K. Goal attainment scaling as a measure of clinically important change in nursing-home patients. Age Ageing 1999;28(3):275-281.

58. Schuling J, de Haan R, Limburg M, Groenier KH. The Frenchay Activities Index. Assessment of functional status in stroke patients. Stroke 1993;24(8):1173-1177.

59. Arrindell WA, Ettema JHM. Symptom Checklist Handleiding bij een multidimensionele psychopathologie-indicator. Lisse: Swets & Zeitlinger B.V. 2003.

60. McColl MA, Davies D, Carlson P, Johnston J, Minnes P. The community integration measure: development and preliminary validation. Arch Phys Med Rehabil 2001;82(4):429-434.

61. Krupp L, LaRocca N, Muir-Nash J, Steinberg A. The fatigue severity scale. Application to patients with multiple sclerosis and systemic lupus erythematosus. Arch Neurol 1989;46(10):1121-1123.

62. Jorm AF, Jacomb PA. The Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE): socio-demographic correlates, reliability, validity and some norms. Psychol Med 1989;19(4):1015-1022.

63. Schreurs PJG, van de Willege G, Brosschot JF, Tellegen B, Graus GMH. De Utrechtse Coping Lijst (UCL): omgaan met problemen en gebeurtenissen. Lisse: Swets & Zeittinger B.V. 2003.

64. Robinson BC. Validation of a Caregiver Strain Index. J Gerontol 1983;38(3):344-348.

65. Liu C, McNeil JE, Greenwood R. Rehabilitation outcomes after brain injury: disability measures or goal achievement? Clin Rehabil 2004;18(4):398-404.

66. Boothroyd RA, Banks SM, Evans ME, Greenbaum PE, Brown E. Untangling the web: an approach to analyzing the impacts of individually tailored, multicomponent treatment interventions. Ment Health Serv Res 2004;6(3):143-153.

67. Hibbard MR, Cantor J, Charatz H, Rosenthal R, Ashman T, Gundersen N, et al. Peer support in the community: initial findings of a mentoring program for individuals with traumatic brain injury and their families. J Head Trauma Rehabil 2002;17(2):112-131.

68. Cicerone KD, Azulay J. Perceived self-efficacy and life satisfaction after traumatic brain injury. J Head Trauma Rehabil 2007;22(5):257-266.

69. Beutler L, Moos R, Lane G. Coping, treatment planning, and treatment outcome: discussion. J Clin Psychol 2003;59:1151-1157.

70. Geurtsen GJ, Martina JD, van Heugten CM, Geurts AC. A prospective study to evaluate a new residential community reintegration programme for severe chronic brain injury: the Brain Integration Programme. Brain Inj 2008;22(7):545-554.

71. Sarajuuri JM, Kaipio ML, Koskinen SK, Niemela MR, Servo AR, Vilkki JS. Outcome of a comprehensive neurorehabilitation program for patients with traumatic brain injury. Arch Phys Med Rehabil 2005;86(12):2296-2302.

Chapter 9

The practical use of Goal Attainment Scaling for people with acquired brain injury who receive cognitive rehabilitation

Abstract

Objective
To examine the feasibility of, and clinical experiences with Goal Attainment Scaling (GAS) when used for the evaluation of cognitive rehabilitation in people with acquired brain injury.

Design
A prospective observational longitudinal study.

Setting
A 21-week cognitive rehabilitation programme and a cognitive programme with varying length in two different Dutch rehabilitation centers.

Subjects
Forty-eight consecutive patients with acquired brain injury enrolled during a 15-month period.

Interventions
Cognitive rehabilitation programme.

Main measure
GAS; the number of goals was counted; time to set goals was recorded; the number of different domains in which goals were set was counted; GAS score was calculated at baseline, one week after the end of the cognitive rehabilitation programme and at six months follow-up; clinical experiences that could be useful for both clinical and research practice were recorded.

Results
The mean (SD) age of the patients was 46.1 (10.7) years; 29 (60%) were male; 186 goals were set with a mean (SD) number of 4 (1) goals per patient. It was possible to set at least three realistic goals per patient within 30 minutes. Most goals were set in the cognitive domain (i.e. memory and attention), followed by the behavioural domain (i.e. fatigue and aggression).

Conclusion
It proved possible to set three goals within an acceptable time-frame, to involve patients in the goal-setting procedure, to set realistic goals, and to set goals within relevant domains. Based on clinical experiences, GAS is less feasible for research when patients lack insight, or suffer from comorbidity or mood problems.

Introduction

Acquired brain injury can lead not only to physical deficits, but also to deficits in cognition, emotion, behaviour and psychosocial functioning [1-3]. After spontaneous recovery, most people with brain injury still experience problems in their daily lives, like forgetfulness, lack of initiative, anger and anxiety.

Cognitive rehabilitation is aimed at improving patients' ability to function in daily life [4]. Cognitive rehabilitation is based on well-described principles, such as compensatory training to adapt to the presence of a cognitive deficit [3-6]. As described in Ben-Yishay et al. [7], the general goals of cognitive rehabilitation are: to improve the self-awareness of strengths and weaknesses; to improve coping and compensating skills; to improve personal organization, social skills and effectiveness, emotional and behavioural self-management, participation in social, work, and leisure activities and health maintenance. It is important that these principles and general goals are in line with the personal needs and preferences of the patient. GAS is a method that takes these personal needs into account [8].

GAS is a measurement method for programme evaluation in mental health centres [9] that allows for both the individualization of patient goals according to the needs of each patient, and the standardization of measurement [10]. GAS reflects actual improvement in a patient's functional ability [11] and can be adapted to any level or domain of the International Classification of Functioning, Disability, and Health (ICF) [12]. Hence GAS is a method that can be used to cover any aspect of illness.

Due to the increasing emphasis on a more client-centered approach in health care in general, GAS has recently been used in patient groups suffering from various disorders including brain injury [9]. Measurement properties like reliability, validity, and responsiveness of GAS when used in people with brain injury who are receiving cognitive rehabilitation, were studied in advance and have proved promising [9, 10, 13-15]. Bouwens et al. [16] reviewed the literature on the applicability of GAS in psychogeriatric patients with cognitive disorders. They found that the use of GAS was feasible both with regard to the involvement of patient and/or caregiver in the goal-setting procedure as well as with regard to the possibility of implementing multiple domains in the goal-setting procedure. The feasibility of GAS in people with acquired brain injury who receive cognitive rehabilitation is however not well established, as the sample sizes of the above-mentioned studies were small, or the results were inconsistent and the clinical experiences were recorded insufficiently. Information about the usefulness of GAS in this population is important since a method should not only depend on good psychometric value. As GAS is a promising clinical method for interventions, its feasibility needs to be tested and clinical experiences with it in practice need to be examined.

Therefore, the aim of this study is to examine the feasibility of, and clinical experiences with GAS when used for the evaluation of cognitive rehabilitation in people with acquired brain injury.

Methods

Participants

This study is based on patients with acquired brain-injury who were referred to two ongoing cognitive rehabilitation programmes by neurologists, general practitioners and rehabilitation physicians. The programmes were provided by two Dutch rehabilitation centres; rehabilitation centre Blixembosch in Eindhoven and rehabilitation centre Hoensbroeck in Hoensbroek. Both cognitive rehabilitation programmes were part of normal service delivery. Patients enrolled in the programmes between September 2006 and December 2007.

The criteria for referral to the programme are the following: patients have sustained an acquired brain injury at least 3 months earlier, patients are older than 18 years, the patient experiences cognitive, emotional, and/or behavioural problems interfering with daily functioning for which information and advice and treatment are necessary to assure adaptation to these problems, the patient has sufficient insight in his or her own problems in order to benefit from a low frequent treatment intensity, communication in daily life is not hindered (the patient can read and write), the patients' social behaviour is adequate in order to function in a group programme, there are no treatment goals in the sensori-motor area of functioning (any more). Patients with primary psychological or psychiatric disorders, dementia, or a whiplash trauma are excluded from the programme.

The aims of both cognitive rehabilitation programmes are the following: 1) to help the patient and caregivers to gain insight in the consequences of the brain damage; 2) to offer strategies to handle cognitive deficits in their daily life; 3) to learn social skills in order to live with the consequences, 4) to learn how to control emotional reactions and 5) to enhance self-efficacy. With these aims, the impairments as such are not treated, but the patients and their caregivers learn to live with, manage, cope with and accept these consequences in order to maximize participation in society. The intensity of the programmes differed.

Measures

Goal Attainment Scaling (GAS)

GAS is used as a method to measure level of attainment of individual goals. This method is described in Appendix 1. Table 1 shows an example of a GAS follow-up guide.

Neuropsychological tests

A neuropsychological test battery consisting of twelve neuropsychological tests was administered to identify the level of cognitive functioning. For this study, we present mean (SD) scores of 5 well known neuropsychological tests of which adequate norms are available: the Auditory Verbal Learning Test (AVLT) [17] to measure verbal episodic memory; part A of the Trail Making Test (TMT) [18] to measure basal speed; the D2 [19] to measure selective and sustained attention; the action plan test, the rule shift test, the zoo map test and the six element test of the Behavioural Assessment of Dysexecutive Syndrome (BADS) [20] to measure executive functioning; and the Matrices reasoning of the WAIS-III [21] to measure general abstract reasoning. The level of everyday functioning was assessed via an interview

with the patient and caregiver separately about 87 instrumental activities of daily life (IADL). For this study we used data about preparing a meal, shopping, telephone use and activities with regard to mail, appointments and finances. We counted the number of patients who had problems with these activities.

Table 1. An example of a Goal Attainment Scaling follow-up guide of an adult brain-injured patient

	Goal Areas		
	Dealing with impaired memory	**Acceptance of consequences injury**	**Coping with aggression**
Much less than expected (- 2)	Poor short term memory, no use of memory aids	* No acceptance of disease	Becomes aggressive in all situations of misunderstanding
Somewhat less than expected (-1)	* Uses memory aids when another person cues to do so	Received information on coping strategies, but no feeling of coping	* Becomes aggressive in almost all situations of misunderstanding
Programme goal (0)	√Able to use memory aids with external cue like watch	√Learned how to cope with disease, but no full acceptance	Learned how to stay calm, but does not succeed in all situations
Somewhat better than expected (+1)	Able to use memory aids spontaneously	Almost full acceptance	√Rarely becomes aggressive
Much better than expected (+2)	No need for memory aids	100% acceptance of disease	Never becomes aggressive anymore

* Admission level
√ Discharge/follow-up level

Procedure

At baseline (one week before the start of the programme), a clinical researcher who was not involved in the cognitive rehabilitation programme, but who was familiar with the content of the programme, constructed GAS follow-up guides for all patients, either at home or at the rehabilitation centre. The team members of the programme received a copy of the goals so they knew what goals were set by each patient. At baseline the neuropsychological test battery was administered to all patients in a quiet testing room. The IADL-interview was done with the patient and caregiver separately.

The time point that was set as the target for goal attainment was one week after the end of the programme. At six months follow-up, the level of attainment was measured again to see if patients remained at the same level as was attained one week after the end of the programme. If the caregiver was present during the goal-setting procedure and the caregiver was influencing the patient, the clinical researcher made it clear that the patient

determined whether the goal was important or not. The clinical researcher scored the short-term level of attainment on the GAS follow-up guide with the patient one week after the end of the programme. Due to practical reasons, the scoring of the level of attainment was done by phone. The researcher read out loud the descriptions for each level of attainment and asked the patient which description was applicable at time of measurement. At six months follow-up, the clinical researcher made a second phone call to score the level of attainment again. The Ethics Committee approved this procedure.

Feasibility aspects and clinical experiences

To investigate the feasibility of GAS, the number of goals set was counted; time to set goals was recorded; the number of different domains in which goals were set was counted; GAS scores were calculated at baseline, one week after the end of the cognitive rehabilitation programme and at six months follow-up; and clinical experiences that could be useful for both clinical and research practice were recorded, such as the possibility to involve the patient in the goal-setting procedure.

The standards of the feasibility aspects were based on Bouwens et al. [16]:

1. The number of goals was considered sufficient when at least three goals were set.
2. The time required for setting the goals was considered good when it took a maximum of 30 minutes to construct a GAS follow-up guide of at least three goals.
3. Domains included in the goals had to cover at least two of the following general domains in order to be relevant: cognition, emotion and behaviour. As these three domains are the core focus of cognitive rehabilitation, they were self-evidently relevant to the method.
4. The achievement of a mean GAS score of 50 (SD = 10, range 40-60). The achievement of a mean GAS score of 50 one week after the end of the programme is the criterion by which we judge if the goals overall had been realistic.

