DEVELOPMENT AND VALIDATION OF AN INTELLIGENT COGNITIVE ASSESSMENT SYSTEM (ICAS) FOR PERSONS WITH CEREBRAL VASCULAR ACCIDENT (CVA)

YIP CHI KONG

Ph.D

THE HONG KONG POLYTECHNIC UNIVERSITY

2010

THE HONG KONG POLYTECHNIC UNIVERSITY

DEPARTMENT OF REHABILITATION SCIENCES

Development and Validation of an Intelligent Cognitive Assessment System (ICAS) for Persons with Cerebral Vascular Accident (CVA)

Yip Chi Kong

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

September 2009

CERTIFICATE OF ORIGINALITY

I hereby declare that this thesis is my own work and that, to my best knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

____________________ (Signed)

YIP Chi Kong ________ (Name of student)

DEDICATION

To my dearest father, mother and wife

ABSTRACT

Background: Advancing technology and the rapidly increasing use of personal computers have speeded up the development of innovative assessment procedures and at a lower and affordable cost. One of the possible applications in rehabilitation is the use of computer-adaptive testing (CAT) for administering testing items that are adaptable to the patient's ability level. CAT was proposed as an effective means to provide accurate and quick screening of cognitive deficits in persons with stroke in the present study. It was further developed into an Intelligence Cognitive Assessment System (ICAS) that enhances outcome prediction by adding artificial intelligence.

In this project, ICAS was designed to be embedded with three special features that are not all found in typical cognitive assessments. Firstly, ICAS is a CAT designed for a comprehensive assessment of cognitive abilities for stroke survivors. Secondly, ICAS's scoring system has been developed by modern psychometrics, using the Rasch model, in arriving at a linear ratio scale on which scores can be compared directly at different time points or between different patients. Thirdly, artificial neural networking (ANN), an artificial intelligence approach, was used to reinforce ICAS's predictive ability with regard to functional outcomes in stroke survivors. The aim of this project was thus to develop and validate this newly developed ICAS for stroke rehabilitation.

Method: Three operational phases of study were conducted to achieve specific objectives of the project. Phase I was to investigate the content validity of the assessment items of ICAS. An expert panel review of the ICAS software and its trial run among 14 stroke survivors were initially carried out. This phase also served as a pilot study to provide preliminary data for implementation of Phases II and III. In Phase II, the test item difficulty measures, item structure (construct validity), and item stability of the ICAS were investigated. Cognitive functions of another group of 30 stroke subjects were assessed by the ICAS and by the Chinese versions of the Mini-Mental Status Examination (MMSE-CV) and the Neurobehavioral Cognitive Status Examination (NCSE-CV) respectively. Phase III investigated the psychometric properties of the ICAS and built up an ANN model for predicting functional outcomes of stroke survivors which was based on the ICAS results and other demographical characteristics. The cognitive functions of a third batch of 66 subjects were assessed by both the ICAS and MMSE-CV. Demographics and clinical data such as age, gender, types of stroke, lesion side, residual upper limb function, and residual self-care function (as indicated by the initial post-stroke Modified Barthel Index or MBI) were collected together with the ICAS score. They served as predictors to forecast the MBI value at discharge stage using a specific ANN model.

Result: In the Phase I study, the content validity of the ICAS was established and the Intraclass Correlation Coefficient (ICC(2,k)) among the panel members for the agreement with content relevance was 0.972 ($p < 0.01$). In addition, 58 out of 65 ICAS testing items got good to excellent rating in the content relevance rated by the panel members. In the Phase II study, the Rasch analysis of the 65 testing items revealed that the item difficulty measures of the ICAS ranged from –4.3 to 5.8. Only 3 items fell outside the INFIT statistics with criteria from 0.6 to 1.3. If the criteria were readjusted to 0.5 to 1.5, all the items fitted the INFIT criteria. For the OUTFIT statistics, 10 of the items fell outside the range at the 0.5 to 1.5 level. However, the principle component analysis of the residual revealed that 66.1% of the variance could be explained by the model. The unexplained variance explaining the first contrast was 3.7%. These findings indicated that the ICAS testing items were unidimensional in nature. The stroke subjects' abilities were also found to be statistically significant and highly correlated with the MMSE-CV and the NCSE-CV scores. In the Phase III study, the correlation of ICAS with MMSE-CV was 0.757 ($p < 0.001$). Both the test–retest reliability of ICAS (Cronbach's alpha = 0.878) and the correlation of the test–retest (0.789; $p < 0.001$) were satisfactory. Finally, a cutoff score of 3.02 was found to be able to determine the existence of cognitive impairment, indicating sensitivity of 80.5% and a false positive rate of 4%. In the ANN prediction model,

there was a high correlation between the observed discharge MBI value and the model predicted discharge MBI value (correlation coefficient = 0.85; $p < 0.001$).

Conclusion: The results suggested that the ICAS items fit the Rasch model and items are unidimensional to measure the cognitive functions for stroke survivors. Moreover, the Rasch based cognitive ability score is a linear ratio scale which might be as valid and useful as MMSE-CV and NCSE-CV in measuring cognitive function in stroke patients. Secondly, the ANN prediction model was found to be effective in predicting the functional outcomes after stroke rehabilitation, based on demographic data and residual cognitive and physical functions. These pieces of information can be useful for treatment planning and to predict a home discharge programme for better recovery and/or reintegration into the community. Thirdly, the psychometric properties of the ICAS were initially established. It can be an efficient alternative tool in determining cognitive impairment in stroke survivors for rehabilitation and related research studies. Lastly, future study can be further validated by increasing the test items in ICAS and among other neuro-disability groups.

PUBLICATIONS ARSING FROM THE THESIS

Publication

Yip, C.K. & Man, D.W.K. (in press). Validation of the Intelligent Cognitive Assessment System (ICAS) for Stroke Survivors. *Brain Injury.*

Yip, C.K. & Man, D.W.K. (under review). Development of an Intelligent Cognitive Assessment System (ICAS) using the Rasch model. *Psychological Assessment.*

Yip, C.K. & Man, D.W.K. (2009). Validation of a computerized cognitive assessment system for persons with stroke: a pilot study. *International Journal of Rehabilitation Research,* 32(3), 270-278.

Conference Presentation

Yip, C.K. & Man, D.W.K. (2009). Item stability in the Computer Cognitive Assessment System (CCAS). Oral presentation at Pacific Rim Objective Measurement symposium 2009, July 28 – 30, Hong Kong.

Yip, C.K. & Man, D.W.K. (2008). Psychometric properties and application of a computerized cognitive assessment system (CCAS) for patient with stroke. Oral presentation at The 5th World Congress for NeuroRehabilitation, Sept 24-27, Brazil. Neurorehabilitation and Neural Repair, 22(5), 558.

Yip, C.K. & Man, D.W.K. (2008). Computer adaptive test (CAT) for persons with stroke: protocol development. Oral presentation at The 5th World Congress for NeuroRehabilitation, Sept 24-27, Brazil. Neurorehabilitation and Neural Repair, 22(5), 557.

Yip, C.K. & Man, D.W.K. (2008). Prediction of functional outcomes by artificial neural network (ANN) for persons with stroke. Oral presentation at The 5th World Congress for NeuroRehabilitation, Sept 24-27, Brazil. Neurorehabilitation and Neural Repair, 22(5), 557.

Yip, C.K. & Man, D.W.K. (2008). Validity and Reliability of Cognitive Computer Adaptive Test for Patient with Stroke. Paper presented at Hospital Authority Convention 2008, May 5-6, Hong Kong. Hospital Authority Convention 2008 Program book, p38.

Yip, C.K. & Man, D.W.K. (2008). Road to cognitive computer adaptive test. Oral presentation at Pacific Rim Objective Measurement symposium 2008, July 31 – Aug 3, Tokyo.

Yip, C.K. & Wong, Y.C. (2007). Prediction of functional outcome for patient with stroke by artificial neural network model. Oral presentation at The 4th Asia Pacific Occupational Therapy Congress, June 23-26, Hong Kong.

Yip,C.K. & Man,D.W.K. (2007). Development of computer adaptive cognitive assessment. Paper presented at The 4th Asia Pacific Occupational Therapy Congress, June 23-26, Hong Kong.

Yip, C.K. & Man, D.W.K. (2007). Intelligent cognitive assessment system (ICAS) – A better way to assess. Oral presentation at 2007 World Congress on Ageing and Dementia in Chinese Communities, Mar 7-10, Hong Kong. Asian Journal of Gerontology & Geriatrics, 2(1), 35-36.

ACKNOWLEDGEMENTS

I would like to express my gratitude to Dr. David Man – my chief supervisor, for his kindly guidance and extensive support during the entire study.

I would also like to thank Mr. Steven Chong – Department Manager of Occupational Therapy of TWGHs Wong Tai Sin hospital, for his fully support to conduct research in the hospital and the departments.

Besides, thank you very much to Miss Elaine Chan, Mr. Ray Li, Mr. Victor Wong, Miss Stella Chan, Miss Catherine Chan, without their help, this study cannot be conducted smoothly.

Finally, I wish to thank the subjects for their participation in the study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ADL	Activities of Daily Living
AI	Artificial Intelligent
ANN	Artificial Neural Network
AUC	Area Under The Curve
BP	Back-Propagation
BP–ANN	Back-Propagation Artificial Neural Network
CAT	Computer Adaptive Test
CIND	Cognitive Impairment No Dementia
CTT	Classic Test Theory
CVA	Cerebral Vascular Accident
FTHUE	The Functional Test for Hemiplegic Upper Extremity
FTHUE-HK	The Functional Test for Hemiplegic Upper Extremity Hong Kong Version
HCI	Human Computer Interaction
ICAS	Intelligent Cognitive Assessment System
IRT	Item Response Theory
MBI	Modified Barthel Index
MCI	Mild Cognitive Impairment
MLP	Multi-Layer Perceptrons
MMSE	Mini Mental Status Examination
MMSE – CV	Chinese Version of Mini Mental Status Examination
NCSE	Neurobehavioral Cognitive Status Examination
NCSE – CV	Chinese Version of Neurobehavioral Cognitive Status Examination
NIHSS	National Institutes of Health Stroke Scale
OT	Occupational Therapists
PCA	Principal Components Analysis
TBI	Traumatic Brain Injury
TIA	Transient Ischemic Attack
ULFT	Upper Limb Function Test

Chapter 1 Introduction

Stroke is the most important etiology of neurological disability and handicap in the Western world (Rosamond et al., 2007). In Hong Kong, stroke is rated the fourth leading disease-related cause of death. There were a total of 26,167 hospital admissions for stroke events. The recent age-specific mortality rate for age 65 or above was 1,602 per 100,000 (Hospital Authority, 2008). Stroke survivors usually face disabilities in motor, sensory, perceptual, and cognitive functioning. They may also face a lifelong struggle with cognitive dysfunctions (Desmond, Moroney, Sano, & Stern, 1996) including vision, memory, attention processes, spatial orientation, problem solving, behaviour management, and emotional difficulties such as anxiety and depression (Gourlay, Lun, Lee, & Tay, 2000; Zhang et al., 2001). These cognitive deficits may also interfere with daily living and reemployment (Lee et al., 2003). Moreover, stroke survivors may have a higher risk of dementia (Tatemichi, Desmond, Stern, Sano, & Bagiella, 1994), especially the older stroke survivors (Lowery et al., 2002). Around 25% of them were reported to have developed dementia at twelve months post-stroke (Barba et al., 2000 ; Linden, Skoog, Fagerberg, Steen & Blomstrand, 2004). Moreover, the prevalence rates of post-stroke cognitive impairment were found to be around 30% to 40% from 3 months to 3 years post-stroke (Patel, Coshall, Rudd, & Wolfe, 2003). A Hong Kong study also showed that 69% of stroke survivors developed cognitive impairment as indicated by the

Chinese version Mini Mental Status Examination (MMSE-CV) using a cutoff score of 19 or below (Luk, Chiu, & Chu, 2008).

Cognitive impairment after stroke can lead to profound functional limitation (Wheatly, 2001). The presence of cognitive problems is crucial in recovery from stroke, and can be a predictor of functional outcome (Paolucci et al., 1996). A strong correlation was reported to exist between cognitive status and rehabilitation success, with better outcomes being achieved in cognitively intact elderly stroke survivors (Heruti et al., 2002; Luk, Chiu, & Chu, 2008). Therefore, accurate and effective cognitive assessment is deemed important in successful rehabilitation planning for stroke survivors (Neistadt, 1994; Donovan, Kendell, Heaton, Kwon, Velozo & Duncan, 2008). Specific clinical reasons for assessing cognition include baseline measurement to monitor change for treatment planning, problem identification, and discharge planning, and to identify patients who would benefit more from detailed neuropsychological evaluation (Radomski, 2008). In addition, assessment results could be used as a predictor for discharge planning.

However, existing cognitive assessments have drawbacks that deserve attention. First of all, most traditional standardized cognitive assessments worth conducting usually require a set of assessment tools, have parallel forms, need to specify the age range, and may be restricted to certain diagnostic groups of the test-takers. The results are usually norm referenced and need to be interpreted by professionals or trained personnel after the test-takers have completed the tests. The interpretation may be complicated and time consuming. Secondly, even though the administration procedures of the assessment are standardized, test administrators could make human errors during the assessment procedures and therefore the actual abilities of test-takers may be underestimated or overestimated. In addition, the results in the format of presenting subcomponents in cognitive functions are good for general treatment planning but not as good as scoring methods for research. A composite score cannot serve as a quick reference by designating medical and rehabilitation team in clinical decision-making. Thirdly, most of these assessments have been developed in Western countries. Cultural discrepancy is inevitable when these tests are applied to the Eastern countries and will affect the reliability of the results. Finally, routine tests and pre-trial screening of high-risk elderly people for cognitive impairment are constrained by the limitations of currently available cognitive function tests (de Jager, Budge, & Clarke, 2003). Typical assessments may be constrained in testing people with cognitive disabilities due to the "learning effect" (by repeating the test multiple times and having no parallel forms); the "ceiling effect": presented items are too easy, especially for high-functioning individuals (Pasquier, 1999; Cullum et al., 2000; de Jager et al., 2002), and the "floor effect": the presented test items are too difficult. The key reasons are that the design of those tests may not consider the test-takers' ability levels and the assessor may choose an inappropriate test for the test-takers. Therefore, there is a pressing need for a simpler, shorter, but comprehensive assessment of cognitive functions for the purpose of longitudinal tracking and monitoring of treatment effects (Simon, Doniger, Dimant, & Dwolatzky, 2007).