Statistical analyses

Descriptive statistics of the patient characteristics were given, the time required for setting the goals, and the number of goals were calculated. The total number of goals that was set per general domain and per specific domain as well as the number of general domains covered by the goals was calculated for the total group. Mean GAS scores were performed per time of measurement.

Percentages were calculated for the number of goals for each level of attainment one week after the end of the programme and at six months follow-up. Independent-Samples T test was performed to analyse the difference between length of the rehabilitation programmes. Statistical analyses were performed with the Statistical Package for Social Sciences (version 16.0) with an alpha level set at 0.05 for all analyses.

Results

Forty-eight patients fulfilled the inclusion criteria for this study. Data at baseline were complete for all 48 patients; one week after the end of the programme this number was 44 (four patients stayed in the open-ended programme); and at six months follow-up this number was 40 (since this is an ongoing study, three patients could not have had the follow-up measurement after six months). The mean (SD) length of the programme as provided by rehabilitation centre Blixembosch was 7.6 (3.8) months; the length of the programme as provided by rehabilitation centre Hoensbroeck was 5 months for each participant. The length of treatment between the two rehabilitation centres differed significantly (F = 58.4, t = 3.3, df = 42, p < 0.05). One patient, who was treated for the consequences of his traumatic brain injury, had a stroke between the end of the programme and six months follow-up. His level of functioning decreased due to this stroke and therefore original goals could not be measured reliably at six months follow-up. The GAS score of this patient was excluded from the six months follow-up. Table 2 shows the patients' characteristics. More males than females participated in the study and the majority had had a stroke or a traumatic brain injury. The mean scores for the cognitive domains were all above cut off. Most patients experienced problems on instrumental activities of daily life.

Table 3 shows a descriptive analysis of goal numbers and GAS scores at each time point. When controlling for length of treatment, the change of GAS scores between baseline and follow-up did not differ between the two rehabilitation centres (p > 0.05). The details of this analysis are described separately below.

Number of goals and time required for setting the goals

Forty-eight patients set 186 goals in total. The mean number of goals set for the patients was 4 (1), ranging from 2-6. Four patients set two goals; these patients could not come up with more goals. It took less than 30 minutes to set these two goals for all four patients. The other 44 patients had three or more goals. It was possible to set at least three goals within 30 minutes for 24 of them. For 18 patients, the goal-setting procedure took between 30 and 45 minutes and in two cases it took one hour to set goals. Reasons for exceeding the 30-minute time-frame were: lack of insight, emotional and communication problems, problems with specifying goals and having more than three goals.

Domains included in the goals

Goals were set in four general domains: cognition, emotion, behaviour and other. Table 4 shows the number of goals for each general domain and specific domain. Most goals were set in the cognitive domain. Memory and attention/concentration were the cognitive categories in which most goals were set, followed by strategy use and planning/organizing. In the emotional domain, most goals were set for acceptance. Aggression/temper and fatigue were the behavioural domains mentioned most frequently.

Table 2. Patient characteristics (N = 48)

Age in years (mean (SD), range)	46.1 (10.7), 23-67
Male (number (%))	29 (60)
Time since brain injury in years (mean (SD), range)	2.1 (3.0), 0-13
Level of education (number (%))	
- Low	12 (25)
- Medium	19 (40)
- High	17 (35)
Cognitive functioning (mean (SD), range)	
- Memory (AVLT, decile 0-10, cut off < 2)	5.0 (3.0)
- Basal speed (TMT A, percentile 0-100, cut off < 5)	32.1 (32.5)
- Attention (D2, percentile 0-100, cut off < 5)	38.8 (34.0)
- Executive function (BADS, standard score, cut off -2)	
o Zoo map	-0.1 (1.2)
o Rule shift	-0.4 (1.3)
o Six element test	-0.3 (1.3)
o Action plan test	-0.2 (1.8)
- General intellectual functioning (WAIS III, standard score, cut off -2)	0.2 (1.1)
Everyday functioning (number of patients experiencing problems with IADL (%))	
- Preparing a meal	37 (77)
- Shopping	25 (52)
- Telephone use	38 (79)
- Activities with regard to mail, appointments and finances	43 (90)
Diagnosis (number (%))	
- Stroke	15 (31)
- Traumatic brain injury	15 (31)
- Subarachnoidal haemorrhage	6 (13)
- Brain tumour	5 (11)
- Epilepsy	1 (2)
- Viral infection	2 (4)
- Sinus thrombosis	1 (2)
- Hypoxia induced brain injury	2 (4)
- Mixed type	1 (2)

AVLT: Auditory Verbal Learning Test; TMT: Trail Making Test; BADS: Behaviour Assessment of the Dysexecutive Syndrome; WAIS III: Wechsler Adult Intelligence Scale, third edition; IADL: Instrumental Activities of Daily Living

Table 3. Descriptive analysis of goal numbers and Goal Attainment Scaling scores at each time point when used in people with acquired brain injury*

Number of goals per patient	4 (1), 2-6
Number of general domains covered by the goals	2 (1), 1-3
Goal attainment score at baseline	32.4 (3.8), 22.6 – 37.6
Goal attainment score at the end of the programme (n = 44)	51.4 (10.7), 31.4 – 71.1
Goal attainment score at six months follow-up (n = 40)	52.5 (11.0), 31.4- 71.1

* Values are presented as mean (SD), range for a total group of 48 patients, unless stated otherwise

To what extent were goals realistic?

The mean (SD) GAS score at baseline for 48 patients was 32.4 (3.8). One week after the end of the programme this score was 51.4 (10.7) (N = 44). The mean GAS score slightly increased at six months follow-up (52.5 (11.0) for 40 patients (Table 3)). Since both scores fall within the 40-60 range, it can be stated that goals on both measurements were realistic. The mean (SD) number of general domains covered by the goals per patient is 2 (1).

Some goals were not attained because the goals changed along the way according to the patients. This was the case for 10 goals, divided over six patients. During cognitive rehabilitation, two of these patients had learned to cope with their problems in other ways than had been described in the goals in advance (i.e. two patients had learned to focus on only one person during a conversation, while the goal was to follow a conversation of more than two people). At follow-up, these new coping styles could not be scored; therefore the predefined goal was not met. The other four patients found that the former goals were no longer an issue and that they had learned what was of real importance to them (e.g. one patient did not consider it any longer a problem if things did not go according to plan).

Level of attainment

Baseline data were complete for 48 patients, covering 186 goals; 44 (24%) of these goals were set at the '-2' level and 142 (76%) at the '-1' level. Longitudinal data at one week after the end of the programme were complete for 44 patients, covering 169 goals; 118 (70%) of these goals were attained (47 (28%) were attained, 63 (37%) were higher than expected, and 8 (5%) were much higher than expected), and 51 (30%) of these goals were not attained (37 (22%) were lower than expected, and (14 (8%) were much lower than expected). Longitudinal data at six months follow-up were complete for 40 patients, covering 155 goals; 112 (72%) of these goals were attained (47 (30%) were attained, 48 (31%) were higher than expected, and 17 (11%) were much higher than expected), and 43 (28%) of these goals were not attained (31 (20%) were lower than expected, and 12 (8%) were much lower than expected).

Table 4. Overview of content of domains in which goals were set

Cognition	n	Emotion	n	Behaviour	n	Other	n
Memory	24	Self-esteem	2	Impulsivity/inhibition	4	Work related	4
Attention/concentration	17	Fear	4	Restless	4	Education about consequences acquired brain injury	6
Planning/organizing	11	Acceptance	17	Assertiveness	9	Peer support	1
Structure	5	Mood	3	Aggression/temper	11	Physical	1
Time pressure	1	Preoccupation	1	Social skills	1	Administration	6
Strategy use	14	Deal with emotions	1	Affective feelings to others	8		
Dual tasking	1	Burden	7	Fatigue	13		
Flexibility	3	Enjoy life again	2	Take some rest	3		
				Talkative	1		
				Making choices	1		
Total	**76**		**37**		**55**		**18**

Clinical experiences

It took more time to set three goals in the first cases of the goal-setting procedure. After some practice, it became easier to set goals. Patients were asked to formulate goals in all 48 cases. The first reaction of many patients was that their goal was to become the person they had been before their brain injury. Since this goal is too broad and unrealistic in most cases, the clinical researcher had to explain to the patients that the goals had to be more specific.

Some patients had difficulties with the abstract nature of GAS. By providing these patients with some examples of possible goals and using verbal explanations for the different levels (e.g. 'much higher than expected') instead of scores (e.g. '-2'), they all understood the goal-setting procedure. In most cases (n = 33) the clinical researcher was alone with the patient when setting goals. In the other 15 cases, the caregiver was present.

In two cases, the caregiver came up with goals important for the relationship, like affection and sharing opinions. In these cases, we consulted both the caregiver and the patient before deciding to include the goals suggested by the caregiver. During the follow-up measures, the clinical researcher asked the patient's permission to also discuss the level of attainment of the goals suggested by the caregiver with the caregiver.

In two cases, the caregiver helped the clinical researcher in convincing the patient to set realistic goals and to motivate the patient to think of what he or she wanted to attain. Four cases lacked insight in their problems. Although the clinical researcher discussed lack of insight with three of them, they still wanted to set an unrealistic goal as a test to see if they could not do the things they could do before. The unrealistic level was set at +1 or +2 and the expected level was set at '0'. The fourth person denied all problems. After some discussion about daily life functioning with the clinical researcher and the caregiver, this patient formulated with three goals.

Besides some difficulties in setting goals with patients, we also experienced difficulties in measuring the attainment of goals with 2 patients. During a phone call these patients were asked to what extent they thought their goals had been met. The clinical researcher read the possibilities out loud. One patient mentioned that he found it difficult to indicate the level attained, since the descriptions were too abstract for him.

One patient was depressed at follow-up. She noticed that her depression could have influenced the outcome. She was negative about herself at the time of follow-up. In some cases, setting goals was confronting. Thinking about the things the patients had to change or adapt made them realize that things were going differently now. For some people this idea was motivating, while for others, this was disappointing.

Discussion

It proved possible to set three goals within an acceptable time-frame, to involve patients in the goal-setting procedure, to set realistic goals, and to set goals within relevant domains. We discovered that setting goals is difficult when patients have insufficient insight into their problems, experience emotional and communication problems or have difficulty specifying goals. Measuring the level of attainment is problematic when comorbidity occurs in between

the measurements; when patients have mood problems like depression; and when goals change along the way.

The strength of our study lies in the relatively large sample size. Only Rockwood et al. [15] had a comparable sample size. Furthermore, we tested the feasibility of, and clinical experiences with GAS in this population using well-defined feasibility aspects and the person who set the goals and assessed the level of attainment was not involved in the cognitive rehabilitation programmes. This procedure ensures independent assessment of the levels attained. In most studies goals were set and the level of attainment was assessed by the treatment team itself or by an independent researcher who discussed level of attainment with the treatment team. The clinical researcher in our study did not discuss with the treatment team what goals were set and what level was attained. Instead she directly asked the patient what goals the patient wanted to attain and what level was attained. The clinical researcher was familiar with the content of the rehabilitation programme and with setting goals. By providing the rehabilitation team with the goals it was ensured that the team knew the goals as well.

Some other studies that used GAS as a method for measuring outcome scored the follow-up measurements by phone as well [22, 23]. This approach, however, slightly differs from most studies in which level of attainment is scored after observation of the patient by the rehabilitation team or the assessor. A disadvantage of a follow-up by phone is the impossibility to see non-verbal cues. Moreover, people may feel uncomfortable when they have to discuss personal topics by phone without seeing the clinical researcher. An advantage of the use of the phone is that is it time effective. Our experience was that most patients understood the procedure and were able to report their level of functioning during a phone call. Furthermore, the nature of most goals set did not allow direct observation (e.g. level of acceptance). Therefore, it was not necessary to actually see the patients. By formulating the levels of attainment in a clear manner at baseline, it was possible for the researcher to assess level of attainment by phone. When goals have an observable nature, such as number of meters that a person can walk after an intervention, it is not possible to do the follow-up assessment by phone.

The patient reported the level of attainment. Therefore, it could have been possible for patients to provide socially desirable answers that did not reflect their actual level of attainment. However, as most levels of attainment could not be observed, the best way to assess level of attainment is by asking the patient.

The goal-setter has to be sure that the goals set by the individual do not go beyond the scope of the research or clinical intervention. So, to a certain extent, the goal-setter has to guide the patient. This guidance makes GAS an interactive approach that requires clinical skills, like being able to confront patients when they choose unrealistic goals or lack insight into their problems. For clinical purposes, this interactive process is part of the intervention. For the purpose of research, however, this interactive process before the intervention has started may influence the actual intervention.

In the case of some patients, confronting them with their unrealistic goals enhanced their insight into their problems. This aspect is relevant for clinical practice but not for

research purposes; the goal-setting procedure in the case of patients who lack insight is an interactive, therapeutic process that influences the actual research intervention.

Moreover, goals could be set with the patient alone. Assistance of the caregiver might be helpful when the patient cannot formulate any goal. When caregivers have specific wishes or needs, separate caregiver goals or system goals can be set as well.

GAS is an individualized approach and based on the needs of the patient. This individualized aspect is the strong point of the method and provides information about the level of satisfaction of the patient. As we have found in our study, however, some patients reported that they were satisfied with their level of functioning after the programme although they did not attain their predefined goals. Therefore, in clinical practice, GAS is recommended to use complementary and not as a replacement of the other instruments. In research, GAS should be compared with traditional standardized measures, such as the Frenchay Activities Index (FAI) [24] and the Stroke-Adapted Sickness Impact Profile (SA-SIP) [25].

Although GAS is an interesting method for both clinical and research aims, it is not yet well understood and recently started to be used in other settings than psychiatry. Several aspects need to be taken into account by those who consider using it. With regard to the goal-setting procedure, goal-setters must be aware of the patient's emotional status, level of insight, communication skills, and the capacity to specify goals. With respect to the level of attainment, it is important to identify possible depression and comorbidity.

It requires practice and clinical skills to learn to apply GAS. We recommend practising the goal-setting procedure before actually using it. It pays off to invest in the proper training of staff and the result is a useful tool.

Hurn et al. [26] reviewed the literature on clinimetric aspects of GAS when used as measure of outcome by clinicians working with patients in physical and neurological rehabilitation settings. They found strong evidence for the reliability, validity and sensitivity of GAS. Our study adds relevant information on the practical use of GAS. We recommend GAS for use in clinical practice when measuring clinical change. As GAS reflects individual needs, it can be useful for quality of life as well.

References

1. Ettenhofer ML, Abeles N. The significance of mild traumatic brain injury to cognition and self-reported symptoms in long-term recovery from injury. J Clin Exp Neuropsychol 2008;11:1-10.
2. Greenwood R. The consequences of brain injury: classification assessment of outcome. Neuropsychol Rehabil 1999;9(3-4):231-240.
3. Sohlberg MM, Mateer CA. Cognitive rehabilitation: an integrative neuropsychological approach. New York: The Guilford Press 2001.
4. Rehabilitation of persons with traumatic brain injury. NIH Consens statement online;16(1):1-41.
5. Prigatano GP. Principles of neuropsychological rehabilitation. Oxford: Oxford University Press 1999.
6. Wilson BA. Neuropsychological rehabilitation. Annu Rev Clin Psychol 2008;4:141-162.
7. Ben-Yishay Y, Prigatano GP. Cognitive remediation. In: Rosenthal M, Griffith ER, Bond MR, Miller JD, editors. Rehabilitation of the adult and child with traumatic brain injury. Philadelphia: FA Davis 1990:393-400.
8. Kiresuk T, Sherman RE. Goal Attainment Scaling: a general method for evaluating comprehensive community mental health programs. Community Ment Health J 1968;4(6):443-453.
9. Malec JF. Goal Attainment Scaling in Rehabilitation. Neuropsychol Rehabil 1999;9(3/4):253-275.
10. Joyce BM, Rockwood KJ, Mate-Kole CC. Use of goal attainment scaling in brain injury in a rehabilitation hospital. Am J Phys Med Rehabil 1994;73(1):10-14.
11. Fisher K. Assessing clinically meaningful change following a programme for managing chronic pain. Clin Rehabil 2008;22(3):252-259.
12. World Health Organization. International Classification of Functioning, Disability and Health. Geneva; 2001.
13. Goodyear DL, Bitter JA. Goal attainment scaling as a program evaluation measure in rehabilitation. Journal of Applied Counseling 1974;5(1):19-26.
14. Lannin N. Goal attainment scaling allows program evaluation of a home-based occupational therapy program. Occupational Therapy in Health Care 2003;17(1):43-54.
15. Rockwood K, Joyce B, Stolee P. Use of goal attainment scaling in measuring clinically important change in cognitive rehabilitation patients. J Clin Epidemiol 1997; 50(5):581-588.
16. Bouwens SFM, van Heugten CM, Verhey FRJ. Review of Goal Attainment Scaling as a useful outcome measure in psychogeriatric patients with cognitive disorders. Dement Geriatr Cogn Disord 2008;26(6):528-540.
17. Rey A. L'examen clinique en psychologie [The clinical examination in psychology]. Paris: Presses Universitaires de France 1958.
18. Reitan R. Trail Making test: Manual for administration, scoring, and interpretation. Bloomington: Indiana University 1956.

19. Brickenkamp R. Aufmerksamkeits-Belastungs-Test (test d2) [The d2 test of attention]. Gottingen: Hogrefe 1962.
20. Wilson BA, Alderman N, Burgess PW, Emslie H, Evans JJ. Behavioural Assessment of the Dysexecutive Syndrome (BADS): Thames Valley Test Company 1996.
21. Wechsler D. Wechsler Adult Intelligence Scale-III. San Antonio, TX: The Psychological Corporation 1997.
22. Stuifbergen AK, Becker H, Blozis S, Timmerman G, Kullberg V. A randomized clinical trial of a wellness intervention for women with multiple sclerosis. Arch Phys Med Rehabil 2003;84(4):467-476.
23. Stuifbergen AK, Becker H, Timmerman GM, Kullberg V. The use of individualized goal setting to facilitate behavior change in women with multiple sclerosis. J Neurosci Nurs 2003;35(2):94-99,106.
24. Schuling J, de Haan R, Limburg M, Groenier KH. The Frenchay Activities Index. Assessment of functional status in stroke patients. Stroke 1993;24(8):1173-1177.
25. Van Straten A, de Haan RJ, Limburg M, Schuling J, Bossuyt PM, van den Bos GA. A stroke-adapted 30-item version of the Sickness Impact Profile to assess quality of life (SA-SIP30). Stroke 1997;28(11):2155-2161.
26. Hurn J, Kneebone I, Cropley M. Goal setting as an outcome measure: a systematic review. Clin Rehabil 2006;20(9):756-772.
27. Malec JF, Smigielski JS, DePompolo RW. Goal attainment scaling and outcome measurement in postacute brain injury rehabilitation. Arch Phys Med Rehabil 1991;72(2):138-143.
28. Gordon JE, Powell C, Rockwood K. Goal attainment scaling as a measure of clinically important change in nursing-home patients. Age Ageing 1999;28(3):275-281.

Appendix 1. Description of Goal Attainment Scaling (GAS)

The GAS method consists of a six-step process. To select the goals that are relevant to the individual patient in step 1 the patient and/or caregiver is either interviewed by a clinician to identify problem areas and to determine goals for those areas in which intervention is planned. Team members can also set goals themselves after interviewing the patient and/or caregiver about problem areas. When GAS is used for research to test the effect of an intervention, the setting of goals by a clinical researcher who is not involved in the intervention is recommended. To ensure reliable calculation with a standard formula, a minimum of three goals is recommended [27].

Step 2 is optional. It provides the opportunity to weight the goals. Several methods exist to do so. We set all weigths at 1.

In Step 3 a follow-up time period before the start of the evaluation is determined, at which point the clinician will measure the level of goal attainment.

In step 4 the expected outcome, i.e. the goal, is set. Outcome levels are specified in observable terms in such a way that laymen are able to score the follow-up guides (Table 1). Levels do not overlap. The expected outcome is scored with '0'.

Step 5 consists of completion of the other scale levels, using possible outcomes including much less than expected (-2), somewhat less than expected (-1), somewhat better than expected (+1) and much better than expected (+2).

In step 6, GAS baseline levels are noted on the follow-up guide. If a clinically relevant deterioration is plausible, the description of the current status can be scored as '-1', with the worsened state scored as '-2'. If the problem is at its worst, the current level of functioning is scored as '-2'.

Patients receive an intervention and at the defined follow-up time (e.g. end of intervention), the patient is scored at the attained level. The overall GAS score at baseline and follow-up is calculated with a standard formula (Appendix 2). The formula produces an overall score that is an average of the outcome scores, adjusted for the relative weighting assigned to goals, the varying number of goals, and the expected inter-correlation among the goal scales.

The hypothetical mean GAS score at follow-up is 50 (SD = 10); consistently high or low follow-up scores indicate that goals were too easy (scores > 50) or too difficult to attain (scores < 50) [28]. A GAS score at follow-up of 50 indicates that all predefined goals met the expected outcome at follow-up.

Appendix 2. Formula to calculate the Goal Attainment Scaling scores

$$\text{Goal Attainment Scaling score} = \frac{50 + 10 \sum (w_i x_i)}{\sqrt{\left(0.7 \sum w_i^2 + 0.3 \left(\sum w_i\right)^2\right)}}$$

where xi = the attainment level and wi = the weight assigned to the goal area. Because the expected outcome/programme goal is assigned the score of 0, if all goals are met, the outcome score would be 50.

Chapter 10

General discussion

The main aim of this thesis was to investigate the ecological aspects of neuropsychological assessment by exploring the association between the domains which are affected in people with brain dysfunction (e.g. dementia and acquired brain injury). The findings of this investigation have been divided into three parts; the first part of the thesis shows the association between cognition and the daily life functioning of people with different types of dementia and of psychogeriatric patients with cognitive disorders; the second part of the thesis presents findings with regard to the ecological validity of the neuropsychological assessment on the basis of which conclusions are drawn about the daily life functioning of people with cognitive disorders due to dementia or acquired brain injury; and the third part of the thesis presents findings with regard to the applicability of individual goal-setting in patients with cognitive disorders due to brain dysfunction.

In this last chapter the main findings will be summarized, followed by a discussion of methodological considerations, theoretical issues and implications, and concluding with recommendations for future research.

Summary of findings

Part I: Association between cognition and daily life functioning

The first study (Chapter 2) examined in this thesis, in which the association between cognitive status and daily life functioning in people with different dementia subtypes is investigated, reveals that this association is related to the type of dementia. The strongest association was found with regard to frontotemporal dementia. In the case of Alzheimer's disease, vascular dementia and Parkinson's dementia the association was found to be moderate, and in the case of primary progressive aphasia it was absent. Furthermore, the scores on both cognition and daily life functioning differed between the dementia subtypes.

The finding that the association between cognition and daily life functioning depends on the dementia subtype underlines the need for a multidomain approach. The results support the importance of undertaking a functional as well as a cognitive assessment, not only when it comes to assessing patients' needs with regard to care, but also as a possible diagnostic tool, since the correlations differ according to dementia subtype.

The association between cognitive measures and a scale for instrumental activities of daily living (IADL) which involves the direct observation of psychogeriatric patients with cognitive disorders performing familiar IADL activities, is examined in Chapter 3. This association was in general found to be moderate, with an explained variance between 27% and 44%. It was higher in patients with vascular dementia in comparison to those with Alzheimer's disease; and higher in patients with mild dementia in to comparison to patients in the moderate-to-severe stages. Apathy was the only neuropsychiatric symptom that significantly influenced the association.

These findings underline the need for direct observation of daily activities in dementia patients. Another finding was that a traditional measure for assessing dementia severity, the Mini-Mental State Examination (MMSE) [1], has less value when it comes to the prediction of daily life functioning. With regard to the International Classification of Functioning, Disability, and Health (ICF) model [2], the MMSE is useful for the body function &

structure level (impairment), but of less value for the activities (limitation), participation (restriction) level and environmental factors.