With the advance of technology proliferating, clinical information automation has become available to improve the documentation process and sharing of medical information between medical and allied health professionals, and thus can provide a complete medical record efficiently. The application of computers as an assessment tool in the medical field has been an interesting research topic of the past decade. The reliability, validity, user-acceptance, and cost-effectiveness of computer assessments have been well documented and equivalence with traditional methods has been shown (Handel, Ben-Porath, & Watt, 1999; Schulenberg & Yutrzenka, 1999; Webb, Zimet, Fortenberry, & Bylythe, 1999; Epstein & Rotunda, 2000; Epstein & Klinkenberg, 2001; Yamanaka et al., 2005). Many recently developed computerized versions of assessments were based on their paper-and-pencil versions. Unfortunately, they merely served as electronic versions, instead of being computerized assessments, as they did not fully utilize the available multimedia characteristics of a computer. In short, a computerized assessment is better able to present test items to the test-taker according to his or her abilities; that is, the assessment could be adapted to the test-taker's abilities. Moreover, it can transform different cognitive constructs into a score in a linear unidimensional scale for easy comparison, and thus facilitate the rehabilitation team's treatment planning. Furthermore, through the easy-to-use and powerful calculation abilities of modern personal computers, a model of a non-linear relationship could also be used to predict treatment outcomes accurately. These accurate findings could greatly enhance treatment monitoring, discharge planning, and welfare resources management.

Thus, the present study aimed to develop and validate a newly computerized cognitive assessment programme labelled as an intelligent cognitive assessment system (ICAS) for stroke survivors. ICAS is computer adaptive test (CAT) software (which is adapted to stroke patients' responses), has a test-item bank, and can present test items according to patients' cognitive abilities. It can be run on any typical personal computer with a Windows XP[1] operating system and equipped with a touch screen and speaker. Based on the Rasch model,[2] it can transform different cognitive domain scores into an overall score. Finally the ICAS score could be used to predict the post-stroke functional outcomes together with other documented predictors (e.g.

---

[1] Windows XP is a trademark of Microsoft corporation

[2] The Rasch model is a model of Item Response Theory and will be discussed further in Chapter 2

physical factors) through the use of artificial neural networking (ANN).[3]

To achieve the aim of the study, there were four specific objectives which have been completed in three continuous phases of the study:

- The Phase I study developed and validated the items bank of ICAS.
- The Phase II study investigated the item difficulty and unidimensionality of the items bank of ICAS for stroke survivors.
- ICAS's psychometric properties and the optimal cutoff score for stroke survivors with cognitive impairment were obtained in the Phase III study. Building up and validation of an ANN to predict the functional outcomes of stroke survivors were also accomplished.

In this study, ICAS has integrated the advantages of computer assessment, modern test theory, and ANN. It has utilized fully the multimedia, simulation, and calculation powers of a computer to provide stimulating rich environments that facilitate cognitive function assessment of stroke survivors. Easy-to-interpret results can be obtained and serve as an indicator of both outcomes and progress. ICAS was also used for correlation with the recognized gold standards of cognitive assessments. Its additive advantage when compared with the traditional gold standard cognitive assessment could thus be demonstrated. It was proposed that:

---

[3] ANN is part of artificial intelligence, which will be further discussed in Chapter 2

1. ICAS is an efficient cognitive assessment as it utilizes computer adaptive testing techniques.
2. As it provides immediate results which are transformed from different cognitive components into a score on a linear ratio scale, it can then serve as an outcome indicator for both stroke survivors' progress and programme documentation.
3. It is equipped with multimedia features that simulate daily tasks as assessment materials and thus can improve its ecological validity.
4. It is equipped with an ANN prediction model, which is good at dealing with non-linear relationship modelling for better forecasting of patients' functional outcomes.

There are a total of six chapters in this research report. After this introductory chapter, the next chapter will give a literature review on the relevant components for developing the ICAS such as cognition and cognitive deficits after stroke, assessment approaches, computerized assessments, use of human computer interaction, the item response theory in the design of ICAS, and the use of artificial intelligence in outcome prediction. Chapter 3 illustrates a proposed structural model for the development of the ICAS and Chapter 4 states the research methodology and data analysis of each of three phases of study. Chapter 5 presents the results and Chapter 6 presents the overall discussion of the study and recommendations for further study.

Chapter 2 Literature Review

In developing the intelligent cognitive assessment system (ICAS) for stroke survivors, the basic question is how to make sure it is valid, relevant, user-friendly, and can enhance cognitive assessment in clinical settings. Therefore, a thorough and critical review of the current literature pertaining to cognition, post-stroke cognitive deficits, comparison of cognitive assessment approaches, rationale for using computerized assessment, the advantages of modern testing theory and the ANN system should be conducted. This chapter presents the summary of this process and demonstrates the pertinent need for a more sophisticated tool for the assessment of cognitive functions for stroke survivors.

Stroke and its impact on cognitive function

Stroke is the second most common cause of acquired cognitive impairment and dementia and contributes to cognitive decline in the neurodegenerative dementias (O'Brien et al., 2003; Linden, Skoog, Fagerberg, Steen & Blomstrand, 2004). Cognitive impairments are incapacitating consequences of ischaemic stroke. Surprisingly, cognition is most often omitted from models' prediction outcomes during the early stage after ischaemic stroke (Kalra & Crome, 1993) even though it is an important predictor for the successful treatment and rehabilitation (Paolucci et al, 1996; Khateb, Annoni, Lopez, Bernasconi, Lavanchy & Bogousslavsky, 2007). Stroke frequently produces cognitive dysfunctions (Desmond et al, 1996). More than a third of stroke survivors failed four or more of the neuropsychological measures whereas only 4% of stroke-free controls recorded the same result. Among those who failed four or more of the neuropsychological measures, half suffered from stroke related dementia (Tatemichi, Desmond, Stern, Paik, Sano, & Bagiella, 1994).

A longitudinal study of prevalence rates of cognitive impairment following stroke found residual cognitive impairments in more than a third of stroke survivors when followed up at three months, one year, two years, and three-year intervals (Patel, Coshall, Rudd, & Wolfe, 2003). Older stroke survivors are at even higher risk of developing dementia (Lowery, Ballard, & Rogers et al., 2002). In a cohort study of older adults without stroke, transient ischaemic attack (TIA), or dementia, cognitive function, and incident cognitive decline were associated with risk of stroke (Elkins, O'Meara, Longstreth, Carlson, Manolio, & Johnston, 2004). Twenty-five percent of stroke survivors develop dementia within 12 months of having a stroke (Barba, Martinez-Espinosa, Rodriguez-Garcia, Pondal, Vivancos, & Del Ser, 2000). Similar condition was found in Hong Kong, the cognitive impairment among non-demented post stroke patients is common, 21.8% have cognitive impaired but no dementia (Tang et al, 2006) and 20% patient developed dementia in post stroke three month (Tang, et al, 2004).

In a study investigating the neuropsychological characteristics of mild vascular cognitive impairment and dementia after stroke (Stephens et al, 2004), there was no significant cognitive impairment between stroke patients and the normal control group. There were however, significant differences in executive function and speed of processing and perception. Meanwhile, in a comparison of stroke patients with no cognitive impairment and with vascular cognitive impairment no dementia (CIND), the domains with significant differences were memory, executive functioning, and language expression. In further comparison of vascular CIND with post-stroke dementia, the domains with significant differences were abstract thinking and executive functioning. The results indicate that attention and executive function are frequent impairments in stroke patients, but deficits of memory, orientation, and language are more indicative of CIND and dementia.

The most frequently reported post-stroke cognitive consequences were attention deficit, aphasia, short-term memory deficit, executive dysfunction, and long-term memory dysfunction (Lesniak, Bak, Czepiel, Seniow, & Czlonkowska, 2008). Cognitive domains were found to have significant predictive validity with respect to functional outcome including intellectual function, language, memory, perception and visuospatial construction, attention, and psychomotor-function (Zandvoort, Kessels, Nys, Haan, & Kappelle, 2005). This is of interest as orientation and attention are components of primary cognitive capacity and deficits of these components following stroke have been found to predict functional outcomes (Pedersen, Jorgensen, Nakayama, Raaschou, & Olsen, 1996; Robertson, Ridgeway, Greenfield, & Parr, 1997). This includes decreased performance of basic and higher level activities of daily living (ADL) (Pedersen et al., 1996). Specifically, deficits in focused attention have been shown to pose particular challenges for patients with regard to engaging in effortful, extended cognitive rehabilitation and reintegrating into social and vocational activities (Anderson, Winocur, & Palmer, 2003). Memory, the third component of primary cognitive capacity, has also been shown to have significant predictive validity with respect to functional outcomes following stroke (Zandvoort, Kessels, Nys, Haan, & Kappelle, 2005). The impact of stroke on executive functions has also been documented (Robertson, Ridgeway, Greenfield, & Parr, 1997). Sustained attention is closely related to executive function and could predict the recovery of function (Manly & Robertson, 1997). Rapport and co-workers (1998) suggested that the influence of motor and sensory impairments on falls was moderated by executive functions. In another community-based study, a first-ever stroke of mild to moderate severity was associated with a significant risk of cognitive impairment at three months, even in the absence of clinical aphasia (Srikanth, Thrift, Saling, Anderson, Dewey, Macdonell, & Donnan, 2003).

Similar findings in Hong Kong were also shown by Chan and his colleagues (2002). They investigated the cognitive profile of Chinese post stroke patients by using the Chinese version of the Neurobehavioral Cognitive Status Examination (NCSE-CV). The stroke survivors performed significantly more poorly than their control counterparts on the NCSE-CV. The stroke patients performed most poorly in orientation, attention, and calculation sub-tests; however, the low level of literacy of the local elderly population and language structure tended to slightly alter their cognitive profiles (Chan, Lee, Fong, Lee, & Wong, 2002).

After a brief review of cognitive deficits after stroke, the next section will briefly outline what cognition is for the purpose of developing ICAS testing items.

What is cognition?

Cognition is all the mental processes which allow us to perform meaningful activities in everyday living (Grieve & Gnanasekaran, 2008). It can be broadly defined as the acquisition and use of knowledge (Neisser, 1967; Grieve & Gnanasekaran, 2008) and it is the ability of the brain to process, store, retrieve, and manipulate information (Prigatano & Fordyce, 1986). Hence, cognition is crucial in supporting an individual's self-identity, roles, tasks, and occupations (Duchek & Abreu, 1997). Cognition consists of an interactive hierarchy (Ben-Yishay, cited in Goldstein & Levin, 1987) Attention, orientation, and memory are the basis of higher cognitive function. Higher cognitive functions include a fund of knowledge, the ability to manipulate old knowledge, social awareness, judgement, and abstract thinking (Strub & Black, 1977).

Cognition can be classified into three levels: primary cognitive capacity, higher level thinking ability, and meta processing ability or executive function (Radomski, 2008). The components of primary cognitive capacity are orientation, attention, and memory including the sensory registry and short- and long-term memory. These are the prerequisites for higher-level thinking and influencing the meta-processing (Radomski, 2008). The components of higher-level thinking are reasoning, concept formation, and problem solving, while the components of executive functioning are volition, planning, purposeful action (initiation of action), and effective performance (Radomski, 2008). The individual components of each level of cognition cited above are outlined in Appendix I.

After reviewing the cognitive deficits after stroke, we will thoroughly review the assessment approaches in the next section.

Review of assessment approaches for cognitive functions

A large variety of approaches or methods are currently used to evaluate cognitive functions. They can be broadly divided into standardized assessment and non-standardized assessment. Standardized cognitive assessments are designed to assess specific cognitive functions and are administered in a standardized fashion as defined by their test manuals. They are the main cognitive assessment tool used and are also known as psychometric tests. Such tests have normative scores with which patient performance can be compared. Standardized assessments can be classified further in terms of their assessment strategies. This is further elaborated later in this section. Non-standardized assessments differ from standardized assessments. The methods include interviews and observation to detect the signs and symptoms based on the brain pathology or behavioural neurology. Using non-standardized assessments, a skilled assessor can gather a large amount of information. It is possible to identify deficits in a number of cognitive domains through spending time talking with a patient. Short bedside tests can also be useful in highlighting the presence of certain cognitive disorders. However, the severity of impairment and certain forms of more subtle cognitive dysfunction may not be detected by these assessment approaches (Evans, 2003).

The main cognitive assessment tools are standardized or known as psychometric tests. This means that the testing items and tasks were designed to assess specified cognitive functions and are administered in a standardized fashion defined in a test manual. Such tests have been given to a normative sample group so that the performance of the patients can be compared.

As addressed earlier in this section, standardized assessment can be classified from the viewpoint of assessment strategy. This includes the bottom-up performance component approach, top-down functional approach, and integration of the two approaches.

Bottom-up performance component approach (Duchek & Abreu, 1997)

As cognitive function is composed of several different components, assessments may include a battery of subtests, each of which measures different components of cognitive functioning. This approach typically reflects a bottom-up performance component emphasis wherein cognition is reduced to subcomponents and the assessment of the subcomponents is used to create an in-depth analysis of specific deficit areas. The bottom-up approach thus reflects assessment at an impairment level of analysis but not at the functional level.