The strengths of both these studies are the clinical settings from which the patients were selected and the large sample sizes (N = 615 and 118, respectively). The large samples were collected from the Maastricht Memory Clinic (MMC) [3] and the Maastricht Evaluation of a Diagnostic Intervention for Cognitively Impaired Elderly (MEDICIE) study [4]. Both samples contain research data that are standardized and collected by experienced clinicians using established research criteria. These aspects ensure that findings are reliable and relevant for clinical practice.

The number of different psychogeriatric diagnoses in the study reported in Chapter 3 was high, ensuring a heterogeneous sample. Theoretically, the heterogeneity within the sample could have biased the association between cognition and daily life functioning. However, as became clear in Chapter 2, this association is the same for Alzheimer's disease and vascular dementia (Pearson r = - 0.60 and - 0.56 respectively). Since the majority of the sample of the study reported in Chapter 3 had either Alzheimer's disease or vascular dementia, the effect of heterogeneity on the results is expected to be minimal.

A limitation of both studies is that these are historical cohorts. Therefore it was not possible to conduct prospective research.

Part II: Ecological validity of neuropsychological assessment and the integration of domains

The association between cognition and daily life functioning is relevant for neuropsychologists, since one of the aims of a neuropsychological assessment is to examine the practical consequences of cognitive impairments for a patient's daily life functioning. Therefore, it is important that neuropsychological tests are ecologically valid. A description of the aspects related to the ecological validity of neuropsychological tests is given in Chapter 4.

Summarizing recent literature, the review in Chapter 4 concludes that the association between neuropsychological tests and the daily life functioning of people with cognitive disorders is moderate. Direct observation of patients performing daily life activities in their own environment might be a good outcome measure for daily life functioning, but studies that investigate whether this is the case are lacking. The association between neuropsychological tests and daily life functioning might be influenced by several factors that challenge the ecological validity of the tests. Examples of these factors are the artificial testing environment and non-cognitive factors.

The ecological validity of neuropsychological assessment is investigated in a study (described in Chapter 5), which explores the extent to which experienced neuropsychologists can predict daily life functioning in people with acquired brain injury solely on the basis of neuropsychological data. This study revealed that the level of agreement between neuropsychologists about the problems expected during the performance of observed daily life activities varied considerably. The accuracy of the predicted daily life problems was in general insufficient, i.e. neuropsychologists were not able to predict the daily life functioning of people with acquired brain injury solely on the basis of neuropsychological data. Thus, neuropsychological data should only be relied on with a clear understanding of

the limitations when neuropsychologists are asked to predict daily life functioning. Moreover, direct observation is recommended as a complementary tool, since this provides important and unique information that is not elucidated by the neuropsychological data.

In Chapter 6 an international cooperative project is described. The project attempts to enhance the ecological validity by integrating neuropsychological data and observational data in patients with brain dysfunction. The aim of the procedure was to decompose daily life activities in order to identify the underlying neuropsychological aspects that are relevant for patient care and the diagnostic process. The sensitivity of the neuropsychological assessment is moderate (77%) with regard to identifying problematic daily life activities and its specificity is low (15%), while its accuracy was 60%. The new assessment procedure did provide insight into the underlying neuropsychological difficulties of patients with acquired brain injury and dementia.

The direct observation of patients in their own environment as a means to gain better insight into their daily life functioning, which is described in Chapter 5 and 6, introduces an innovative approach. This approach provides new insights into the ecological validity of neuropsychological assessment.

Unlike most studies, which used instruments that were merely derivatives of daily life functioning [5, 6], such as informant-based questionnaires or self-reports, the studies discussed in Chapter 5 and 6 rely on the observation of everyday functioning. However, it is necessary to note that although the individual approach based on the observation of relevant and important activities per patient is important for good clinical practice, its standardization would be complex. It is difficult to select a small sample of standard activities that are relevant and that could be observed in all patients, as would be required in order to improve the validation process of direct observation. Furthermore, an activity can be performed in different ways, each of which can be equally effective. Moreover, it is imaginable that the physical presence of an observer might influence performance.

The proper quantification of everyday activities is complicated. Qualitative judgements, however, are also valuable. The use of direct observation, description and deductive analysis of themes can contribute significantly to an objective understanding of complex behaviour in daily life functioning [7].

The neuropsychological and observational data of 48 patients with acquired brain injury are presented in Chapter 6. Due to practical considerations, only those activities were observed at patients' homes that took less than one hour to perform and could be observed in and around the house. Therefore, the sample of observed activities is somewhat biased.

Part III: Individual goal-setting

A problem that arises when investigating the ecological aspects of neuropsychological assessment is that personal preferences or individual goals are often not taken into account by clinicians and researchers and this may bias both scientific and clinical results. A way to deal with these problems is available in the form of Goal Attainment Scaling (GAS) [8]. In the final part of this thesis the practical use of GAS in people with brain dysfunction is investigated. The applicability of GAS in with cognitive disorders is examined in Chapter

7 by reviewing the available literature on the clinimetric aspects and the feasibility of GAS, when it is used as an outcome measure for psychogeriatric patients.

The review in Chapter 7 revealed that it is useful to involve the demented patient and/or caregiver in the goal-setting process. Moreover, GAS can cover multiple domains relevant to the psychogeriatric population such as cognition, daily life activities and mood. Inconsistent results were found for responsiveness, content validity, inter-rater reliability and construct/convergent validity. The number of studies that have investigated the use of GAS is small and they have been mainly initiated by one particular research group. Therefore it is too early yet to conclude that GAS is an applicable method for setting goals in this population.

The effectiveness of a post-acute outpatient community re-entry cognitive rehabilitation programme for patients with acquired brain injury and their relatives is investigated in Chapter 8. GAS was used as one of the primary outcome measures. Patients did improve significantly on individual goals ($p < 0.05$) between baseline (T0) and short-term follow-up (T1). The level of attainment did not change between T1 and long-term follow-up (T2). Goals were mostly set in the cognitive and behavioural domains. At T2 the level of social activities improved and patients reported less fatigue ($p < 0.05$). The other primary and secondary outcome measures showed no significant differences between T0, T1 and T2. Based on this result the use of GAS as an outcome method for such programmes is recommended.

Chapter 9 investigated the feasibility of GAS and the clinical experiences with GAS as an outcome method when it is used for the evaluation of cognitive rehabilitation in people with acquired brain injury. The use of GAS proved feasible: it proved possible to set three goals within an acceptable time-frame; to involve patients in the goal-setting procedure; to set realistic goals; and to set goals within relevant domains.

A strength of the longitudinal studies described in Chapter 8 and 9 was the low number of loss to follow-up. In Chapter 8 the percentage of loss to follow up for the GAS analysis was 0%, and 15% for the analyses with the questionnaires. For Chapter 9 the percentage was 2% (1 patient, recurrent CVA). These numbers are much lower than generally reported in comparable studies [9].

Role of the neuropsychologist

Traditionally, the role of the neuropsychologist has been to give support in diagnosing and localizing brain pathology. Nowadays the neuropsychologist is more frequently involved in questions about the daily life functioning of patients and their ability to return to work [10, 11]. In addition, the neuropsychological profile of a patient with dementia or an acquired brain injury is often of significant importance in the diagnostic trajectory [12].

When patients with brain dysfunction need rehabilitation to help them deal with the impairments resulting from their injury, knowledge about the underlying cognitive processes is necessary. Affected cognitive domains may interfere with non-affected domains and can result in problematic functioning. For example, a person's memory may function normally, but if his attention skills are affected it will be difficult for this person to remember things he did not pay attention to. These cognitive processes can be seen as a black box that can become manifest by testing cognition.

Neuropsychologists have the background and training required to open the black box and to interpret deficits in terms of the likely handicapping effect of those deficits in relation to everyday functioning. Standardized neuropsychological tests are used in a laboratory setting to identify patients' cognitive strengths and weaknesses. The neuropsychologist estimates the level of compensation skills and future functioning on the basis of the interactions between these strengths and weaknesses. The studies that include direct observation of patients performing daily life activities (Chapter 5 and 6), show that knowledge about the cognitive processes that underlie daily life activities provides relevant information for neuropsychologists which is not elucidated by laboratory tests alone.

Furthermore, knowledge about the level of independence of the patient is required. Occupational therapists have the knowledge and skills to identify patients' level of independence and to recommend occupational therapy. Occupational therapy is based on the Model of Human Occupation. This model seeks to explain the occupational functioning of persons. It focuses on how persons choose, order, and perform activities with regard to everyday occupational behaviour [13]. The occupational therapist uses several tools to observe daily life activities, such as the Assessment of Motor and Process Skills [14] and the Arnadottir OT-ADL Neurobehavioral Evaluation [15]. Occupational therapy is concerned with the assessment and rehabilitation of functional skills, including the understanding of the impact of cognitive deficits on daily life functioning [16].

The neuropsychological approach and the occupational approach can be used synergistically to aid the further rehabilitation of people with brain dysfunction. The functional cognitive assessment undertaken by the neuropsychologists will evaluate the cognitive deficits that interfere with daily life activities in order to determine the compensatory strategies that are required to reduce this interference. Information obtained from the direct observation of daily life activities can validate the predictions based on the tests [16].

This synergistic assessment procedure was the focus of the PROFINTEG project (Chapter 6). Its value is difficult to quantify, but all participants felt that the decomposition of daily life activities provides a better understanding of underlying neuropsychological aspects. This improved understanding supports individualized treatment. It is clear that neuropsychology makes unique contributions by offering a comprehensive understanding of brain-behaviour relationships and a critical evaluation of cognitive assessment and rehabilitation practices [17]. Neuropsychologists should be more aware of their distinctive skills and their specific contribution within the team. It is important to make effective use of their expertise when it comes to understanding how cognitive deficits interfere with daily functioning and to use their skills to interpret test results.

Ecological validity of neuropsychological tests and the role of direct observation

Our findings of a moderate association between neuropsychological tests and daily life functioning are not new. Remarkably, however, neuropsychologists still prefer to use these traditional neuropsychological tests when it comes to answering questions about the ability of patients to function without restrictions in daily life. Rabin et al. [18] investigated the actual use of non-traditional, ecologically oriented measures by neuropsychologists. Their

conclusions were surprising: approximately one third of the neuropsychologists in their sample used at least one ecologically oriented test. More important, these instruments were much less frequently used than traditional measures. Clinicians, who had worked most of their time in a rehabilitation setting, were more likely to use ecologically oriented instruments than those who did not work in such settings. Another remarkable result in the study of Rabin et al. was that there was no association between rehabilitation-related referral questions and the use of ecologically oriented instruments. These results imply that there still is a need to promote ecologically oriented instruments.

The ultimate approach with regard to measuring ecological validity does not exist. The assessment procedure in which a neuropsychological assessment was combined with the direct observation of patients performing complex activities in their own environment (Chapter 6) adds to the existing knowledge and provides interesting insights into the problem of ecological validity. This procedure is not a substitution of common practice but an improvement of what neuropsychologists normally do. The choice of the outcome measure is an inherent limitation of all research on the ecological validity of neuropsychological tests, as one can never really know the level of true everyday ability [19]. Direct observation of daily life activities is a method that comes very close to what neuropsychologists really want to measure when confronted with referral questions with regard to daily life functioning.

Individual goals

Since Goal Attainment Scaling (GAS) can be used as a method to measure the actual everyday functioning of an individual, it is tempting to state that it is an ecologically valid instrument. The individual character, or idiosyncratic nature of GAS is a strength in the case of clinical practice but it is a problem when it comes to conducting research since the uniqueness of GAS makes it difficult to test its validity [20].

Dementia versus acquired brain injury

This thesis presents the findings from studies of patients with dementia and of patients with acquired brain injury. While both of these disease entities are examples of brain dysfunction, these groups differ with respect to aspects such as prognosis, age, and severity of illness. Dementia is a neurodegenerative disease common in old age, while acquired brain injury is characterized by a stable course and mostly present in young and middle-aged adults. Differences exist also with regard to the affected domains: both groups are characterised by cognitive problems, but dementia is mainly characterised by memory problems. Moreover, cognitive impairments that interfere with IADL activities are part of the definition of dementia. In acquired brain injury the interference of cognition with daily life is often more subtle. Since people with acquired brain injury are often of working age, the consequences of their brain dysfunction are different in comparison to those of people with dementia who often are retired and frequently live in nursing homes.

Despite these differences, both populations have to deal with the impairments that result from the affected domains and the interaction between these domains. As this interaction is the main focus of the various investigations in this thesis, both groups were integrated in its scope.

Clinical implications

Importance of association between cognition and daily life functioning

It became clear in Chapter 2 that both cognitive as well as functional data are important for assessing patient care and for supporting diagnostic procedures in patients with brain dysfunction. Moreover, discrepancies between cognitive tests and instruments for measuring daily life activities provide important information that can guide the clinician in choosing the appropriate treatment. Therefore, clinicians need to pay attention to both domains when confronted with this population.

Improving ecological validity of neuropsychological assessment

As it is known now that the association between most traditional neuropsychological tests and everyday cognitive functioning is moderate, neuropsychologists should be encouraged to use other tools with higher levels of ecological validity when they are confronted with referral questions that deal with the daily life functioning of patients. Examples of such tools are neuropsychological tests developed with ecological validity in mind (verisimilitude) in combination with a detailed description and observation of the everyday activities of the individual patient.

To further improve the ecological validity of neuropsychological assessment, clinicians can state clear hypotheses about which cognitive functions are expected to be affected in everyday life based on the neuropsychological assessment, which can subsequently be tested (e.g. with direct observation). Clinical reasoning by testing a hypothesis enables the identification of neurobehavioral impairments that interfere with task performance and the understanding of factors that underlie functional dependence. Testing these hypotheses will guide the clinician in providing the appropriate treatment and education for patient and caregiver. Retrospectively, combining neuropsychological test data with observation of a patient's ability to perform everyday activities, improves the making of ecologically relevant predictions [16]. Moreover, combining neuropsychological and observational data will lead to more reliable predictions and to interventions that are better adapted to patients' needs.

Direct observation

Direct observation of everyday activities in a patient's own environment is not always necessary but certainly indicated when a patient lives alone or when the caregiver cannot give a clear view of how the patient acts. The IADL scale as presented in Chapter 6 might be a useful tool with which to identify those patients who complain about their cognitive difficulties and who are motivated to undergo rehabilitation for their impairments. These activities can subsequently be observed in the patient's environment.

Although it is time-consuming to observe patients in their own environment, it is useful to invest time and effort in these observations. Moreover, one video-recorded observation can be used by all the disciplines involved in the rehabilitation team, such as neuropsychology, occupational therapy, physiology as well as speech and language therapy, which is efficient. A videotape makes it possible to score performance after the observation instead of during the observation, which increases the reliability of scoring since it is difficult to score during an

observation. Furthermore, the findings based on direct observation can be used to set goals; during the observation problems might show up that need treatment and GAS can be used to operationalize these needs.

Recommendations for future research

This thesis provides insight into the ecological aspects of neuropsychological assessment. With regard to the association between on the one hand the several ways that exist for assessing dementia severity and on the other hand the consequences of dementia for daily life, it has become clear that a deeper understanding is required of the influence of neuropsychiatric symptoms as well as of the individual's and/or caregiver's preferences.

As the role of the neuropsychologist changes from one in which diagnosing and localizing brain pathology is central to one which focuses on the answering of questions related to daily life activities and the ability to return to work, ecologically oriented instruments that better support this changing role need to be developed. Furthermore, the inclusion of existing tests or approaches, such as the assessment procedure developed by the PROFINTEG project (Chapter 6), could be considered. Moreover, as Rabin et al. [18] showed that neuropsychologists are not willing to replace their use of standardized, traditional tests with instruments that use the verisimilitude approach, ecologically oriented instruments need to be promoted more vigorously. A way to accomplish this is to develop an up-to-date overview of existing ecologically oriented instruments which is made available to clinicians so that they are better informed about their options when they want to predict the daily life functioning of people with cognitive impairments due to brain dysfunction.

The next step with regard to the assessment procedure which was examined in Chapter 6, involves a further investigation of its benefits. If this procedure is to be applied in the future, insurance companies should be provided with a detailed overview of the procedure's costs and with studies showing its effectiveness.

One way of improving the statistical analysis of direct neuropsychological observation is to analyse one specific instrumental daily life activity within a large homogeneous population. To ensure a large research population, recruiting could take place across European countries. An example of a sophisticated technique for such a study is the multi-facet Rasch analysis. This technique was used to standardise the Assessment of Motor and Process Skills (AMPS) [21]. The multi-facet Rasch technique is an item response theory (IRT) model that measures latent traits, and thus focuses on individual elements of each variable [22].

An important inclusion criterion for the direct observation of instrumental daily life activities is that the activity needs to be relevant and important for the individual patient. The IADL scale as presented in Chapter 6 is a useful tool for identifying the relevance of these activities. Future research should use such tools when investigating daily life functioning of the individual patient. Moreover, as the presence of an observer might influence the patient's performance of an everyday activity and the activity is observed just once, the observational process could be improved and made more reliable by asking patients and/or caregivers to take notes about problematic activities in a diary for one week. This would include noting the level of help a caregiver had to provide and the level of subjective burden experienced by both the patient and the caregiver. A way to do this is to use a method such as the Experienced

Sampling Method (ESM) [23]. ESM is a one-day self-assessment technique [24] in which participants wear a watch that randomly beeps a few times a day. Participants are asked to write down their thoughts or feelings in a diary every time they hear the beep. ESM is frequently used in patients with schizophrenia.

In order to further investigate the effectiveness of a post-acute outpatient community re-entry cognitive rehabilitation programme for patients with acquired brain injury, a randomized controlled trial will be the next step.

Taking these ecological aspects of neuropsychological assessment into account will optimize the rehabilitation and future care of the individual patient with cognitive impairments.

References

1. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975;12(3): 189-198.

2. World Health Organization. International Classification of Functioning, Disability and Health. Geneva; 2001.

3. Verhey FR, Jolles J, Ponds RW, Rozendaal N, Plugge LA, de Vet RC, et al. Diagnosing dementia: a comparison between a monodisciplinary and a multidisciplinary approach. J Neuropsychiatry Clin Neurosci 1993;5(1):78-85.

4. Wolfs CAG, Kessels A, Dirksen CD, Severens JL, Verhey FRJ. Integrated multidisciplinary diagnostic approach for dementia care: randomised controlled trial. Br J Psychiatry 2008;192(4):300-305.

5. Burgess PW, Alderman N, Evans J, Emslie H, Wilson BA. The ecological validity of tests of executive function. J Int Neuropsychol Soc 1998;4(6):547-558.

6. Farias ST, Harrell E, Neumann C, Houtz A. The relationship between neuropsychological performance and daily functioning in individuals with Alzheimer's disease: ecological validity of neuropsychological tests. Arch Clin Neuropsychol 2003;18(6):655-672.

7. Manchester D, Priestley N, Jackson H. The assessment of executive functions: coming out of the office. Brain Inj 2004;18(11):1067-1081.

8. Kiresuk T, Sherman RE. Goal Attainment Scaling: a general method for evaluating comprehensive community mental health programs. Community Ment Health J 1968;4(6):443-453.

9. Corrigan JD, Harrison-Felix C, Bogner J, Dijkers M, Terrill MS, Whiteneck G. Systematic bias in traumatic brain injury outcome studies because of loss to follow-up. Arch Phys Med Rehabil 2003;84(2):153-160.

10. Kibby MY, Schmitter-Edgecombe M, Long CJ. Ecological validity of neuropsychological tests: focus on the California Verbal Learning Test and the Wisconsin Card Sorting Test. Arch Clin Neuropsychol 1998;13(6):523-534.

11. Sbordone RJ, Long CJ, editors. Ecological validity of neuropsychological testing. 2nd ed. Boca Raton: St. Lucie Press 1996.

12. Chaytor N, Schmitter-Edgecombe M. The ecological validity of neuropsychological tests: a review of the literature on everyday cognitive skills. Neuropsychol Rev 2003; 13(4):181-197.

13. Kielhofner G. Conceptual foundations of occupational therapy. Philadelphia: F.A. Davis Company 1992.

14. Fisher A. Assessment of motor and process skills. 3th edition. Fort Collins, CO: Three star press 1999.

15. Arnadottir. The Brain and behavior. Assessing cortical dysfunction through activities of daily living. St Louis: Mosby 1990.

16. Bennett TL. Neuropsychological evaluation in rehabilitation planning and evaluation of functional skills. Arch Clin Neuropsychol 2001;16(3):237-253.

17. Spooner DM, Pachana NA. Ecological validity in neuropsychological assessment: a case for greater consideration in research with neurologically intact populations. Arch Clin Neuropsychol 2006;21(4):327-337.

18. Rabin LA, Burton LA, Barr WB. Utilization rates of ecologically oriented instruments among clinical neuropsychologists. Clin Neuropsychol 2007;21(5):727-743.

19. Chaytor N, Schmitter-Edgecombe M, Burr R. Improving the ecological validity of executive functioning assessment. Arch Clin Neuropsychol 2006;21(3):217-227.

20. Donnelly C, Carswell A. Individualized outcome measures: a review of the literature. Can J Occup Ther 2002;69(2):84-94.

21. Fisher AG. The assessment of IADL motor skills: an application of many-faceted Rasch analysis. Am J Occup Ther 1993;47(4):319-329.

22. Stahl JA, Lunz ME. A comparison of generalizability theory and multi-faceted Rasch measurement. Midwest Objective Measurement Seminar; 1992; Chicago, IL; 1992.

23. De Vries MW. The experience of psychopathology: investigating mental disorders in their natural settings. Cambridge, UK: Cambridge University Press 1992.

24. Myin-Germeys I, Krabbendam L, Jolles J, Delespaul PA, van Os J. Are cognitive impairments associated with sensitivity to stress in schizophrenia? An experience sampling study. Am J Psychiatry 2002;159(3):443-449.

Summary

The main aim of this thesis was to investigate ecological aspects of neuropsychological assessment. This was done by exploring the association between domains such as cognition and daily life functioning, affected in people with brain dysfunction (e.g. dementia and acquired brain injury). The general introduction (Chapter 1) provides background information on the affected domains in patients with brain dysfunction, ecological validity and its importance for research and clinical practice and individual goal-setting. Furthermore, the rationale, the aims and the outline of the thesis are presented.

The first part of the thesis describes the association between cognition and daily life functioning. The second part of the thesis focuses on the ecological validity of neuropsychological assessment and the integration of cognitive, emotional and behavioural domains and daily life functioning. The third part describes the setting of individual goals and the level of attainment of these goals for patients with cognitive impairments due to psychogeriatric disorders or acquired brain injury.

Chapter 2 describes the association between cognitive status and daily life functioning in people with different types of dementia. Cross-sectional data were used from 615 patients with dementia who were referred to the Maastricht Memory Clinic of the Maastricht University Medical Centre. Scores on the Mini-Mental State Examination showed strong correlation with Blessed Dementia Scale scores in cases of frontotemporal dementia (r = -0.80); moderate correlation in cases of Alzheimer's disease (AD), vascular dementia (VaD), and Parkinson's dementia (PD) (range r = -0.50 to - 0.60); while no correlation was found in primary progressive aphasia (PPA) cases. The association between cognition and daily life functioning varied among dementia subtypes for AD, VaD, FTD and PD. Furthermore, the overall scores on both domains differ between dementia subtypes, indicating that different types of dementia are characterized by a specific pattern of cognitive status and daily life functioning. These findings underline the need for multidomain assessment in patients with dementia.

Chapter 3 presents the investigation of the association between cognition and direct observation of daily life functioning in a psychogeriatric population with cognitive disorders and to examine the influence of neuropsychiatric symptoms on this association. Cross-sectional data of 118 patients with cognitive disorders were used; data on cognition (MMSE, CAMCOG), global severity (GDS), daily life functioning (IADL), and neuropsychiatric symptoms (NPI) were collected and analyzed using correlation and regression analyses. Different combinations of the severity measures were tested for their ability to predict the AMPS process ability scores. Scores on the MMSE, CAMCOG and GDS were moderately associated with the AMPS process ability score. These measures explained between 27% and 44% of the variance in the AMPS score. The presence of apathy influenced the association between the cognitive measures and the AMPS score. Commonly used measures of dementia severity are only moderately associated with observation of performance on daily activities. This underlines the need for direct observation of daily activities in dementia patients.