Within the bottom-up performance component approach, Goldstein (1987) describes neuropsychological tests that can be further classified as either comprehensive or specialized. Comprehensive assessments include a battery of subtests which measure various components of cognitive, psychomotor, and perceptual functions. Specialized assessments include tests which measure the same aspects of more specific cognitive components, such as attention, memory, and so on.

Top-down functional approach (Duchek & Abreu, 1997)

The focus of the top-down functional approach is at the macro or functional level. Cognition is not reduced to subcomponents in order to identify specific deficit areas. Instead, cognitive function is inferred from everyday activities in order to identify abilities and deficits in occupational performance. Functional cognitive assessment is often administrated through non-standardized methods such as interviews, observation, and the performance of functional tasks. However, in recent decades, standardized assessments based on the top-down functional approach have been developed. Their ratings are based on the patients' performance of functional tasks in accordance with well defined and standardized guidelines and rating scales. This assessment process requires the rater to be appropriately trained before administering the assessment in order to accurately identify the cognitive impairments.

An integrative functional approach (Duchek & Abreu, 1997)

This approach interfaces the bottom-up approach and top-down approach with the assumption that a relationship exists between the specific components and performance at a functional level and that this relationship is neither unidirectional nor causal (Abreu, Duval, Gerber, & Wood, 1994). An example is the association between the cognitive component of attention and the functional performance of driving a car. This relationship can therefore be conceptualized as a multidimensional and complex assessment.

In summary, different assessment approaches in standardized assessment have their advantages and disadvantages. The bottom-up performance component approach assesses cognitive functions at an impairment level but the results are difficult to generalize to the functional level. The deficit components are identified but no information is given on which functional tasks the client can perform safely. On the other hand, using a top-down functional approach, the functions of the test-taker are well assessed but the rater needs special training in administrating assessment, which is based on this kind of assessment approach for ensuring the standardization of assessment procedures and interpretation of results from the functional level to the impairment level. Even though the integrative functional approach gathers the advantages of the previous two assessment approaches, raters using this assessment approach still need special training. However, when we engage the computer in the integrative functional approach (computerized integrative functional approach), the assessment procedures can be standardized for a function run in a computer-simulated environment. There can be pre-set rating criteria, stopping criteria, and interpretation criteria for the computer to execute. As a result, it could remove the disadvantages including the need for special training for administration of the assessment and interpretation of the result. Different assessment approaches are also compared in terms of standardized procedure, interpretation of results, and level of assessment results (see Table 1 for details).

Table 1 Comparison of different assessment approaches

	Bottom-up performance component approach	Top-down functional approach	Integrative functional approach	Computerized integrative functional approach
Administration procedure	Easy to administer in standardized procedure	Need special training to administer in standardized procedure	Need special training to administer in standardized procedure	Easy to administer in standardized procedure
Interpretation of result	Interpretation according to test manual	Interpretation requires special training	Interpretation requires special training	Interpretation by computer
Level of assessment result	Impairment level	Functional level; requires special training in impairment level	Both functional and impairment level	Both functional and impairment level

Computerized assessment

The application of the computer as an assessment tool in the medical field is constantly being enhanced by the rapid advancement of computer technology. Therefore, it is not surprising that over the past decade there has been increasing research interest in this field. It is well documented that computerized assessments are reliable, valid, user-friendly, and cost effective. They have also been shown to be equivalent to traditional methods (Handel, Ben-Porath, & Watt, 1999; Webb, Zimet, Fortenberry, & Blythe, 1999; Schulenberg & Yutrzenka, 1999; Epstein & Rotunda, 2000; Epstein & Klinkenberg, 2001; Yamanaka et al., 2005). Examples of well documented computerized cognitive assessments including Cambridge Neuropsychological Test Automated Battery (CANTAB) (Morris, Evenden, Sahakian & Robbins, 1987), Cognitive Drug Research (CDR) (Simpson, Surmon, Wesnes & Wilcock, 1991) and Mindstreams^{TM} (NeuroTrax Crop, NY).

The application of new technologies is most successful when it affords a capability superior to the procedures that it is replacing. Technological success also depends upon interface consistency and core transportable functions that readily permit searching, editing, integrating, visualizing, displaying, and storing (Chute, 2002). Application of computer technology in elderly had been demonstrated effective in delivery of knowledge, medication instruction and they reported satisfactory to computer-based technology (Rippey et al, 1987; Leirer et al, 1988; Ogozalek, 1993). Elderly clients with very little prior computer experience have successfully learned computer-based information about health management and disease-related self-care and have reported satisfactory with computer based technique (Ogozalek, 1993). In addition, the interaction effect between human and computer also affects the use of computerized assessments. The following section will review the advantages and disadvantages of computerized assessment; human computer interaction (HCI) and the application of the theory of HCI in a computer-based cognitive assessment will be discussed in the subsequent sections.

Advantages of computerized assessment

Current literature suggests that computerized assessments enhance the automation of gathering clinical information and improvement in the assessment and documentation process (Wenzel, 2002). In a review of computerized assessment, Collie and his colleagues concluded that computerized cognitive tests and test batteries that are designed specifically for the detection of very mild cognitive dysfunction offer both practical and scientific advantages over conventional neuropsychological tests (Collie, Barby, & Maruff, 2001).

Firstly, it is possible for computerized assessments to have random alternative forms where the stimulus capacity can be controlled and the presentation of hierarchical and repetitive items challenges can be varied from simple to complex, contingent upon success (Rizzo & Buckwalter, 1997). This can minimize test-takers' frustration and loss of dignity when working on tasks once accomplished with ease. A specific example of this kind of computerized assessment is the computer adaptive test (CAT), which is a newly developed assessment strategy that provides more advanced test administration procedures (Hol, Vorst, & Mellenbergh, 2008). The CAT is embedded with an item bank and can select test questions that are most likely to obtain the information about the respondent (Hahn, Cella, Bode, Gershon, & Lai, 2006). Each question is thus adapted to the respondent's ability level. After each response, the test-taker's estimated ability is updated immediately and CAT selects subsequent items to reflect the test-taker's new estimated ability level (Van der Linden & Glas, 2000). In other words, based on previous responses, the CAT only selects questions that are at the test-takers estimated ability level and skips items that are too easy or too hard (Wainer, 2000). As a result, the CAT provides an efficient assessment system that is more precise than traditional assessment systems (Hahn, Cella, Bode, Gershon, & Lai, 2006). With the advancement of computer technology, there is increasing scope for the development of the CAT as a critical tool for cognitive assessments which has important clinical implications.

Secondly, computerized assessments can assist in recording and scoring examinee responses and, more importantly, suggesting the next item to be administered. The computer assessment can score the right or wrong answer, but still leaves the examiner the option of overriding the programme or probing the examinee for more details, storing the scored response for future use. Moreover, one of the commonest errors in test administration is that examiners stop testing at the wrong time, a mistake that can potentially result in erroneous test sores. This can be a purely mechanical decision that can be left to a computer programme, again with the option for the examiner to override the computer for specific clinical reasons. When the stopping criterion is reached, the computer prompts the examiner to begin the next subtest. Furthermore, a truly adaptive approach is that testing can focus very quickly on the items that are at the examinee's ability level, so more information is obtained per item administered. Testing can also focus on areas that appear to be of particular concern (Thorndike, 1999). Another major positive aspect of computerized assessment is that it has the potential to enhance efficiency in the clinical setting. It has been suggested that computerized assessments can save substantial amounts of time by decreasing the probability of errors being made either by respondents while filling out hand-written sheets or by assessors when hand-scoring items (Allard, Butler, Faust, & Shea, 1995). Computers offer efficiency advantages over tradition formats, such as reducing transcription errors and making possible new measurement options such as interactive branching, personalized probes, and the provision of explanatory material and online help (Richman, Kiesler, Weisband, & Drasgow, 1999). Another study demonstrated that multimedia version assessment yielded a more positive reaction and the test-taker perceived more content and predictive validity and felt that it provided more relevant information. In addition, they felt more enjoyment and satisfaction with the assessment process (Richman-Hirsch, Olson-Buchanan, & Drasgow, 2000). It has been postulated that if an assessment deals with sensitive or personal information, respondents may be more willing to reveal their true feelings to a computer than to a human being, which may lead to more informative results when computerized assessments are used (Hofer, 1985). Computerized assessments are also likely to be able to predict aspects of the criterion space (e.g. interpersonal relations) that are not easily predicted by cognitive ability (Richman-Hirsch, Olson-Buchanan, & Drasgow, 2000). Computer assessments have been shown to provide incremental validity when used in conjunction with tradition cognitive ability tests (Olson-Buchanan et al., 1998). Finally, computer-based program have been developed to accommodate persons with both physical and cognitive disabilities. These program provided elderly with opportunities for enhancing social interaction, diminished feelings of isolation and improved self-esteem.

Disadvantages of computerized assessment

A number of disadvantages of computerized assessment have been raised in the literature too. Firstly, there may be some individual discomfort with computers and consequent awkwardness when dealing with them (Hofer, 1985). It has also been suggested that even though many individuals have become quite familiar with computers through word processing and financial software packages, a mild degree of computer-phobia may still exist. This could potentially reduce the perceived usefulness of computer-based assessments for test-takers or clinicians (Schatz &

Browndyke, 2002). Secondly, the majority of current computerized assessments do not take into consideration human–computer interactions. These include potentially critical non-verbal cues such as speech pattern, vocal tone, and facial expression (Butcher, Perry, & Atlis, 2000). Also, the automated nature of current computerized assessments does not allow the examiner to interrupt or stop the assessment. This rigidity ultimately decreases the ability of test-taker to "test the limits" or the ability of examiners to be more flexible with their evaluations. By nature, computerized assessments present stimuli through either visual or auditory modalities. Thus, they do not allow for the collection of spontaneous verbal responses and eliminate the ability to test verbal functioning. Computer-based assessment may not address the dynamic needs of clients with "challenging" behaviour disorders or symptoms. These kinds of neurobehavioural presentations may require the clinician to alter the order, schedule, or pace of the assessment. Such alterations may not be possible with current computer-based techniques (Schatz & Browndyke, 2002). Thirdly, computerized assessments generally collect responses through mouse clicks or keyboard responses, which may be severely limiting for individuals with physical or motor control deficits such as stroke and traumatic brain injury patients.

Finally, it has been suggested that due to the excessive generality of results produced by computerized assessments, they have a high potential for misuse due to their increased availability (Butcher, 1987).

Consideration of Human–Computer Interactions in the development of the ICAS

The interaction between respondent and hardware has been an important consideration when developing the ICAS for stroke survivors. Human–Computer Interactions (HCI) is the complex study of how individuals use, design, and implement interactive computer systems and how computers affect individuals, organizations, and society. In the presence of stroke survivors, insight into the field of HCI shaped the design of the ICAS.

Developers of computer-based products are keenly aware of the crucial role that HCIs play in the marketability of products. They realize that consumers expect products to be highly effective and useful, have easy to learn interfaces, and create a pleasurable experience (Myers et al., 1996). This has resulted in a notable history of research into HCIs.

Since the early nineties, HCI development and research have gone through a number of stages, moving from a focus on a dialogue between humans and computers to a focus on work settings (Grudin, 1990). With the development and widespread use of network technologies, HCI then moved towards a new state, characterized by network and social design, where the HCI became “socialized” (Wellman, Haase, Witte, & Hampton, 2001). The focus of HCI also included a research focus on human–human interaction mediated by computer and network technologies. There are various approaches to studying the HCI and those frameworks, approaches, and theory guided the development of computerized assessment with efficient human–computer interaction. The cognitive model framework, distributed cognition approach, and interaction design, for instance, are the theoretical constructs of HCIs on which the development of the ICAS for stroke survivors was based.

Cognitive model framework in HCI

This refers to computer programmes implemented with the core resources of a cognitive architecture. Cognitive architectures are relatively complete proposals regarding the structure of human cognition and are generally believed to be capable of modelling cognitive activities. Cognitive architecture provides the resources for developing cognitive models that simulate human performance of cognitive skills. Cognitive modelling is mostly an iterative methodology similar to learning cycles which go through successive cycles of theory building, computational artefact construction, and empirical evaluation (Emond & West, 2003). It also takes into account interaction of all three elements of cognition, artefact, and task, known as the cognition–artefact–task triad (Gray & Altmann, 2001). These three elements are generally required to model human–computer interaction tasks. Cognition simulates the cognitive performance of a human performing a task; task simulation provides the task as well as the interface that will be used by the cognition; and artefact (a linkage mechanism) simulates human perception and action, so that the cognitive model can communicate with the task simulation (Ritter, Baxter, & Jones, 2000). This cognitive model has previously been successfully applied in many domains including perception and attention, learning and memory, problem solving, and decision making. Hence, the cognitive model framework is suggested to be highly applicable to developing a computerized cognitive assessment such as the ICAS in the present study.

Distributed cognition approach (Hollen, Hutchins, & Kirsh, 2000) in HCI

This extends the reach of what is considered to be cognition beyond the individual to encompass the interaction of individuals with each other, resources and materials, and the environment (Hollen, Hutchins, & Kirsh, 2000).The theory of distributed cognition seeks to understand the organization of cognitive systems. There are three tenets of the distributed cognition approach. The first is that cognitive processes are socially distributed across the members of a group. The second tenet is that cognition is not an incidental matter and that we are locked into causal relationships with our environment. The final tenet is that the study of cognition is not separable from the study of culture, because individuals exist in complex cultural environments. These three tenets shaped the development of the ICAS for stroke survivors by ensuring that the usability of the programme across different social groups, the interactivity between humans and computers, and the cultural relevance of the programme were integrated into the programme design.

Interaction design in HCI

This design approach informed the development of the ICAS for stroke survivors by offering a design approach that is empathetic, fun, motivational, aesthetic, helpful, and supportive. It extends the traditional central design approach towards empathy, fun, motivation, aesthetics, helpfulness, and support. The aim of interactive design is to create products that are usable, useful, and enjoyable (Preece, Rogers, & Shape, 2002). It integrates insights from ethnographic studies of practices and social environments in which the technology is used and studies into the interaction between user and technology.