Chapter 4 is an introduction to Part II and provides a description on the investigation of the ecological validity of neuropsychological tests. From research it appears that the association between neuropsychological tests and daily life functioning is moderate. This implies that daily life functioning of people with brain dysfunction cannot be explained

solely by neuropsychological tests. Several factors (e.g. mood, testing environment) may influence the association between neuropsychological tests and daily life functioning. Direct observation of patients in their own environment may provide useful additional information to plan treatment and future care complementary to neuropsychological tests.

In Chapter 5 the ecological validity of neuropsychological assessment is investigated by exploring the extent to which experienced neuropsychologists can predict daily life functioning of people with acquired brain injury based on neuropsychological data only. Eight neuropsychologists were asked to make predictions about daily life functioning of brain-injured patients based on neuropsychological test data, data on emotion, behaviour and personality. The level of agreement between neuropsychologists was calculated and the neuropsychologists' predictions were compared to the data provided by videotapes of observation of the participants performing one or more relevant, familiar and important daily life activities in their own environment. A clinical researcher scored the performance of the activities. The level of agreement between neuropsychologists about expected problems during the performance of observed daily life activities varied considerably. The accuracy of the predictions was 60%. The only category with a sufficient accuracy of the predictions was the overall performance of an activity (i.e. success or failure). This exploratory study showed that the ecological validity of neuropsychological tests is enhanced when taking the neuropsychologists' decision-making process into account. However, a substantial part of daily life functioning remains unexplained by the neuropsychological assessment. Direct observation of a patient performing daily life activities provides important and unique information, not elucidated by neuropsychological data only and is therefore recommended in addition to neuropsychological test assessment.

Chapter 6 describes an international cooperative project called PROFINTEG that attempts to enhance the ecological validity by integrating neuropsychological data and observational data in patients with brain dysfunction. The sensitivity of the neuropsychological assessment is moderate (77%) with regard to identifying problematic daily life activities and its specificity is low (15%). This means that these tests cannot discriminate well enough between patients who show problems in daily life and who do not. The qualitative description showed that the integrated approach aided the treatment of the patients since the approach provided relevant information not elucidated by the neuropsychological test assessment. In conclusion, the integrated approach has added value for identifying the impact of neuropsychological deficits on concrete and everyday life activities in people with brain dysfunction.

Chapter 7 provides a systematic review on the feasibility of Goal Attainment Scaling (GAS) in a psychogeriatric population with cognitive disorders. GAS is a method to set individual goals and measures the level of attainment of the goals. Eight databases were screened on literature about the clinimetric aspects and the feasibility of GAS when used as method to measure level of attainment of individual goals for psychogeriatric patients with cognitive disorders. Two authors independently reviewed all the data. Ten studies were included. Mixed results were found for clinimetric aspects such as responsiveness, content validity, inter-rater reliability and construct/convergent validity. GAS was feasible with regard to the possibility to involve the patient and/or caregiver in the goal-setting procedure and with regard to the possibility to implement multiple domains. The feasibility aspects that were less

well established are the possibility to set at least three realistic goals per patient in less than 30 minutes and the involvement of a blinded assessor. GAS proved to be useful on important aspects of an outcome method for psychogeriatric patients with cognitive disorders. Since other relevant aspects showed mixed results and the number of studies investigating the use of GAS in psychogeriatric patients with cognitive disorders is small, the evidence is not strong enough yet to state that GAS is an applicable method in this population.

Chapter 8 shows the results of a prospective cohort study on the effectiveness of a post-acute outpatient community re-entry cognitive rehabilitation programme for patients with acquired brain injury and their relatives. GAS was used as one of the outcomes to investigate whether patients attain their predefined goals after the intervention. Participants were twenty-seven patients with acquired brain injury and their relatives. The primary outcome measures were cognitive failures (cognitive failure questionnaire, CFQ), quality of life (stroke adapted sickness impact profile, SA-SIP) and individualized goals (Goal Attainment Scaling, GAS). There were no significant differences on CFQ; the SA-SIP physical functioning first deteriorated significantly (T1), but improved at follow-up (T2). At T2 the level of social activities improved and patients reported less fatigue ($p < 0.05$). Patients did improve significantly on individual goals ($p < 0.05$) between T0 and T1. The level of attainment did not change between T1 and T2. Goals were mostly set in the cognitive and behavioural domains. The programme has a positive effect on the individual goals set by the patients. This effect did not result in a better participation or a better quality of life of either the patients or their relatives.

Chapter 9 presents the results of a prospective observational longitudinal study to examine the feasibility of, and clinical experiences with Goal Attainment Scaling (GAS) when used for the evaluation of cognitive rehabilitation in people with acquired brain injury. Forty-eight consecutive patients with acquired brain injury enrolled during a 15-month period. The mean (SD) age of the patients was 46.1 (10.7) years; 29 (60%) were male; 186 goals were set with a mean (SD) number of 4 (1) goals per patient. It was possible to set at least three realistic goals per patient within 30 minutes. Most goals were set in the cognitive domain (i.e. memory and attention), followed by the behavioural domain (i.e. fatigue and aggression). It proved possible to set three goals within an acceptable time-frame, to involve patients in the goal-setting procedure, to set realistic goals, and to set goals within relevant domains. Based on clinical experiences, GAS is less feasible for research when patients lack insight, or suffer from comorbidity or mood problems.

Chapter 10 reflects the general discussion of the results of the studies described in this thesis. Methodological and theoretical considerations are described. Clinical implications are provided and recommendations for further research are made.

Samenvatting

Het doel van dit proefschrift is het evalueren van ecologische aspecten van neuropsychologisch onderzoek. Hiertoe werd het verband onderzocht tussen de verschillende domeinen die aangedaan zijn bij mensen die hersenschade hebben opgelopen ten gevolge van een niet-aangeboren hersenletsel of neurodegeneratieve aandoeningen zoals dementie. Voorbeelden van de aangedane domeinen zijn cognitief functioneren en dagelijks functioneren. Hoofdstuk 1 voorziet in achtergrond informatie met betrekking tot de aangedane domeinen bij mensen met hersenschade, ecologische validiteit en het belang hiervan voor onderzoek en het stellen van individuele doelen. Verder worden de rationale, de doelen en de opbouw van het proefschrift gepresenteerd.

In het eerste deel van het proefschrift wordt het verband tussen cognitie en dagelijks functioneren onderzocht (Hoofdstuk 2 en Hoofdstuk 3). In het tweede deel van het proefschrift wordt gekeken naar de ecologische validiteit van neuropsychologisch onderzoek en de integratie van domeinen (Hoofdstuk 4, Hoofdstuk 5 en Hoofdstuk 6). Het derde deel beschrijft het opstellen van individuele doelen en de mate waarin deze doelen behaald worden bij mensen met cognitieve problemen ten gevolge van psychogeriatrische aandoeningen of niet-aangeboren hersenletsel (Hoofdstuk 7, Hoofdstuk 8 en Hoofdstuk 9).

Hoofdstuk 2 beschrijft het verband tussen de cognitieve status en het dagelijks functioneren van mensen met verschillende typen dementie. Cross-sectionele data werden gebruikt van 615 patiënten met dementie die verwezen waren naar de geheugenpoli van het Universitair Medisch Centrum in Maastricht. Scores op de Mini-Mental State Examination (MMSE) toonden een sterk verband met scores op de Blessed Dementia Scale bij patiënten met frontotemporale dementie (FTD) (r = -0.80); een matig verband bij patiënten met de ziekte van Alzheimer (ZvA), vasculaire dementie (VaD), en Parkinson's dementie (PD) (range r = -0.50 tot - 0.60); terwijl er geen verband werd gevonden bij patiënten met primaire progressieve afasie (PPA). Het verband tussen cognitie en dagelijks functioneren varieerde dus tussen ZvA, VaD, FTD en PD. Verder verschilden de overall scores op beide domeinen tussen de verschillende typen dementie. Dit impliceert dat verschillende typen dementie gekarakteriseerd worden door een specifiek patroon van cognitieve status en dagelijks functioneren. Deze bevindingen onderstrepen de behoefte aan een multidomein aanpak bij mensen met dementie.

In hoofdstuk 3 wordt ingegaan op onderzoek naar het verband tussen maten voor de ernst van dementie en directe observatie van de uitvoering van alledaagse activiteiten bij mensen met psychogeriatrische problemen. Tevens werd de invloed van neuropsychiatrische symptomen op dit verband bekeken. Cross-sectionele data van 118 patiënten met cognitieve stoornissen werden gebruikt; gegevens over cognitie (Mini-Mental State Examination (MMSE); Cognitive component of the revised Cambridge Examination for Mental Disorders of the Elderly (CAMCOG), globale mate van ernst (Global Deterioration Scale (GDS)), dagelijks functioneren (Instrumental Activities of Daily Living Scale (IADL)) en neuropsychiatrische symptomen (Neuropsychiatric Inventory (NPI)) werden verzameld en geanalyseerd met correlatie en regressie analyse. Directe observatie van alledaagse activiteiten was de onafhankelijke variabele (Assessment of Motor and Process Skills (AMPS; proces vaardigheden). Scores op de MMSE, CAMCOG en GDS waren matig geassocieerd met de scores op de AMPS proces vaardigheden. Deze maten konden tussen de 27 % and 44% van

de variantie op de AMPS verklaren. De aanwezigheid van apathie beïnvloedde het verband negatief tussen cognitieve maten en de AMPS score. Veelgebruikte maten voor de ernst van dementie zijn matig geassocieerd met directe observatie van het uitvoeren van alledaagse activiteiten. Deze bevinding benadrukt de behoefte aan directe observatie van het uitvoeren van alledaagse activiteiten van mensen met dementie.

Hoofdstuk 4 is een introductie voor het tweede deel van het proefschrift en geeft een beschrijving van het begrip ecologische validiteit van neuropsychologische tests. Uit de literatuur blijkt dat het verband tussen neuropsychologische tests en dagelijks functioneren matig is. Dit impliceert dat het dagelijks functioneren van mensen met hersenschade niet volledig verklaard kan worden door neuropsychologische tests alleen. Verschillende factoren (bijvoorbeeld stemming en testomgeving) kunnen het verband tussen neuropsychologische tests en dagelijks functioneren beïnvloeden. Directe observatie van patiënten in hun eigen omgeving kan voorzien in aanvullende informatie om behandeling en toekomstige zorg te plannen complementair aan neuropsychologische tests.

Hoofdstuk 5 behandelt het onderzoek naar de ecologische validiteit van neuropsychologisch onderzoek door te exploreren in welke mate ervaren neuropsychologen het dagelijks leven van mensen met niet-aangeboren hersenletsel kunnen voorspellen op basis van enkel neuropsychologische gegevens. Acht neuropsychologen werden gevraagd voorspellingen te doen over het dagelijks functioneren van patiënten met niet-aangeboren hersenletsel op basis van neuropsychologische test data, gegevens over emotie, gedrag en persoonlijkheid. De mate van overeenstemming tussen de neuropsychologen werd berekend en de voorspellingen van de neuropsychologen werden vergeleken met de gegevens van gefilmde directe observatie van patiënten tijdens de uitvoering van een of meerdere relevante en bekende dagelijkse activiteiten in de thuisomgeving van de patiënt. De mate van overeenstemming tussen de neuropsychologen over de verwachte problemen tijdens de uitvoering van de geobserveerde activiteiten varieerde aanzienlijk. Het percentage accurate voorspellingen was 60%. De enige categorie met een voldoende percentage aan juiste voorspellingen was het uiteindelijke resultaat van de activiteit (wel of niet slagen in de uitvoering). Deze explorative studie laat zien dat de ecologische validiteit van neuropsychologische tests verhoogd wordt wanneer het beslissingsproces van de neuropsycholoog meegenomen wordt. Echter blijft een substantieel gedeelte van het dagelijks leven onverklaard door het neuropsychologisch onderzoek. Directe observatie van een patiënt tijdens de uitvoering van een dagelijkse activiteit geeft belangrijke en unieke informatie die niet aan het licht komt wanneer er alleen neuropsychologisch onderzoek gedaan wordt. Daarom is directe observatie aanbevolen naast het neuropsychologisch test onderzoek.

Hoofdstuk 6 beschrijft de ontwikkeling en de implementatie van een geïntegreerde aanpak waarin neuropsychologische tests en functioneren in het echte dagelijkse leven worden gecombineerd. De geïntegreerde aanpak bestaat uit een neuropsychologisch onderzoek, de identificatie van dagelijkse activiteiten en directe observatie van mensen met hersenschade tijdens het uitvoeren van een relevante activiteit in zijn of haar eigen omgeving. De benadering is ontwikkeld in samenwerking met onderzoekers uit België en Nederland. De kwantitatieve beschrijving van de benadering zoals beschreven in dit hoofdstuk is gebaseerd op Nederlandse patiënten met niet-aangeboren hersenletsel; de kwalitatieve beschrijving van de benadering is

gebaseerd op Belgische patiënten met dementie. De sensitiviteit van het neuropsychologisch onderzoek is matig (77%) met betrekking tot het identificeren van problematische dagelijkse activiteiten. De specificiteit is laag (15%). Dit resultaat betekent dat deze tests niet voldoende kunnen discrimineren tussen patiënten die problemen laten zien in het dagelijks leven en zij die geen problemen laten zien. De kwalitatieve beschrijving laat zien dat de geïntegreerde benadering de behandeling ondersteunt aangezien de benadering voorziet in relevante informatie die niet aan het licht komt door neuropsychologisch testonderzoek. De geïntegreerde benadering heeft toegevoegde waarde voor het identificeren van de impact van neuropsychologische tekorten op concreet en alledaags functioneren van mensen met hersenschade.

Hoofdstuk 7 laat de resultaten zien van een systematische review naar de bruikbaarheid van Goal Attainment Scaling (GAS) in een psychogeriatrische populatie met cognitieve stoornissen. GAS is een methode om individuele doelen te stellen en de mate waarin deze doelen zijn behaald te meten. Acht databases zijn afgezocht op literatuur over klinimetrische aspecten en de bruikbaarheid van de GAS wanneer deze gebruikt werd als uitkomstmaat voor psychogeriatrische patiënten met cognitieve stoornissen. Twee auteurs hebben onafhankelijk van elkaar de data bekeken. Dit resulteerde in de inclusie van tien studies. Inconsistente resultaten werden gevonden voor de responsiviteit, content validiteit, inter-rater validiteit en construct/convergente validiteit. Het is mogelijk gebleken de patiënt en/of diens naaste te betrekken bij het opstellen van de doelen. Daarnaast is gebleken dat doelen binnen verschillende domeinen kunnen worden opgesteld. De mogelijkheid om ten minste drie realistische doelen per patiënt op te stellen in minder dan 30 minuten behoeft verder onderzoek evenals de betrokkenheid van een geblindeerde onderzoeker bij het meten van het niveau van de behaalde doelen. GAS bleek bruikbaar te zijn op belangrijke aspecten voor een uitkomstmaat bij psychogeriatrische patiënten met cognitieve stoornissen. Aangezien andere relevante aspecten inconsistente resultaten lieten zien en het aantal studies dat het gebruik van de GAS onderzocht bij patiënten met psychogeriatrische stoornissen klein is, is er nog onvoldoende bewijs om te zeggen dat de GAS een bruikbare uitkomstmaat is in deze populatie.

Hoofdstuk 8 laat de resultaten zien van een prospectieve cohort studie naar de effectiviteit van een poliklinisch cognitieve revalidatie programma voor mensen met niet-aangeboren hersenletsel en hun naasten. GAS werd gebruikt als één van de uitkomstmaten om te bepalen of de deelnemers hun vooraf opgestelde individuele doelen behaalden na de interventie. Zevenentwintig patiënten en hun naasten deden mee. De primaire uitkomstmaten waren cognitieve vergissingen (Cognitive Failure Questionnaire, CFQ), kwaliteit van leven (Stroke Adapted Sickness Impact Profile 30, SA-SIP30) en individuele doelen (Goal Attainment Scaling, GAS). Er werden geen significante verschillen gevonden op de CFQ en de SA-SIP30 en de secundaire uitkomstmaten op alle meetmomenten. De patiënten gingen wel significant vooruit op hun individuele doelen ($p < 0.05$). De meeste doelen werden gesteld binnen de cognitieve en gedragsmatige domeinen. Het programma had een positief effect op de individueel gestelde doelen bij patiënten. Dit effect resulteerde niet in een betere participatie of een betere kwaliteit van leven bij zowel de patiënten en hun naasten.

Hoofdstuk 9 laat de resultaten zien van een cross-sectionele studie naar de toepasbaarheid

van de GAS als uitkomstmaat bij cognitieve revalidatie programma's voor mensen met niet-aangeboren hersenletsel. Doelen werden opgesteld met patiënten die deelnamen aan een cognitief revalidatie programma. Achtenveertig patiënten namen deel aan het onderzoek. Er werden tezamen 186 doelen opgesteld met een gemiddelde van 4 doelen per deelnemer. Het was mogelijk om ten minste drie realistische doelen te stellen per patiënt binnen een half uur. De meeste doelen werden opgesteld in het cognitieve domein (bijvoorbeeld geheugen en aandacht), gevolg door gedrag (bijvoorbeeld vermoeidheid en agressie). Met GAS bleek het mogelijk te zijn om drie doelen op te stellen binnen een acceptabel tijdpad, om de patiënten te betrekken bij het opstellen van hun eigen doelen, om realistische doelen te stellen en om doelen op te stellen binnen relevante domeinen. Gebaseerd op klinische ervaringen is de GAS minder geschikt als uitkomstmaat voor onderzoek wanneer patiënten geen inzicht in hun eigen problemen hebben, er comorbiditeit is, of wanneer er stemmingsproblemen zijn.

In hoofdstuk 10 worden de belangrijkste resultaten van de studies in dit proefschrift bediscussieerd. Er wordt ingegaan op methodologische en theoretische overwegingen. Klinische implicaties worden gegeven, evenals aanbevelingen voor verder onderzoek.

Publications

Bouwens SFM, van Heugten CM. Visuo-spatieel neglect: hoe te meten? Een kritisch overzicht van de literatuur. Neuropraxis 2005;4:109-116.

Bouwens SFM, van Heugten CM, Verhey FRJ. Ecological validity of neuropsychological tests: what do they say about functioning in daily living? [Abstract]. Fourth Dutch Endo-Neuro-Psycho (ENP) Meeting, Doorwerth, The Netherlands, June 2005.

Bouwens SFM, van Heugten CM, Aalten P, Wolfs C, Baarends E, van Menxel D, et al. Predictive value of different severity measures for interference with daily life functioning of psychogeriatric patients [Abstract]. Fifth Dutch Endo-Neuro-Psycho (ENP) Meeting, NVN Satellite Symposium, Doorwerth, the Netherlands, June 2006.

Bouwens SFM. Goal Attainment Scaling (GAS) [Abstract]. Werkgroep CVA Nederland, Utrecht, The Netherlands, April 2007.

Kempen GI, Meier AJ, Bouwens SFM, van Deursen J, Verhey FRJ. Psychometrische aspecten van de Nederlandstalige versie van het Telefonische Interview Cognitieve Status (TICS). Tijdschr Gerontol Geriatr 2007;38(1):38-45.

Bouwens SFM. Rechtstreekse ergotherapeutische observatie bij dementie: de toegevoegde waarde van de AMPS [Abstract]. Karel V Symposium, Alzheimer Research Nederland, Utrecht, The Netherlands, September 2007.

Bouwens SFM, van Heugten CM, Aalten P, Wolfs CAG, Baarends EM, van Menxel DAJ, et al. Relationship between measures of dementia severity and observation of daily life functioning as measured with the Assessment of Motor and Process Skills (AMPS). Dement Geriatr Cogn Disord 2008;25:81-87.

Bouwens SFM, van Heugten CM, Scholte PAJM, Verhey FRJ. Do demented patients benefit from rehabilitation for their everyday problems? [Abstract]. 36th Annual International Neuropsychological Society (INS) Meeting. Waikoloa, Hawaii, USA, February 2008.

Bouwens SFM, van Heugten CM, Verhey FRJ. Psychometric properties of Goal Attainment Scaling used for people with cognitive disorders: a systematic review [Abstract]. 36th Annual International Neuropsychological Society (INS) Meeting. Waikoloa, Hawaii, USA, February 2008.

Bouwens SFM, van Heugten CM, Verhey FRJ. Het belang van directe observatie bij patiënten met dementie. Tijdschrift voor Neuropsychiatrie en Gedragsneurologie 2008;7(3):69-71.

Bouwens SFM, van Heugten CM, Verhey FRJ. Review of Goal Attainment Scaling as a useful outcome measure in psychogeriatric patients with cognitive disorders. Dement Geriatr Cogn Disord 2008;26(6):528-540.

Bouwens SFM, van Heugten CM, Verhey FRJ. Goal Attainment Scaling as an outcome measure for people with acquired brain injury who receive cognitive rehabilitation. Clin Rehabil (In press).

Bouwens SFM, van Heugten CM, Verhey FRJ. Association between cognition and daily life functioning in dementia subtypes. Int J Geriatr Psychiatry (In press).

Bouwens SFM, van Heugten CM, Ponds RWHM, Verhey FRJ. The ecological validity of neuropsychological assessment with regard to daily life functioning after acquired brain injury: a different approach (Submitted).

Rasquin SMC, Bouwens SFM, Dijcks B, Winkens I, Bakx WG, van Heugten CM. Effectiveness of a post-acute outpatient community re-entry cognitive rehabilitation programme for patients with acquired brain injury (Submitted).

List of abbreviations

AAT	Aachen Aphasia Test
ABI	Acquired Brain Injury
ACL	Aphasia Check List
AD	Alzheimer's Disease
ADAS-cog	Alzheimer's Disease Assessment Scale Cognitive Subscale
AMPS	Assessment of Motor and Process Skills
ANOVA	Analysis of Variance
AVLT	Auditory Verbal Learning Test
BADL	Basic Activities of Daily Living
BADSL	Bristol Activities of Daily Living Scale
BADS	Behavioural Assessment of the Dysexecutive Syndrome
BAFQ	Brock Adaptive Functioning Questionnaire
BCET	Biber Cognitive Estimation Test
BCRS	Brief Cognitive Rating Scale
BDS	Blessed Dementia Scale
BI	Barthel Index
BNT	Boston Naming Test
CAMCOG	Cognitive component of the CAMDEX
CAMDEX	Cambridge Examination for Mental Disorders of the Elderly
CDR	Clinical Dementia Rating
CDS	Cornell Depression Scales
CES-D	Center for Epidemiologic Studies Depression (Scale)
CFQ	Cognitive Failure Questionnaire
CFT	Complex Figure Test
CGI	Clinical Global Impression
CIBIC-plus	Clinician's Interview-Based Impression of Change-Plus caregiver input
CIND	Cognitive impairment, No Dementia
CIQ	Community Integration Questionnaire
CIRS	Cumulative Illness Rating Scale
CMT	Contextual Memory Test
COGNISTAT	Neurobehavioural Cognitive Status Examination
COWAT	Controlled Oral Word Association Test
CPS	Cognitive Performance Scale
CSI	Caregiver Strain Index
CVLT	California Verbal Learning Test
DEX	Dysexecutive Questionnaire
DOC-PG	Diagnostic Observation Centre for PsychoGeriatric patients
DRS	Disability Rating Scale
DS	Dementia Scale
DSM	Diagnostic and Statistical Manual of Mental Disorders
EMS	Episodic Memory Scale
ES	Effect Size
FAI	Frenchay Activities Index

FAQ	Functional Activities Questionnaire
FIM	Functional Independence Measure
FRS	Functional Assessment Rating Scale
FSS	Fatigue Severity Scale (FSS)
FTD	Fronto Temporal Dementia
GARU	Geriatric Assessment and Rehabilitation Unit
GAS	Goal Attainment Scaling
GAU	Geriatric Assessment Unit
GCOR	Global Clinical Outcome Rating
GDS	Global Deterioration Scale
GIT	Groninger Intelligence Test
GPT-D	Grooved Pegboard Test, Dominant hand
HABAM	Hierarchical Assessment of Balance and Mobility
HADS	Hospital Anxiety and Depression Scale
IADL	Instrumental Activities of Daily Living
ICC	Intraclass Correlation Coefficient
ICF	International Classification of Functioning, Disability, and Health
ICHII	Iowa Collateral Head Injury Interview
ILS	Independent Living Scales
IMC	Information-Memory-Concentration
IQCODE	Informant Questionnaire on Cognitive Decline
IRR	Inter-rater Reliability
JOLO	Judgement of Line Orientation
KADL	Katz Activities of Daily Living Index
LBD	Lewy Body Dementia
LNS	Letter-Number Sequencing
LOTCA	Loewenstein Occupational Therapy Cognitive Assessment
LR	Likelihood Ratio
MDRS	Mattis Dementia Rating Scale
MDS-PAC	Minimum Data Set for Post-Acute Care
MEDICIE	Maastricht Evaluation of a Diagnostic Intervention for Cognitively Impaired Elderly
MFRS	Montebello Rehabilitation Factor Score
MGAT	Mobile Geriatric Assessment Team
MMC	Maastricht Memory Clinic
MMSE	Mini-Mental State Examination
MUMC	Maastricht University Medical Centre
NHP	Nottingham Health Profile
NINCDS-ADRDA	National Institute of Neurological Diseases and CommunicativeDisorders and Stroke, as well as the Alzheimer's Disease and Related Disorders Association
NPI	NeuroPsychiatric Inventory
OARS IADL	IADL subscale of the Older Americans Resources and Services
PD	Parkinson's Dementia