Choosing an appropriate input mechanism would be an important consideration when developing the ICAS for stroke survivors. This is because input devices are the foundation of human–computer interaction. The basic task of computer input is to move information from the brain of the user to the computer (Jacob, 1996). The first commonly used physical device for computer input was the keyboard. Here, commands are input into the computer in text-string format. Following the development of the graphical user interface in the operating system, the mouse became an easier physical input device. A mouse click can select a command by using a pop-up menu, a fixed menu, multiple clicks, circling the desired command, or even writing the name of the command with the mouse (Jacob, 1996). Further development of technology has resulted in the touch screen as another input option instead of the mouse.

Keyboard, mouse, and touch screen are physical input devices operated by hand. Current technology now allows other body movements such as foot position, head position, and even the directional gaze to be used as computer inputs (Bolt, 1981; Jacob, 1991). Speech is yet another computer input mechanism (Schmandt, 1993). Nowadays, the technology used in virtual reality systems is one of advancement in computer input. Here a magnetic tracker is used to detect and orientate the position of a camera for scene rendering. Additionally, virtual reality gloves and other 3D input devices allow the user to interact with the displayed environment (Jacob, 1996). The input devices discussed above all enhance human–computer interaction to potentially allow individuals with disabilities to use computer systems in an easy and user-friendly way. Hence, the pros and cons of these devices should also be taken into consideration when developing the ICAS for stroke survivors.

Finally, computer feedback analysed as feedback is a vital component of human–computer interaction. Without feedback, there is only a one-way direction for the human to operate the computer rather than two-way interaction between human and computer. In human conversation, we use language, gestures, and body language to inform our conversational partners that we have heard and understood the communication in order to facilitate the continuum of the conversation. These communication expectations, also known as "psychological closure" (Miller, 1968; Simes, & Sirsky, 1988), exist when a human interacts with a computer (Perez-Quinones & Sibert, 1996). Feedback in HCI refers to communication from the system to the users as a direct result of a user's action (Shneiderman, 1987). It can also be used to communicate the state of the system independently of the user's actions. The system must let the user know its current state of processing so that the user does not feel frustrated or locked out of the dialogue. Feedback can be presented in the form of icons, sound, or computer graphics.

Feedback is used not only to let the user know the computer's current state of processing, but also as a prompt for the next step required to complete the tasks. An example of this type of feedback is "Microsoft Help" commonly seen in Microsoft Office. When computer programmes use this type of feedback, it allows for psychological closure just like conversations between humans. This understanding and insight into human–computer interaction from both a theoretical and a physical perspective has provided a sound platform for the development of the ICAS.

Psychometric properties for developing a valid and reliable computerized cognitive assessment tool

In order to develop a valid and reliable cognitive assessment tool, a sound understanding of test theory is required. The concept of test validity corresponds to whether a test or an assessment procedure provides the kind of information needed for a particular interpretation (Franzen, 2000). Put simply, it refers to whether a test or assessment measures what it intends to measure. Test validity is very important because test scores are meaningless unless they refer to a defined realm of observable phenomena. Reliability, on the other hand, refers to the level of consistency or stability of scores elicited by an instrument (Franzen, 2000). It is the extent to which an experiment, test, or measuring procedure yields the same results in repeated trials. Finally, test theory is concerned with methods for estimating the extent to which a specific assessment of psychological function influences measurement in a given situation and with methods for minimizing the errors (Willmes, 2003). There are two main test theories: Classical Test Theory (CTT) (Gulliksen 1987; Lord & Novick, 1968) and Item Response Theory (IRT). CTT will be only briefly discussed. As IRT offers more sophisticated mathematical models, it is viewed as more applicable to the development of the ICAS for stroke survivors. Therefore, IRT is where the emphasis of discussion lies.

CTT, also known as true score modelling, assumes that a subject's observed test score (X) is additively composed of the subject's true performance (T) and random error (E) (Willmes, 2003). This relationship can be summarized as $X = T \pm E$ (Portney & Watkins, 1993). Loosely speaking, the true score can be interpreted as the average of the observed score over an infinite number of repeated test runs with the same test. Practically, it is impossible to repeat the test infinitely. Therefore, the true score is derived directly from the observed test score with consideration of error of measurement.

Reliability can be considered as error of measurement, as it estimates the error in the formula of CTT. Reliability refers to the level of consistency or stability in the values of the scores that an instrument elicits (Franzen, 2000). There are various methods to estimate the reliability such as split half, test–retest, and alternative forms. Reliability measurement is an attempt to estimate the percentage of error variance (Anastasi, 1982). The reliability index is defined as the correlation coefficient which expresses the degree of relationship between the true and observed scores on a test. It is established by repeated measurement using the same assessment.

The concept of test validity basically corresponds to the question of whether a test or an assessment procedure supplies the kind of information needed for a particular interpretation (Franzen, 2000). Put simply, it refers to the question of whether the test or assessment is measuring what we intend to measure. It is very important because test scores are meaningless unless they refer to a defined realm of observable phenomena. Validation procedures are based on the types of evidence that can be offered in support of a test's validity. These types of evidence are generally defined as: content validity, criterion validity, construct validity, and ecological validity.

Content validity indicates the degree to which a test adequately samples from the domain of interest (Franzen, 2000). It requires a test which is free from the influence of factors that are irrelevant to the purpose of the measurement. The determination process of content validity is a subjective process. There are no statistical indices that can assess the content validity (Portney & Watkins, 1993). It can be determined by a review of the test items by a group of experienced experts.

Criterion validity indicates that the outcomes of one instrument can be used as a substitute measure for an established gold standard criterion test (Portney & Watkins, 1993). It is commonly expressed as the correlation between a test score and some external variables, which may be another test that is assumed to measure the same characteristic of interest or future behaviour that is assumed to demonstrate the characteristic of interest (Franzen, 2000). It can be tested as concurrent or predictive validity. Concurrent validity establishes whether the test is considered more efficient than the gold standard. Predictive validity establishes whether the outcome of the test can be used to predict a future criterion score. Construct validity, on the other hand, establishes the ability of an instrument to measure an abstract construct and the degree to which the test reflects the theoretical components of the construct (Portney & Watkins, 1993).

The validity concerns of assessments are expanded to ecological considerations (Franzen, 2000). The ecological validity refers to the test's ability to predict functional performance based on the test result. That is the extent to which a test predicts ability to function in important life tasks (Hart & Hayden, 1986). It has been stated that complete assessment of the ecological validity of an instrument involves investigations of both verisimilitude and veridicality (Franzen & Wilhelm, 1996). Investigation of verisimilitude includes examining the instrument with reference to theoretical consideration and with reference to situation descriptions. Veridicality is the extent to which test results reflect or can predict phenomena in the open environment. Verisimilitude may be more important in the design of an assessment instrument, as during the design procedure, one would need to carefully consider the intended use of the information. Once the instrument has been designed, aspects of veridicality become more important. Therefore, both verisimilitude and veridicality are important components to guide the development of the ICAS.

Item Response Theory (IRT), also known as latent trait theory, is a body of theory describing the application of mathematical models to data from questionnaires and tests as a basis for measuring abilities, attitudes, or other variables. It is a measurement model that relates performance on the behaviour sample to the latent variable (Embretson, 1999). Latent variables refer to variables that are not directly observed but are rather inferred (through a mathematical model) from other variables that are observed and directly measured. IRT provides greater transparency of resulting scores than CTT and hence has important applications in scale development (Coster, Ludlow, & Mancini, 1999).

The basic assumption of IRT is that individuals who possess more abilities are more likely to be able to successfully complete a task requiring those abilities. This feature is known as monotonicity. A relationship that is assumed to be monotonic and positive is the relationship between an individual's performance of the trait in question and his or her probability of succeeding in the task (Thorndike, 1999). More details of features and assumptions of the IRT are presented in Appendix II.

The Rasch model is a family member of IRT and is a probabilistic model that estimates an individual's ability based on the difficulty of test items (Rasch, 1960). The basic assumption of the Rasch model is that the more an individual is able to provide correct responses on a test, the more likely it is that he or she will be able to successfully complete a task requiring the specific ability measured by the test. The Rasch model allows for the measurement of latent variables that cannot be quantified by a measurement device (Caty, Arnould, Stoquart, Thonnard, & Lejeune, 2008). For cognitive constructs, this is done by comparing constructs based on the probability that an individual will pass the measured construct during a specified test. This model includes an item difficulty parameter and a person ability parameter. The mathematical formula is as follows:

$$\text{Log}\,[Pni / (1-Pni)] = Bn - Di,$$

where:

Pni is the probability of person n passing item i,

1-Pni is the probability of person n failing item i,

Bn is the ability of person n, and

Di is the difficulty of item i (Rasch, 1960; Wright & Linacre, 1987).

This overcomes the limitations of traditional psychological measurement (psychometric) methods (Hobart, 2002) by allowing comparisons between potentially varied and complex constructs using the same structure. This model features unidimensional measurement, a linear scale, sample-free calibration, and test-free measurement (Wright & Stone, 1999). This has implications for the development of a computerized cognitive assessment tool, as the Rasch model, when used with the powerful calculation abilities of the computer, provides a means for ensuring construct validity, investigating the difficulty level of test items, and from there reflecting cognitive ability on a linear continuous ratio scale.

Artificial intelligence (AI)

AI is the study of how to make computers do things that (if done by a human) would be perceived to require intelligence (Hancox, Mills, & Reid, 1990). AI is broadly defined as concerned with intelligent behaviour in artefacts. Intelligent behaviour, in turn, involves perception, reasoning, learning, communicating, and acting in complex environments (Nilsson, 1998). There are many special areas of AI, for example automated theorem proving, expert systems, machine learning, machine vision, natural language processing, robotics, and neural networks (Hancox, Mills, & Reid, 1990).

There are two approaches to AI, namely the symbol-processing approach and the sub-symbolic approach. The symbol-processing approach (also known as classical AI) uses logical operations when applied to declarative knowledge bases. This style of AI represents "knowledge" about a problem domain by declarative sentences. Logical reasoning methods are used to deduce consequences of this knowledge. When applied to "real" problems, this approach requires substantial knowledge of the domain and is then often called a knowledge-based approach. It often uses a top-down design method which begins at the knowledge level and processes downward through the symbol and implementation levels (Nilsson, 1998). An example of this approach is the expert system. Expert systems embody large amounts of human knowledge about a highly specific problem and use this knowledge to provide advice on what to do in particular circumstances. They usually have the ability to explain how solutions are reached. The expert system should remain unchanged if the knowledge base is modified and updated, or even if a new base plugged in (Hancox, Mills, & Reid, 1990).

Another approach, the sub-symbolic approach, usually proceeds in the bottom-up style, which starts at the lowest level and works upward. At the lowest level, this approach concentrates on duplicating the signal processing abilities and control system. An example of this approach is neural networks, which are inspired by biological models, and are very interesting and useful for studying the ability to learn. For the same reason, the neural network model was used in this project.

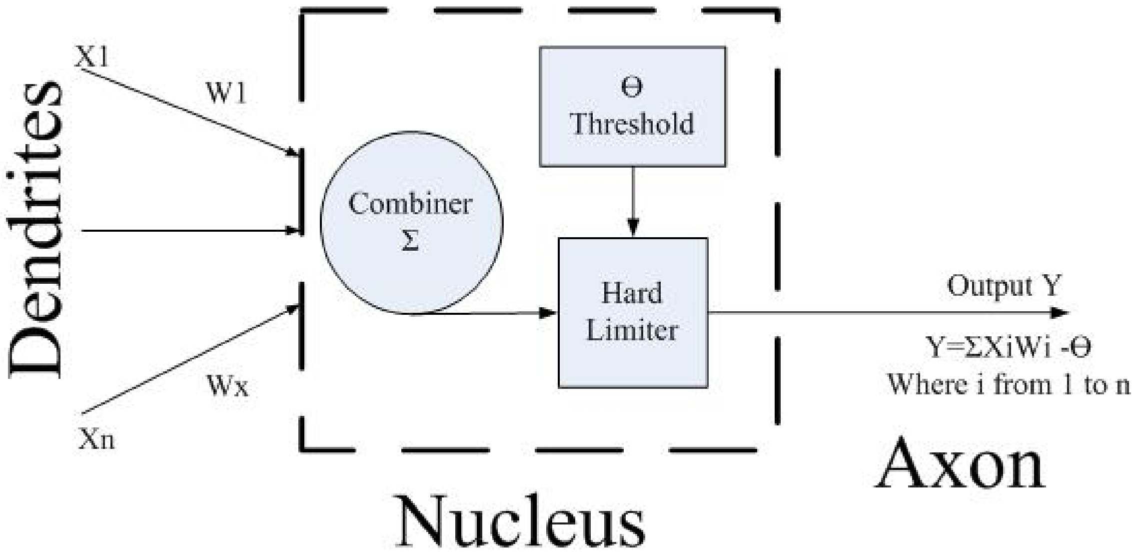
Artificial neural network (ANN)

The first ANN, invented in 1958 by psychologist Frank Rosenblatt and called Perceptron, was intended to model how the human brain processes visual data and learns to recognize objects. Other researchers have since used similar ANNs to study human cognition. Eventually, it was realized that in addition to providing insights into the functionality of the human brain, ANNs could be useful tools in their own right. Their pattern matching and learning capabilities allowed them to address many problems that were difficult or impossible to solve by standard computational and statistical methods. By the late 1980s, many real-world institutes were using ANNs for a variety of purposes (Kay, 2001).

ANN is attempting to model the cognitive architecture of the human mind and focus on the physical architecture of the brain. The power of the ANN is the parallel operation of simple units and the ability to adapt the configuration of the network (Finlay & Dix, 1996). An ANN is a means of processing complex data using multiple interconnected processors and computing paths. Inspired by the architecture of the human brain, ANNs are capable of learning and analysing large and complex sets of data that more linear algorithms cannot easily deal with (Kay, 2001).