PPA	Primary Progressive Aphasia
PROFINTEG	Professional Integration
PSMS	Physical Self-Maintenance Scale
RCF	Rey Complex Figure
RCT	Randomized Controlled Trials
RE	Relative Efficiency
SAILS	Structured Assessment of Independent Living Skills
SC	Star Cancellation
SCL-90	Symptom Checklist 90 items
SCWT	Stroop Colour Word Test
SD	Standard Deviation
SMMSE	Standardized Mini-Mental State Examination
SPMSQ	Short Portable Mental Status Questionnaire
SQLI	Spitzer Quality of Life Index
SRM	Standardized Response Mean
TAP	Tailored Activity Program
TEA	Test of Everyday Attention
TMT	Trail Making Test
UCL	Utrecht Coping List
VaD	Vascular Dementia
WAIS	Wechsler Adult Intelligence Scale
WCST	Wisconsin Card Sorting Test
WMS-R	Wechsler Memory Scale-R

Dankwoord

Na mijn klinische stage en onderzoeksstage in revalidatiecentrum Hoensbroeck in 2004 kreeg ik de mogelijkheid om promotieonderzoek te gaan doen aan de Universiteit Maastricht. Enige aarzeling om me in dit onderzoek te storten kan ik niet ontkennen, maar juist de klinische inslag trok mij over te streep.

Met veel patiëntcontacten, speurwerk in databases en schrijfwerk achter de rug, ben ik trots op het huidige eindwerk. Het voltooien van dit proefschrift was hard werken, maar niet alleen voor mijzelf. De komende ruimte wil ik dan ook gebruiken om een woord van dank uit te spreken naar iedereen die op wat voor manier dan ook een bijdrage heeft geleverd.

Op de eerste plaats wil ik alle deelnemers, hun partners, kinderen en andere familieleden die aan mijn onderzoek hebben deelgenomen hartelijk danken. Zonder jullie inzet zou het maar een karig proefschrift zijn geworden. Ik vond het hartverwarmend om te zien hoe gastvrij ik werd ontvangen bij jullie thuis, zelfs wanneer ik de camera tevoorschijn haalde om het dagelijks leven te filmen.

Dan mijn promotieteam: promotor Frans Verhey en co-promotor Caroline van Heugten. Een gouden duo! Frans, bedankt voor je geweldige begeleiding. Er zijn maar weinig promovendi die zo'n intensief contact met hun promotor hebben. Ik heb ontzettend veel van je geleerd. Je humor werkte relativerend en maakt dat ik met veel plezier terugkijk op de afgelopen vier jaar. Caroline, bedankt dat je mijn co-promotor wilde zijn! Geen vraag was je teveel, ik kon altijd bij je terecht en je bent een echte motivator. Onze Hawaiiaanse whale watch en de hula zal ik niet snel vergeten. Ik voel me bevoorrecht om jou en Frans als promotieteam te hebben.

De leden van de beoordelingscommissie wil ik hartelijk danken voor hun kritische beoordeling en goedkeuring.

Eric, Steve, Martine, Patrick, Sabine, Catherine, Vincianne, Anne et Françoise: je vous remercie de notre coopération du projet PROFINTEG.

Derick, thank you for our inspiring conversations and for being part of my exam committee. I appreciate your invitation to join the expert meeting on goal planning in London last year.

Nico, ontzettend bedankt voor je hulp als ik weer eens niet kon inloggen in mijn mailbox, mijn Mac voor de zoveelste keer 'raar' deed, mijn oude Apple wel heel erg traag was bij het invoeren van gegevens in SPSS of als ik voor de zoveelste keer riep dat ik niet kon printen en je dan altijd als eerste vroeg of ik de netwerkkabel wel ingeplugd had... Op momenten dat Nico er niet was, was Ron de klos. Memorabele momenten waren dat ik laat in de middag nog snel even alles wilde back-uppen, en de hele handel weer eens vast liep. Ron, bedankt voor je geduld.

Sascha, jij bent mede degene geweest die me over de streep getrokken heeft om te gaan promoveren. Ik heb vaak met je aan de telefoon gehangen om alle ins en outs van het 'aio-zijn' te bespreken. Ook toen de stap in het onderzoek eenmaal gezet was, heb ik nog veel gehad aan jouw adviezen. We hebben zelfs samen een onderzoek gedaan, waarvan de resultaten in dit proefschrift beschreven staan.

Pauline, bedankt voor je begeleiding in de beginfase van mijn project. Bij het schrijven van mijn eerste artikel stond ik vaak bij je op de stoep. Ik heb veel geleerd van je schrijfadviezen. Jij wist altijd de zinnen net wat korter en krachtiger te formuleren zodat het artikel wél voldeed aan het maximum aantal woorden.

Claire, Erica, en Debby, erg leuk dat jullie betrokken wilden zijn bij mijn eerste artikel. Daar zijn de nodige versies van voorbij gekomen die jullie ook allemaal van commentaar hebben voorzien.

Danny, bedankt voor je hulp bij het testen en de huisbezoeken, het was een gezellige tijd. Ik vond het een eer om je te mogen toespreken bij je diploma-uitreiking. Gisela, wat heb jij een berg werk verzet (en nog). Leuk om te zien ook dat jij ook gekozen hebt voor promotieonderzoek. Dat gaat je vast goed af! Daarnaast ook een woord van dank aan mijn andere stagiaires/onderzoeksassistenten/apnio's: Inge, Climmy, Athena, Carlijn, Marlien, Trudi, Petra en Kristina, bedankt voor jullie inzet.

Sven, het was erg prettig samenwerken met jou. Ik heb veel gehad aan je steun en luisterend oor, maar zeker ook aan onze gezellige gesprekken over de welbekende promotieperikelen en mijn verdere loopbaan.

Ingrid en Bert, bedankt voor onze interessante gesprekken. Jeanne, Marleen, Heleen en alle andere medewerkers van revalidatiecentrum Blixembosch die hun steentje hebben bijgedragen aan het slagen van het onderzoek aldaar: bedankt!

Henk, Tonny, Math, Rita en alle andere medewerkers van het BreinGrepen project in Hoensbroek: bedankt, het is een mooi project geworden.

Peter van de Sande, bedankt voor je steun en vertrouwen in de fase na mijn promotie.

Rudolf, Marjolein, Yvonne, Jacqueline, Dymphie, Jeanette, Annelien, Brechje, hebben jullie nog zin om filmpjes te bekijken? Bedankt voor de leuke discussies.

Carla, voor jou was geen vraag over tests, normen, Y-factoren en Z-factoren teveel. Bedankt!

Els, ontzettend bedankt voor je vertaalwerk. En natuurlijk voor het soepeltjes laten verlopen van alle afspraken met Frans.

Anky, Laurent, Marco, Judith en Andrü, ik ben blij dat jullie over de centen gingen en dat ik alleen maar braaf ieder kwartaal mijn urenstaatjes hoefde in te dienen.

Wim, Lia en Dick, bedankt voor jullie statistische adviezen.

Willemien en Desiree, super dat jullie het proefschrift zo grondig hebben willen lezen voordat het naar de drukker ging.

Dan nu in willekeurige volgorde allerlei mensen die in welke vorm dan ook er toe hebben bijgedragen dat ik zover ben gekomen om mijn proefschrift verdedigingsklaar te krijgen of juist voor de nodige afleiding hebben gezorgd: Ieke, Chantal, Karin, Martine, Tanja, Pernette, Jane, Angelique, de ergotherapeuten van de afdeling ergotherapie azM, Floortje, Heidi, Heidi, Ed, Sanne, Nikki, Elissa, Annemarie, Celeste, Natalie, Renske, Inez, Rosa, Rosa, Deliane, Sil, Jennifer, Linda, Candy, Jelle, Renate, Martin, Elsa, Inez, Netty, Lisbeth, Esther, Esther, Olga, Leonore, Aukje, Astrid, Anja, Geert, Bart, Eric, Anita, Janneke, Joyce, Lyzel, Margriet, Sebastian, Beatrice en mijn Tilburgse en Eindhovense vrienden.

Ik ben blij dat mijn kleine rode autootje het 13.000 km lang heeft volgehouden om door Brabant en Limburg te scheuren op weg naar het zoveelste huisbezoek.

Jacky, het is een mooi boekje geworden, super bedankt voor het lay-outen. Het was een behoorlijke klus.

Har Naus, de Charleston heeft een mooie plek gekregen in de woonkamer.

Peter Molenaar en Klaartje, Hawaii was top!

Elsbeth, bedankt voor je gezelligheid, steun en niet te vergeten je sportiviteit.

Saartje, super dat je mijn paranimf wilt zijn! Jouw nuchtere kijk op de wereld heeft me erg geholpen de boel te relativeren. Jammer dat we elkaar nu wat minder zien, maar dat mag de pret niet drukken. Het voelt goed om je achter me te hebben staan bij de laatste horde tot het doctorschap.

Jeroen, dat wij als Brabo's tussen al die Limbo's (en alle andere niet-Brabo's) al snel op dezelfde golflengte zaten werd snel duidelijk; Mengvoeders United dreef andere roomies tot wanhoop. Vanaf de dag dat jij als aio begon, kwamen we tegenover elkaar op 1 kamer te zitten. Omdat na twee jaar de gezelligheid wel erg hoog opliep (en we met de jaren wijzer werden), hebben we de tafels omgedraaid zodat we niet meer de hele dag tegen elkaar aan hoefden te kijken. De productiviteit schoot omhoog. Jeroen, bedankt voor de fijne aio-tijd! Dat je mijn paranimf zou worden, was vanaf de eerste dag al duidelijk.

Zussie, bedankt voor het vertalen van de Franstalige stukken! Papa en mama, bedankt voor alle steun en interesse in mijn onderzoek. Mama, bedankt voor het vele corrigeerwerk. Jullie tripjes naar Maastricht waren altijd gezellig.

Lieve Bram, bedankt voor je onvoorwaardelijke steun en je vele inspanningen bij het doorlezen en corrigeren van de ontelbare versies van ieder hoofdstuk. Zonder jou was het niet zo'n mooi proefschrift geworden! Nu is het mijn beurt om jou te helpen bij de beproeving die promoveren heet. Bovenal is het tijd om samen te genieten van alle dingen die niets met promoveren te maken hebben!

Curriculum vitae

Sharon Bouwens werd geboren op 6 februari 1982 te Oirschot. In 2000 behaalde ze haar VWO diploma aan het Eckart College te Eindhoven. Hierna ging ze Psychologie studeren aan de Universiteit van Tilburg. In 2004 studeerde ze af in de major Klinische Gezondheidspsychologie en minor Klinische Neuropsychologie. Haar onderzoekszoekstage vond plaats binnen het Instituut voor Revalidatie Vraagstukken (iRv, nu Vilans) en revalidatiecentrum Hoensbroeck, beiden in Hoensbroek. Van 2004 tot 2008 werkte ze als promovenda bij de vakgroep Psychiatrie en Neuropsychologie van de Universiteit Maastricht. Het onderzoek uit die periode is beschreven in dit proefschrift. Sinds 2005 is ze bestuurslid van de Nederlandse Vereniging voor Neuropsychologie. Sinds december 2008 is ze werkzaam als psycholoog binnen de afdeling Medische Psychologie van het Laurentius Ziekenhuis te Roermond. Daarnaast werkt ze als onderzoeker bij het Clinical Trial Center Maastricht.