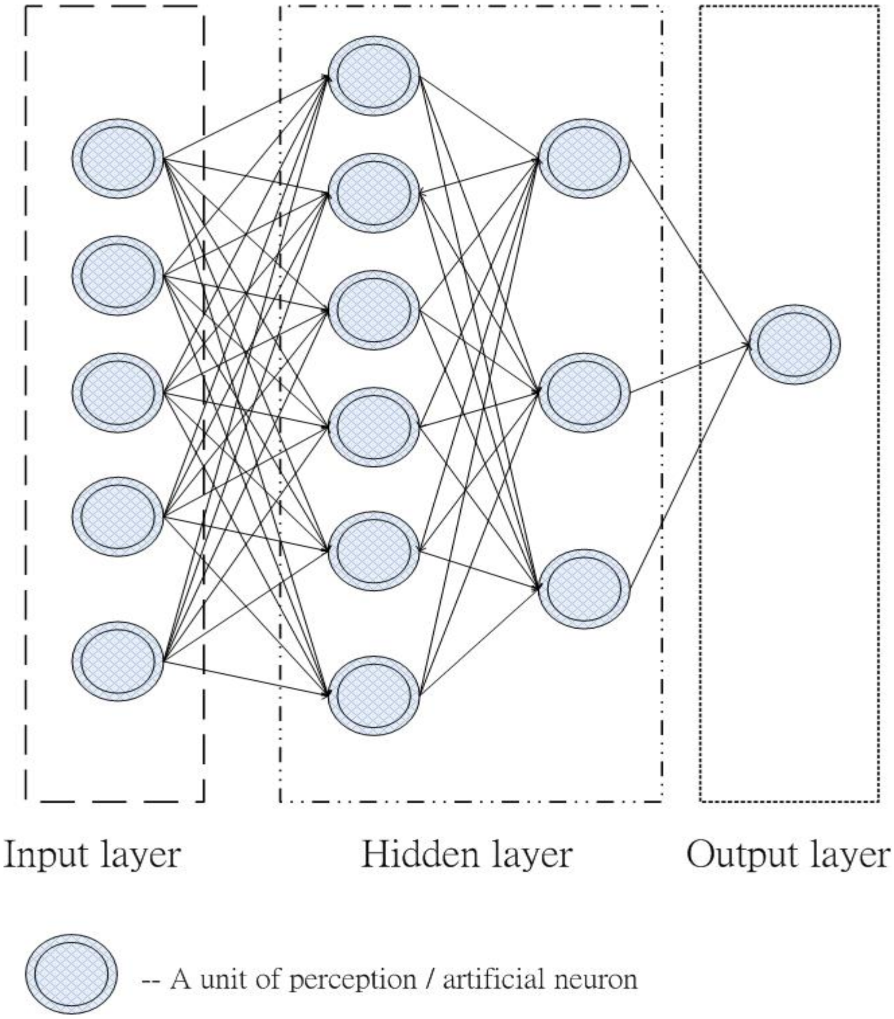
The basic unit of ANN is the perceptron, which simulates the neuron in biological networks (Figure 1). Each input (x) is multiplied by the weight (w) on its connection, which is set randomly to start with. The weighted inputs are then summed by the neuron and compared with the threshold value; if the threshold value is exceeded, the response is “on”, otherwise it is “off”. The perceptron learns by adjusting the weights to reinforce a correct decision or classification and discourage an incorrect one.

Figure 1. Structure of Perceptron unit



Perceptrons are the layered configuration of a neural network. The simplest form of perceptron is called a single-layer perception, or simply a perceptron, which consists of an input layer and an output layer only. Multi-layer perceptrons (MLP) have hidden layers in between the input layer and the output layer (Figure 2).

Figure 2. The ANN formed by multi-layer perceptrons



Types of ANNs are classified by their training style and signal transformation. Network training is the essential process of adjusting the arcs' weights so that they can represent input data in some numerical form within a network (Cheng & Titterington, 1994). ANN training methods can be divided into two approaches, supervised or unsupervised, depending on the availability of a target vector (or a desired output).

Supervised training methods require paired training data that include a target vector. The difference between the target vector and the actual output vector is an error signal, which occurs in the supervised training network. In contrast with the supervised training network, unsupervised training networks do not need a target vector. Input data are transformed to output clusters by unsupervised training networks.

Besides the training method, another classification is by their signal transformation: feed forward or feed backward. In feed backward networks, signals are sent back to the neurons in the previous layers. Thus, the feedback system networks are also called bi-directional networks. On the other hand, the feed forward systems are the networks whose signals are transferred in a forward directional only.

A back-propagation (BP) network is an MLP with a BP algorithm as a systematic training approach. Nodes (neurons) receive input values whether from the previous nodes or from the outside (in case of input nodes). Each of the received vectors is being weighted by the associated arc's weights and summed at each node. Then, the summation of the products is transformed by an activation function into the node's output value, which in turn becomes the input value to nodes in the next layer. The process continues until the output values in the output layer are calculated. Then these actual output values are compared with the target values and the differences (errors) between the target values and the actual networks outputs are computed. The training algorithm is essentially a recursive process of adjusting the arc's weights so that the network achieves minimization of the errors in terms of some error measures such as the sum of squares. In this respect, the BP network is an unconstrained nonlinear minimization problem (Wasserman, 1989; Lippmann, 1987).

Conventional statistical models may present certain limitations that can be overcome by neural networks. Predictive models provide a probability for a predefined classification. The classification can entail the prognosis for a specific condition. Models using linear and logistic regression models are limited. Complex nonlinear relationships among independent and dependent variables cannot be modelled using these methods. Multilayered neural networks are able to solve certain complex nonlinear problems and linear and logistic regression models are not. Therefore, neural networks have been increasingly applied in medicine (Lucila & Todd, 1999).

This lack of interpretability is one of the most criticized features in neural network models. Advocates of the method argue, however, that the existing trade-off between being able to model complex nonlinear functions and interpretation of weights favours neural networks for applications in which the primary goal is to obtain a reliable prediction rather than to get insight into the problems (Lucila & Todd, 1999), and it could be argued that the decisions of medical specialists often seem like a block-box situation to their colleagues (Cross, Harrison, & Sander, 2003).

Summary of literature review

This extensive review of the literature covers the components of cognitive function, the use of item response theory, and the application of ANN. It provides the backdrop to the development of the ICAS for stroke survivors. The structural framework of the ICAS and how important insights and methods can be gained from this critical analysis of the literature are integrated into this assessment system which will be discussed in the next chapter.

Chapter 3 Structural Framework in guiding the development of the intelligent cognitive assessment system (ICAS)

The intelligent cognitive assessment system (ICAS) was a newly developed computer adaptive test (CAT) for stroke survivors. Its development was based on the review of relevant literatures in cognitive function, assessment approaches, human computer interaction, psychometrics and artificial intelligence. This chapter will anatomize the ICAS and discuss its structural framework.

One of the aims of the study was to develop an intellectual cognitive assessment system (ICAS) for stroke survivors. The characteristics of this system would be:

(1) based on the CAT format, such that it estimates the cognitive abilities in a more efficient and precise way (Hahn et al., 2006); and

(2) using the artificial intelligence to predict the functional outcomes for stroke survivors.

To dates, these unique features of ICAS have not yet been found in other similar cognitive assessments: i.e. this study integrated the Computer Adaptive Testing (CAT); Item Response Theory (IRT) and Artificial Neural Network (ANN). As it is a computerized assessment, the Human Computer Interaction (HCI) theory is taken into consideration during the development of the ICAS. Based on the above-mentioned theoretical framework, the study generated an application in the aspect of cognitive function evaluation and through the use of computer in providing a rich media environment context for evaluation purpose. The structural model of ICAS is shown in Figure 3.

Development of the intelligent cognitive assessment system (ICAS)

After extensive review of lectures on the cognitive function after stroke and with taking reference to the existing cognitive assessments and some daily activities, the testing items of the ICAS were constructed based on the 3 level classification namely primary cognitive capacity, high level thinking ability and executive function. The ICAS items got six aspects on the primary cognitive capacity including attention span, orientation to time, semantic memory, working memory, prolonged memory and visual inattention; 6 aspects on the high level thinking including visual recognition, visual interference, abstract thinking, calculation, sequence and similarity categorization; and one aspect on executive function. In each of aspect, there were different items with different level of difficulties on those aspects and totally got 65 items. Some of these items were based on the traditional Chinese culture such as semantic memory of traditional Chinese festivals and some of them were picked up from daily activities such as pressing the door gate lock and using of mobile phone, etc. The items and their corresponding aspects were showed in Table 2.

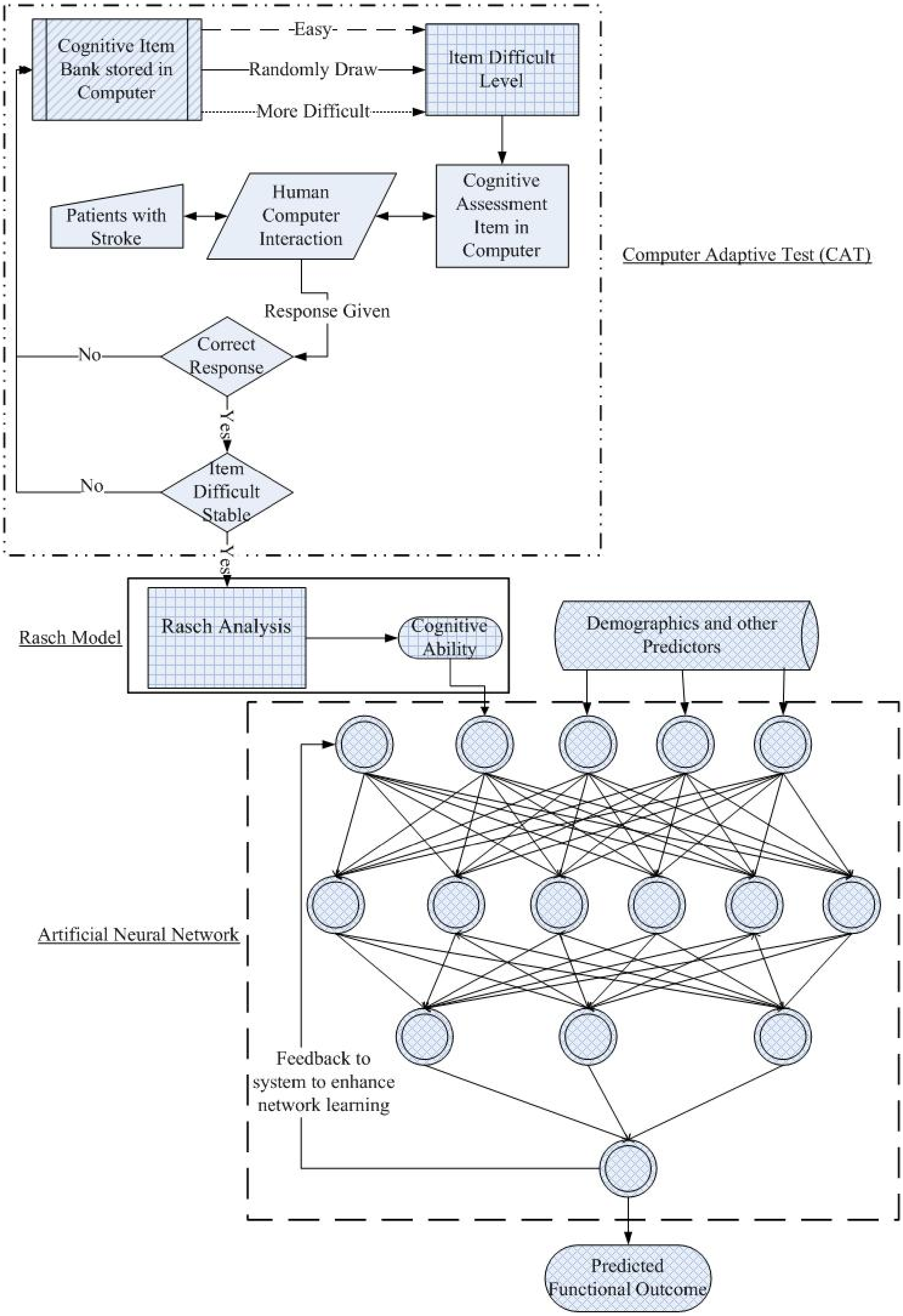
Table 2: Aspects of cognitive function being assessed in ICAS item bank

Items Number	Cognitive Aspects
1-3	Working Memory
4-10	Orientation to time
11-20	Semantic memory
21-27	Calculation
28-30	Visual Recognition
31-40	Abstract Thinking
41-46	Visual Interference
47-51	Attention Span
52-53, 55, 59-62	Executive Function
54	Visual Inattention
56-57	Similarity Categorization
58	Sequence
63-65	Memory

The platform of ICAS was based on Macromedia Flash MX 2004, which provided a rich media and animated platform for the assessment content to be presented to test-takers. The test item content has been selected from the tasks that we often tackle in the daily activities, information we use as general knowledge and specific neuropsychological assessment items. A computerized programme was successfully developed by using of Macromedia Flash MX 2004 and embedded with the multimedia effect, animation and action effect, while the recording of test-takers' responses were controlled by action script 2.0. This programme then became the prototype of the ICAS.

Trial run of ICAS on our target population was done to investigate the difficulty levels of the test items, which was based on the Rasch Model. Then the levels of item difficulty were integrated into the computer programme and by using of action script 2.0 to control the sequence of presented test items to test-takers to achieve the CAT format. After completing the assessment, the recorded responses from test-takers were transformed into logits by using of Rasch Model (for detail, please refer to the methodology session) and the result reflected test-takers' cognitive abilities. Finally, the cognitive abilities, together with other predictors from the literature review, were used to build up a back-propagation artificial neural network (BP–ANN) to predict the functional outcome. After the BP–ANN became "stabilized", it was then integrated into the programme to get a finalized ICAS which could assess the cognitive functions and predict the functional outcomes at one time. Thus, the underlining theories of the ICAS's development were through the integration of the CAT, HCI, IRT and ANN.

Figure 3 The structural framework of ICAS proposed by Yip and Man (2009)



Human computer interaction (HCI) and Computer Adaptive Testing (CAT)

Human computer interaction (HCI) in the ICAS

Having a bank of cognitive assessment tasks (test items) in the computer, the administration format and content were standardized, through the interaction between patients and the computer. Their neurobehavioral responses were recorded accurately by the computer too. Moreover, through the computer platform, test items could be presented in a more interactive manner, and they were rich in media of delivery and simulated the real situation and environment. Therefore, test items in the computer enhanced the ecological validity, so, we anticipated test-takers were more capable of coping real task if they passed or completed the tasks given by the computer. Furthermore, interaction between test-takers and the computer may minimize the anxiety and enhanced their performance when compared with direct assessment or performing the tasks under therapist's supervision.

Furthermore, with considering our client groups were elderly suffered from stroke, the interaction was based on the computer input of that was operated through a touch screen, a digital pen and a much simplified keyboard. With special needs, enlarged keyboard, head mouse pointer or chin control mouse pointer could be adapted so as to ensure test-takers could access the computer freely. The tasks presented in the screen could be repeated if patients were unable to capture the content of the tasks, until responses were successfully input into the computer. As clients' visual deficits were reported to be a problem for senior participant in computer based group (Ogozalek, 1993). The development of the ICAS adopted the design of age-sensitive computer program to accommodate the sensory deficits that occur with aging. The sound effect and bright colors were used in presenting content so as to provide high stimulation, positive feedback to test-takers, especially for elderly clients.

In actual operation, the questions or tasks were given in a multiple-choice format of five answers: one correct answer, three distracters and one "don't know" answer, as suggested by Courtenay & Weidemann in 1985. This arrangement aims to reduce the guessing effect. The application of HCI also guided the development of the computer software of ICAS into an easy-to-use and user-friendly system.

Computer Adaptive Testing (CAT) in the ICAS

The ICAS question bank contained a highly selected 65 items that assessed 13 aspects of cognitive functions: working memory, orientation to time, semantic memory, calculation, visual recognition, abstract thinking, visual interference, attention, executive function, visual inattention, similarity categorization, sequence and memory. The number of items in each assessment area was shown in Table 2. Screen shot of some testing items were also presented in Appendix III.

The ICAS adopted the integrated functional approach for assessment (i.e.

interfaces the bottom-up approach and top down approach with the assumption that a relationship exists between the specific components and performance at a functional level). In addition of basic cognitive functions, test takers needed to perform functional tasks displayed on the computer screen. For examples, boiling water, the use of the telephone and octopus card, etc. Some basic cognitive components were assumed to be able to support test takers' in completing more demanding functional tasks. The items within the same cognitive function got different level of difficulty, and the computer could estimate test-takers' abilities based on their first three attempted items. If a test-taker input the right response, then the computer would present a more difficult item, otherwise, an easier item would be given instead. The ICAS operation would stop once the test-takers' abilities met the stopping criteria of the test. The final linear scale score reflecting test-takers abilities was then presented as the result screen of the test. In other words, test-takers could know their results immediately. The stopping criteria set for the ICAS were namely:

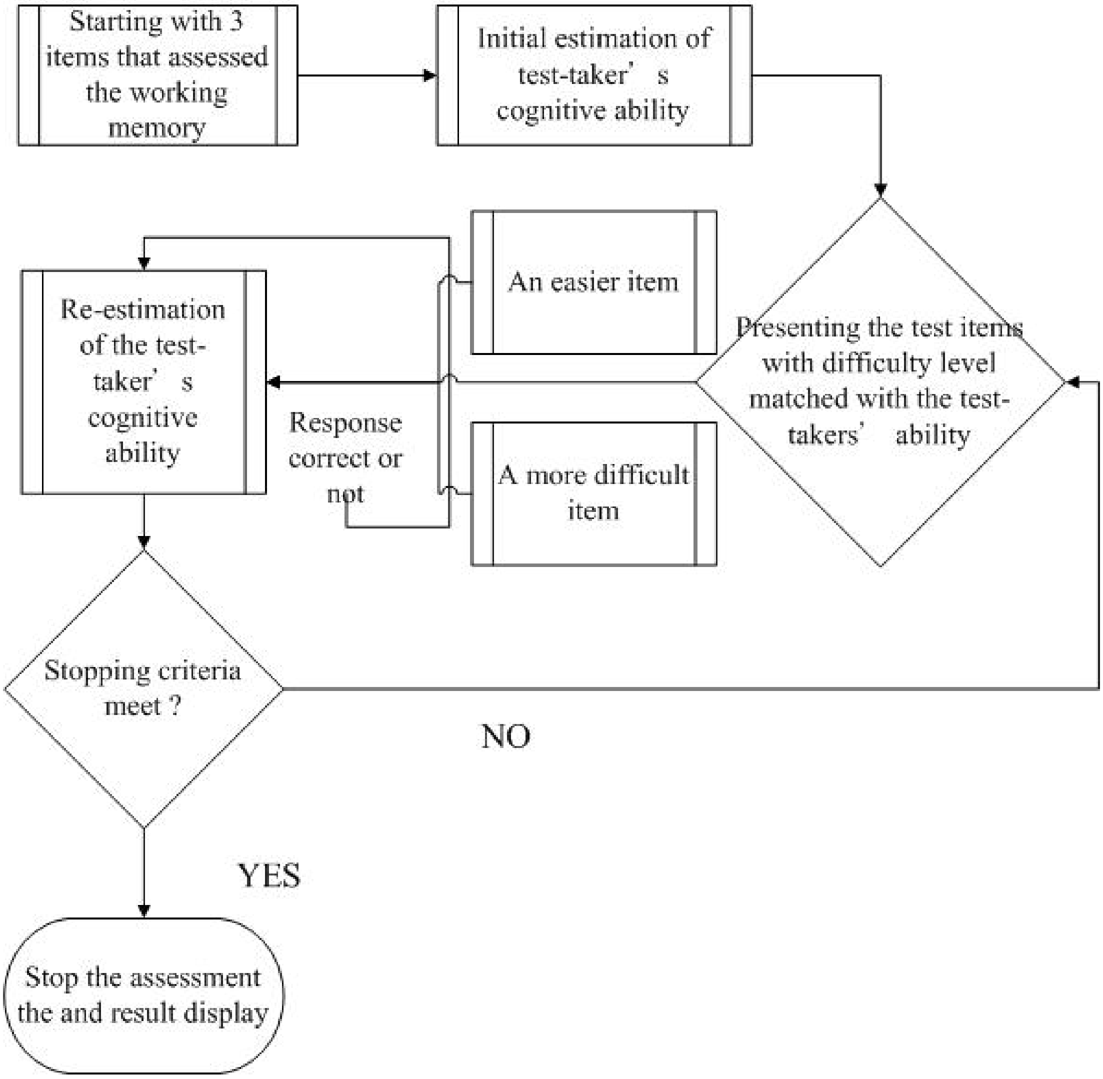
i) All test items were used up;

ii) The standard error was smaller than 0.4 logits (Halkitis, 1993);

iii) The converged cognitive ability estimated (the estimated ability difference between two items) was smaller than 0.02 logits. This also implied the stability in the estimation of the cognitive ability (Wright & Douglas, 1996).

The adaptive procedures of the ICAS were further shown in Figure 4. The CAT feature was considered as a unique feature of the ICAS that could assess the cognitive function of test-takers in an efficiency and accuracy way.

Figure 4 The adaptive procedures of the ICAS purposed by Yip and Man (2009)



Item Response Theory (IRT) as applied in the ICAS

The test item difficulty levels were developed according to the Rasch Model -- a family member of IRT. The scoring system for all items in the ICAS was in dichotomous form, i.e. right answer or wrong answer. After completing the requirement in each item, a test-taker would be scored one for that item, and otherwise zero. Each score was just the observation count of a test taker' response to that task. Each items in the item bank got its own difficulty, and they were presented according to test-takers' performance of the presented item. If a test-taker passed the item of a certain difficulty level, he/she then would be presented with another item of higher difficulty level, otherwise, he/she would have an item of lower difficulty level. As the difficulty levels were different across different items, the cognitive function might not be assessed in different trials of the test. This mechanism reduced test-takers' learning effect of the assessment. Moreover, item selection was based on test-takers' performance when they were extracted from the item bank, this action also reduced the learning effect. Furthermore, as the item presented to test-takers correlated to their abilities, and each item or question got meaningful to them, this arrangement may reduce the time for test administration and increased the test validity (Wainer, 2000). After test completion, test-takers's abilities would be computed, based on Rash model. The raw scores were then converted to logits, which could pull up all cognitive functions into a common linear ratio scale and reflected the test takers' overall cognitive abilities. In this way, cognitive abilities among stroke survivors could be compared directly. And the overall ICAS score could be used to monitor the progress of stroke survivors over time and cognitive rehabilitation therapy' effectiveness

Artificial Intelligence (AI) in the ICAS

In order to develop the AI component within the ICAS, an artificial neural network (ANN) was used. ANN was developed by Matlab Neural Network Toolbox 5.1 (MathWorks, 2007) and it has an advantage in dealing with non-linear relationship in classification and prediction. The data of patients' cognitive abilities, types of stroke, side affected, upper limb function and length of onset from stroke formed a matrix and these information was all entered into the MATLAB to build up an artificial BP network and train up the network. After the network became stable, the BP network was exported into a "com" or "dll" format, which could be controlled by action script 2.0 and finally intergraded into the ICAS to predict the function outcomes.

In summary, the ICAS integrated the several theoretical frameworks to guide development of a new assessment system that could serve the dual purposes of cognitive assessment and prediction of functional outcomes. It may serve as a documentation tool for cognitive functions of stroke survivors. The overall linear ratio scale of different cognitive components could be used to monitor the progress of treatment programmes and for research purpose. Therapists may save time in conducting assessment and generating assessment reports, enriching information for setting or adjusting treatment goals. They may have better preparation to decide on patients' discharge, as the system could quickly provide both the functional outcome forecast and the cognitive function information for consideration.

Chapter 4 Methodology

This chapter provides the basic scientific information in the procedures of developing of the intelligent cognitive assessment system (ICAS). The study objectives, the procedures of the study and the method of statistical analysis, together with sample size planning, will be described.

The study was divided into three phases to achieve its aim and objectives. Phase I study was the pilot study of the ICAS and the content validity of the item bank of ICAS was assessed. Phase II study investigated the item difficulty measures, item structure stability and uni-dimensionality of testing items in the item bank of ICAS. Phase III study examined the cutoff point of the ICAS in screening cognitive impairment for stroke survivors and building up the prediction model through artificial neural network (ANN).

The sub-objectives of the study were further outlined as follows:

1. To develop and validate content validity of the items bank of ICAS (Phase I study).
2. To investigate the item difficulty measures, item stability and unidimensionality of the items bank of ICAS for stroke survivors (Phase II).
3. To investigate the psychometric properties of the ICAS and develop an optimal cutoff score for stroke survivors with cognitive impairment (Phase III).
4. To build up and validate an artificial neural network (ANN) to predict the

functional outcome of stroke survivors from the demographic and cognitive functions (Phase III).

For all phases of the study, the target population was stroke survivors. To increase the coverage of stroke in different recovery stages, subjects were recruited from sub-acute ward, rehabilitation ward and day hospital of a local hospital. We operationally defined the sub-acute phase as a period within 2 weeks after the onset of stroke, and patients were staying in the sub-acute ward of a hospital; the rehabilitation phase (2-8 weeks post onset) when patients were staying in a rehabilitation ward; and the community phase (8 weeks or more post onset) when patients were attending a day hospital or an out-patient service. Subjects were selected according to the inclusion and exclusion criteria which were described in Table 3.

Table 3 Inclusion and Exclusion criteria for sampling of study population

Inclusion Criteria	Exclusion Criteria
1. Age 60 or above	1. Suffered from transient ischemic attack (TIA)
2. Suffered from Stroke and confirmed by CT brain	2. Premorbid diagnosis of vascular dementia or Alzheimer's disease
3. Hemorrhagic stroke or infraction stroke	3. Uncooperative and unable to follow instructions
4. Medically stable	4. Visual or hearing impaired
5. Able to follow verbal instructions	

Potential subjects fulfilled the above criteria were invited to participate in the study. They were screened out by occupational therapists (OT) according to selection criteria of the study after receiving the referral in the sub-acute ward, rehabilitation ward or the day hospital. If potential subjects agreed to participate in the study, subjects were asked to sign a written consent form. After that, they went through the data collection procedures according to the corresponding phases of the study. The general data collection procedures were shown in Figure 4. The variables to be collected in different phases of study were described separately. All the assessments were conducted by individual case therapists that were the routine assessments in the stroke rehabilitation programme in that hospital. All the phases of the study were approved by the Kowloon West Cluster Clinical Research Ethics Committee of the Hong Kong Hospital Authority.

Instrumentations used in the study were described as follows:

Intelligence Cognitive Assessment System (ICAS)

The ICAS was a newly developed CAT in this study. It has an item bank containing 65 testing items and assesses different aspects of cognitive function including orientation to time, visual recognition, visual interference, visual inattention, attention, working memory, semantic memory, calculation, abstract thinking, executive function, similarity and categorization. The content validity of the ICAS has been reported as part of the present study (Yip & Man, 2009). The scoring system of the ICAS is based on the Rasch model. The score refers to the cognitive abilities identified above and is a linear scale that represents 13 different aspects of cognitive function (Yip & Man, 2009). It was used throughout different phases of the study.

Mini Mental Status Examination (MMSE)

The Mini-Mental Status Examination (MMSE) was developed by Folstein and colleagues (Folstein, Folstein, & McHugh, 1975). It is a well-known screening test of cognitive status and covers a wide range of cognitive functions including memory, orientation, visual-spatial copying and language. The MMSE has recently been used as a screening test for Alzheimer's disease, although it is less sensitive to subcortical dementia or dementia secondary to ischemic vascular disease. The Chinese version of MMSE (MMSE – CV) is validated in Hong Kong (Chiu, Lee, Chung, & Kwong, 1994). The maximum MMSE–CV score is 30 and the suggested cut-off point for Chinese elderly is 21 (Chiu, Lee, Chung, & Kwong, 1994). It is a well-known cognitive assessment, and is proposed as one of the standards for establishing the concurrent validity of the ICAS in this study. It was used as a golden standard of cognitive assessment in all the three phases of the study.

Neurobehavioral Cognitive Status Examination (NCSE)

The Neurobehavioral Cognitive Status Examination (NCSE), also known as Cognistat, was developed by the Northern California Neurobehavioral Group (1995) as a screening tool for detecting and characterizing cognitive function. It adopted a screen and metric approach and given a profile score in ten cognitive aspects. This test assesses global cognitive function in five areas: language, construction, memory, calculation and reasoning. Attention and level of consciousness are assessed independently. Language has four subsections: spontaneous speech, comprehension, repetition, constructional ability and naming.

The Chinese version of the NCSE (NCSE – CV) (Chan, Lee, Wong, Fong, & Lee, 1999) was used in the study as another standard to establish the concurrent validity of the ICAS in Phase I and Phase II study. It was used to serve as another golden standard of cognitive assessment, as it was a more detail cognitive assessment with sub-components score only.

Modified Barthel Index (MBI)

The MBI is widely used to measure patient's basic self care performance. The items measured include personal hygiene, feeding, dressing, bathing, toileting, stair climbing, bowel control, bladder control, ambulation, use of wheelchair and bed/chair transfer. It was originally developed by Mahoney and Barthel in 1965 and modified by Shah, Vanclay and Copper in 1989 to increase the sensitivity to detect the changes. The MBI served as a tool for measuring the functional outcomes for patient with CVA in Phase I and Phase III study.

Upper Limb Function Test (ULFT)

The functional test for the hemiplegic upper extremity (FTHUE) (Wilson, Baker, & Craddock, 1984) was developed in Rancho Los Amigo Hospital in California and was a good attempt to evaluate the recovery in the hemiplegic upper extremity from non-use to full hand function. It evaluates upper extremity function as a whole rather than looking at separate parts of the extremity or simply the hand function. The FTHUE has been translated into a Chinese version with the content changed to fit the culture of Hong Kong (Fong et al., 2004). The functional test for the hemiplegic upper extremity Hong Kong version (FTHUE-HK) has a high level of concurrent validity with the FTHUE and the hand sub-score of the Fugl-Meyer Assessment. It is used as a measurement of hand function for patients with stroke and served as a physical component predictor in Phase I and Phase III study.

The details of investigation procedure and statistical analysis of each of the Phase were described in the following sessions.

Phase I Study -- Development of the ICAS and pilot study

Development of ICAS

With review of the literatures in cognition ability and cognitive function level in the daily tasks, the testing items of the cognitive assessment were extracted. Then, the ICAS was developed by using of Macromedia Flash MX 2004 and details of structure of ICAS can be referred back in Chapter 3.

Panel review for the content validity

After completion of the ICAS computer software, it was sent to panel review members together with a questionnaire. The members consisted of an associate professor, a senior medial officer, a manager of an occupational therapy department and several occupational therapists I and II. All of the members had at least five years of experience in the field of cognitive rehabilitation or neuro-rehabilitation. They were asked about the content validity of the ICAS items in measuring the specific cognitive areas and to elicit their comments on the ICAS. A reminder was sent two days before the deadline to ensure high return rate. After collecting the questionnaire and receiving feedback from the panel members, minor adjustments/modifications were made according to panel members' comments.

Pilot testing of the ICAS

Subsequently a pilot testing was carried out on 14 stroke patients that fulfilled the inclusion and exclusion criteria for testing the fluency of ICAS administration of ICAS. Pilot data was also collected for further study. Their cognitive functions were assessed by the ICAS, MMSE–CV and NCSE–CV respectively. Demographic data and the Modified Barthel Index (MBI) score were also collected.

Data Processing and Statistical analysis

Questionnaire results from the panel members were analysed by descriptive statistics. ICC model 2 (ICC(2,k) (McGraw & Wong, 1996) was used to investigate the content validity of the ICAS. Spearman's rho correlation coefficient was used to investigate the concurrence of the ICAS with the MMSE–CV and NCSE–CV. Regression model was used to investigate the cognitive ability of the pilot subjects and to predict the functional outcome in terms of the MBI. Rasch analysis software WINSTEPS (Linacre, 2006) was used to perform the Rasch analysis and to calculate the cognitive abilities and item reliability index. This index indicates the replicability of item replacement along the pathway, that is, whether the same items could be given to another sample of the same size that behaved in the same way and this index could be interpreted in the way like the Cronbach's alpha (Bond and Fox, 2007). Therefore, it was used to establish the construct validity of the ICAS.

Phase II Study -- Development of the item difficult trait level

After Phase I study ICAS was then administrated on the patients who also fulfilled the selection criteria.

The numbers of potential subjects recruited were based on Linacre's (1994) suggestion that a sample size of 30 subjects was adequate to demonstrate item calibration stability within ± 1 logit with a 95% confidence interval. The data obtained in Phase II study were used to develop the item difficult measures of the ICAS test items, based on the Rasch model. Traditional statistical analysis was also performed to test the validity and reliability as a benchmark with NCSE–CV and MMSE–CV.

Data Collection Procedure

The data collection procedure was the same as Phase I study. Another group of 30 subjects were recruited into the Phase II study. They were assessed in cognitive functions by means of the ICAS, MMSE–CV and NCSE–CV. In addition, their demographic data including age, gender, type of stroke and lesion side were collected. All 65 items in the item bank of the ICAS were used when conducting the ICAS assessment, and the raw scores of the items were collected for further analysis.

Data Processing and Statistical analysis

Items responses were analyzed by the Rasch model and software WINSTEPS (Linacre, 2006) was used for Rasch analysis. All responses to the ICAS items were analyzed using the WINSTEPS (Linacre, 2006). The cognitive ability scores of subjects (person measures) and item difficulty measures of the ICAS were thus calculated.

To ensure that the item difficulty measures of the ICAS item bank fit the Rasch model, we performed fit analysis. The INFIT mean-square (INFIT) and OUTFIT mean-square (OUTFIT) statistics of the 65 items in the ICAS item bank were examined. These two sets of statistics served as indicators of the data fit to the Rasch model. They were the mean of the squares of the residuals of those items. The residual was the difference between the Rasch model's theoretical expectation of item performance and the performance actually encountered for that item in the data matrix (Bond & Fox, 2007). The reasonable values for both the FIT mean squares ranged from 0.5 to 1.5, which was an allowable measurement range for clinical observation (Wright & Linacre, 1994). Principal components analysis (PCA) of residuals also helped to confirm if the items were in the same dimension. If the variance explained by the measure was more than 60% and the variance explained by the first contrast is less than 5%, then the unidimensionality of the items was confirmed (Linacre, 2006).

For a good test, the testing items should be constant or invariance no matter in high ability group or in low ability group. Therefore, to further establish the stability of the item measures, we re-grouped the 30 subjects into two groups according to a cut-off point of 21 for the MMSE–CV score (Chiu, et al., 1994). Further analysis of each group was carried out with WINSTEPS, and the item measures were obtained again. Then we compared these measures by t-test and Spearmen's rho (rank correlation coefficient) to investigate the differences and correlations of the measures between the two cognitive-different groups. In addition, we plotted the item difficulty measures of these two groups to investigate the invariance of the ICAS test items.

The cognitive ability scores (person measures) of subjects were correlated with their MMSE–CV and NCSE–CV scores to determine the concurrent validity of the ICAS. In addition to determining the correlation coefficient of all of the NCSE–CV components, the NCSE–CV profile scores were analyzed by WINSTEPS and an overall score that represented the NCSE–CV components was generated. Then, correlation analysis of the cognitive ability score and overall NCSE–CV score was used to further establish the concurrent validity of the ICAS.

Phase III study -- Psychometric properties of ICAS and the ANN model for outcome prediction

After the Phase II study, the item difficulty measures of the ICAS test items were validated and the ICAS was then finalized with adaptive test features. In this study Phase, the psychometric properties of the ICAS, including the concurrent validity to golden standard cognitive assessment – The MMSE–CV, test-retest reliability of the ICAS and the cutoff score for stroke survivors with cognitive impairment were investigated. Moreover, the predicting model based on ANN was developed to predict the functional outcome in terms of MBI for the stroke survivors.

Variables including age, gender, types of stroke (infraction or heamorrhage), side affected, residual upper limb functions, cognitive ability and residual self care ability were used to predict the functional outcome. They were collected together with the result of the ICAS in order to build up the BP–ANN model for prediction of the functional outcome. The correlation between the observed and the predicted outcome were used to serve the indicator of the BP–ANN model. Finally, the trained and stable BP–ANN model was integrated with the ICAS.

Sample size planning

In this phase, based on the result from the pilot testing in phase I and by using a sample size estimation software PASS 2000, a sample size of 62 could achieve 80%

power to detect an R-Squared of 0.1 attributed to 7 independent variable(s) using an F-test at 1% level of significance (alpha = 0.01). Taking around 5% of dropout rate, totally 66 subjects were required.

Data Collection Procedure

The data collection procedure was similar with Phase I and II study, except that subjects needed to be re-assessed by the ICAS within seven days of the first administration of the ICAS to determine the ICAS's test-retest reliability. The cognitive functions of 66 subjects were thus assessed by both the ICAS and MMSE–CV in a random sequence. Demographic data on age, gender, type of stroke and side affected and the Modified Barthel Index (MBI) score at the admission and discharge from the hospital or the OT services were again collected.

Development of Back-propagation ANN (BP–ANN) predicting model

The recruited subjects were divided into two groups equally and randomly, the data from the first group was used to train up the BP–ANN model and the second group was used to test the BP–ANN model. The independence between the two groups was tested by t-test and chi-square. Functional outcome was measured by Modified Barthel Index (MBI) and predicting variables mentioned before.

A three layer BP–ANN with 7 input neurons in the input layer, various neurons in hidden layer and 1 neuron in the output layer were successfully built by the software Matlab Neural Network Toolbox 5.1. There were different neurons in the hidden layer, which were tested to find out the optimal number of neurons in the hidden layer to give the most accurate forecast ability. The correlation of the predicted MBI value to the observed MBI value served as an indicator for the best prediction ability.

All the data from the first group were entered into the BP–ANN network to train the network forecast the functional outcome in term of Modified Barthel Index (MBI).

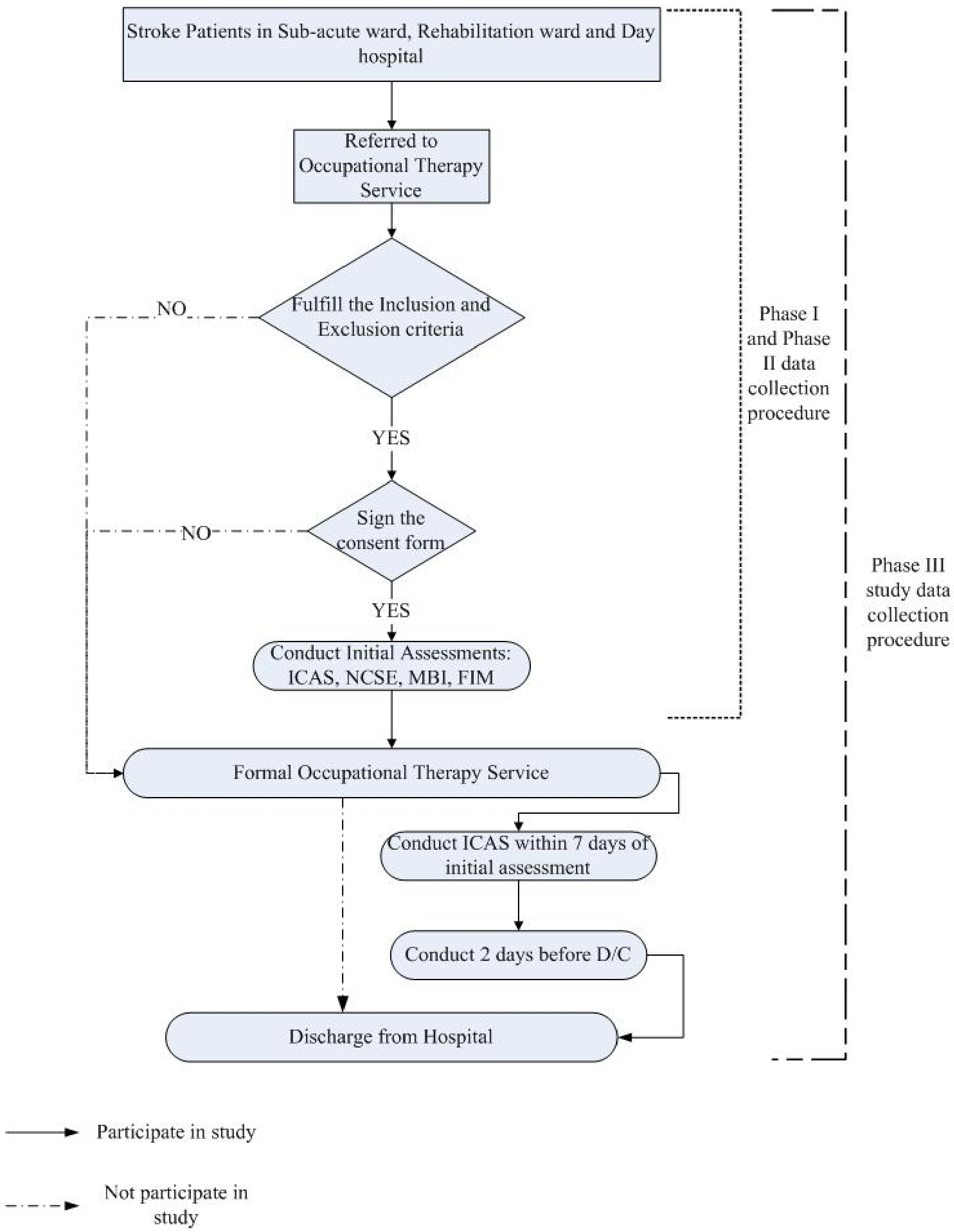
The second group was to verify the stability of the BP–ANN. The predicted values of MBI by BP–ANN were then compared with the clinical measured value of MBI. Correlation coefficient was used to determine the relationship between BP–ANN prediction and actual observation. The network with the highest correlation coefficient between the observed value and predicted value was incorporated into the ICAS and served as the prediction function.

Data Processing and Statistical analysis

Subjects were operationally classified into cognitive impaired group and non-cognitive impaired group according to their performance in the MMSE–CV. A cutoff point of 21 in MMSE–CV score (Chiu, et al., 1994) was adopted in this classification. Descriptive analysis of the demographic characteristics of subjects was followed by a comparison of the two groups using the independent t-test and chi square test. Then, the sensitivity and specificity of the ICAS were estimated using the receiver operating characteristics (ROC) curve. The optimal cutoff score of the ICAS and the system's diagnostic accuracy in correctly identifying cases with cognitive impairment were determined by the area under the ROC curve (Area under curve or AUC). The test-retest reliability of the ICAS was further tested using Cronbach's alpha, and concurrent validity was established by the correlation between the ICAS and MMSE–CV.

In short, three phases of study were conducted to develop and validate the ICAS and incorporate the ANN into the ICAS for prediction purpose.

Figure 5 The flowchart of data collection procedure



Chapter 5 Results

As mentioned in chapter 4, the objectives of the study were achieved in three phases and the results of each phase were presented in that order.

Phase I Study – Development of the ICAS and pilot study

Phase I study was the pilot study which explored the content validity of the item bank of the ICAS. The preliminary concurrent validity of the ICAS with the MMSE–CV and NCSE–CV and prediction the functional outcome from cognitive ability and residual functional status were achieved. The results of the Phase I was presented in two parts. The first part was the content validity demonstrated by an expert panel and the second part was the concurrent validity and prediction.

Content Validity

The response rate was 100% and all questionnaires from the panel members were collected. The mean (S.D.) experience of the panel members in the cognitive rehabilitation was 10.25 (2.32) and ranged from 8 years to 15 years. The result showing all the panel members agreed that 58 items out of total 65 items of ICAS (89%) were good to excellent in the content relevance to the assessed content, except 7 items (11%) got agreement below 87.5%. Item 18 got the lowest agreement on the content relevance, 25% of panel member rated it poor in assessing the semantic memory. They further commented that the picture was too vague to visualize the objects. Other items got “fair” rating, they included item 15, item 26, item 27, item 61, item 64 and item 65. The percentage of agreement among the panel members were shown in Table 4. The Intraclass Correlation Coefficient (ICC(2,k)) among the panel members in their agreement of content relevance was 0.972 with $p<0.01$.

Table 4 The percentage of agreement about items assess corresponding area of cognitive functions

Item	% Poor	% Fair	% Good	% Very Good	% Excellent	% Total
Working Memory						
01	00.0	00.0	37.5	62.5	00.0	100
02	00.0	12.5	37.5	50.0	00.0	100
03	00.0	12.5	37.5	37.5	12.5	100
Orientation						
04	00.0	00.0	12.5	75.0	12.5	100
05	00.0	00.0	12.5	62.5	25.0	100
06	00.0	0.0	25.0	62.5	12.5	100
07	00.0	00.0	25.0	62.5	12.5	100
08	00.0	00.0	00.0	87.5	12.5	100
09	00.0	00.0	12.5	62.5	25.0	100
10	00.0	00.0	25.0	50.0	25.0	100
Semantic Memory						
11	00.0	00.0	12.5	75.0	12.5	100
12	00.0	00.0	25.0	62.5	12.5	100
13	00.0	00.0	25.0	50.0	25.0	100
14	00.0	00.0	25.0	62.5	12.5	100
15*	00.0	25.0	12.5	62.5	00.0	100
16	00.0	12.5	25.0	62.5	00.0	100
17	00.0	00.0	25.0	75.0	00.0	100
18*	25.0	37.5	12.5	25.0	00.0	100
19	00.0	12.5	25.0	62.5	00.0	100
20	00.0	00.0	37.5	62.5	00.0	100
Calculation						
21	00.0	00.0	25.0	62.5	12.5	100
22	00.0	00.0	12.5	75.0	12.5	100
23	00.0	00.0	12.5	75.0	12.5	100
24	00.0	00.0	25.0	62.5	12.5	100
25	00.0	00.0	12.5	75.0	12.5	100
26*	00.0	25.0	37.5	37.5	00.0	100
27*	00.0	25.0	25.0	50.0	00.0	100

continue Table 4

Item	% Poor	% Fair	% Good	% Very Good	% Excellent	% Total
Visual Recognition						
28	00.0	00.0	00.0	75.0	25.0	100
29	00.0	00.0	00.0	75.0	25.0	100
30	00.0	00.0	00.0	87.5	12.5	100
Abstract Thinking						
31	00.0	12.5	25.0	62.5	00.0	100
32	00.0	00.0	25.0	75.0	00.0	100
33	00.0	00.0	25.0	75.0	00.0	100
34	00.0	00.0	12.5	87.5	00.0	100
35	00.0	00.0	25.0	75.0	00.0	100
36	00.0	00.0	25.0	62.5	12.5	100
37	00.0	00.0	50.0	50.0	00.0	100
38	00.0	00.0	25.0	75.0	00.0	100
39	00.0	12.5	37.5	50.0	00.0	100
40	00.0	00.0	12.5	87.5	00.0	100
Visual Interference						
41	00.0	00.0	25.0	75.0	00.0	100
42	00.0	00.0	12.5	87.5	00.0	100
43	00.0	00.0	25.0	75.0	00.0	100
44	00.0	00.0	25.0	75.0	00.0	100
45	00.0	00.0	25.0	75.0	00.0	100
46	00.0	00.0	12.5	87.5	00.0	100
Attention						
47	00.0	00.0	12.5	75.0	12.5	100
48	00.0	00.0	25.0	62.5	12.5	100
49	00.0	00.0	00.0	87.5	12.5	100
50	00.0	00.0	12.5	50.0	37.5	100
51	00.0	00.0	12.5	50.0	37.5	100

Continue Table 4

Item	% Poor	% Fair	% Good	% Very Good	% Excellent	% Total
Executive Function						
52	00.0	00.0	12.5	62.5	25.0	100
53	00.0	00.0	37.5	25.0	37.5	100
55	00.0	00.0	12.5	50.0	37.5	100
59	00.0	00.0	37.5	37.5	25.0	100
60	00.0	12.5	25.0	50.0	12.5	100
61*	00.0	25.0	12.5	62.5	00.0	100
62	00.0	12.5	25.0	62.5	00.0	100
Visual Inattention						
54	00.0	00.0	12.5	50.0	37.5	100
Similarity						
56	00.0	12.5	37.5	25.0	25.0	100
57	00.0	00.0	37.5	37.5	25.0	100
Sequence						
58	00.0	12.5	37.5	12.5	37.5	100
Memory						
63	00.0	12.5	25.0	37.5	25.0	100
64	00.0	37.5	00.0	37.5	25.0	100
65	00.0	37.5	12.5	25.0	25.0	100
Overall	00.0	12.5	12.5	75.0	00.0	100

* Item with rating < 85% in good or above

Pilot field testing of the ICAS

There were 14 subjects (11 male and 3 female) fulfilled the selection criteria, signed the consent form, and recruited into the pilot study. The age ranged from 60 to 86 with mean age 67 (S.D. = 7.5). Two of them were from sub-acute stage, 5 of them were recruited in rehabilitation stage and 7 of them from geriatric day hospital stage. 71.4% of subjects suffered from infarction and 28.6% suffered from hemorrhagic stroke. 42.9% affected the right side function and 57.1% affected left side function. The demographic result was shown in Table 5.

Table 5 Demographic of subjects in the pilot study

Item	Number (N)	Percentage
Gender		
Male	11	78.6
Female	3	21.4
Total	14	100
Type of stroke		
Infarct	10	71.4
Hemorrhage	4	28.6
Total	14	100
Affected side		
Right	6	42.9
Left	8	57.1
Total	14	100
	Mean (standard Deviation)	Range
Age	67.13 (7.472)	60-86

The results of the field test were divided into two parts; the first part was about the validity of ICAS and its prediction of the functional outcome in terms of MBI. All items in the ICAS obtained by the Rasch model and the formation of the cognitive ability measure or ICAS score in short form contributed to the second part of data analysis.

Concurrent Validity

The concurrent validity of ICAS was established by correlations with the two most commonly used cognitive assessments – MMSE–CV and the NCSE–CV. The Spearman's rho correlation coefficient between ICAS score and MMSE–CV score was 0.676 with *p* =0.011. The ICAS score also statistically correlated with the repetition, naming, construction and calculation areas of NCSE–CV score. Detailed results were also presented in Table 6.

Table 6 The Spearman's rho Correlation of Cognitive ability (ICAS score) to MMSE–CV and NCSE–CV

Items	Cognitive Ability (ICAS score)	p value
MMSE–CV*	0.676	0.011
NCSE–CV		
Orientation	0.492	0.124
Attention	0.597	0.053
Comprehension	0.384	0.244
Repetition**	0.850	0.001
Naming*	0.686	0.020
Construction*	0.633	0.037
Memory	0.449	0.166
Calculation*	0.710	0.014
Similarity	0.359	0.278
Judgment	-0.063	0.854

* $p < 0.05$, ** $p < 0.01$

In the regression model, upper limb function and ICAS score explained 77.6% variance of MBI with R^2 = 0.776. The strength of the relationship was tested by ANOVA with $F(2,10) = 13.86$ with $p = 0.003$. Both the regression coefficient of upper limb function and the ICAS score was significant from zero at 5% level (Table 7).

Table 7 Regression model for cognitive ability and upper limb function to functional outcome

Model	Unstandardized Coefficients		Standardized Coefficients		
	B	S.E.	Beta	t	Sig.
Constant	30.683	9.323		3.291	0.011
Upper Limb Function	7.456	1.720	0.731	4.335	0.002
Cognitive Ability (ICAS score)	5.783	2.393	0.408	2.417	0.042

Dependent Variable: MBI

R = 0.881, R Square = 0.776

ANOVA table

Model	Sum of Square	df	Mean Square	F	Sig.
Regression	5986.098	2	2993.049	13.857	0.003
Residual	1727.902	8	215.988		
Total	7714.000	10			

Rasch analysis of ICAS

There were 910 responses (65items from 14 subjects) collected and they were analyzed by software WINSTEPS. The dichotomous Rasch Model was used during the analysis. The results were divided into person measures (the subjects' cognitive abilities or the ICAS score) and item measures (the item difficulty measures of each testing items).

Person Measures

The result showed that all the responses matched with the Rasch model and there were no unspecified elements in the responses data. The cognitive ability measures of patient were shown in Table 8. In assessing the fitness of the data to the Rasch model, the criteria suggested by Wright and Linacre in 1994 were adopted in the interpretation. The INFIT and OUTFIT statistics were between 0.5 to 1.5 or their standardized values between -2 to 2. From Table 8, the result shows that all cognitive ability measures fit the Rasch model. The cognitive ability reliability index was 0.93 and the overall chi square was 11.4 with p = 0.49. It indicated that there was no association among patients' cognitive abilities. All subjects were independent to each others in the cognitive ability measures.

Table 8 Cognitive ability score (Person measures) in Phase I study

Obsvd Score	Obsvd Count	Obsvd Average	Fair-M Avrage	Measure	Model S.E.	Infit MnSq	Infit ZStd	Outfit MnSq	Outfit ZStd	PtBis	Nu Patient
27	59	.5	.41	-.37	.30	1.2	1	1.1	0	.33	1 Patient
24	59	.4	.35	-.64	.30	.9	0	.8	0	.41	2 Patient
13	59	.2	.15	-1.73	.34	1.1	0	1.1	0	.27	3 Patient
44	59	.7	.79	1.33	.36	1.2	1	1.3	0	.40	4 Patient
40	59	.7	.70	.87	.33	.9	0	1.0	0	.51	5 Patient
29	59	.5	.45	-.19	.30	1.0	0	.9	0	.42	6 Patient
36	59	.6	.61	.46	.31	1.0	0	.8	0	.48	7 Patient
48	59	.8	.87	1.91	.41	1.0	0	.8	0	.54	8 Patient
47	59	.8	.85	1.75	.39	.9	0	.6	-1	.60	9 Patient
58	59	1.0	1.00	5.78	1.11	1.2	0	.2	0	.51	10 Patient
47	59	.8	.85	1.75	.39	.8	0	.5	-1	.65	11 Patient
48	59	.8	.87	1.91	.41	.8	0	.7	0	.62	12 Patient
42	59	.7	.75	1.09	.34	1.1	0	.8	0	.48	13 Patient
49	59	.8	.89	2.08	.42	1.1	0	1.2	0	.50	14 Patient
39.4	59.0	.7	.68	1.14	.41	1.0	.1	.9	-.3	.48	Mean (Count: 14)
11.8	.0	.2	.24	1.70	.20	.1	.6	.3	.6	.10	S.D.

RMSE (Model) .45 Adj S.D. 1.64 Separation 3.60 Reliability .93*

Random (normal) chi-square: 11.4 d.f.: 12 significance: .49**

* Cognitive Reliability Index = 0.93

** $p > 0.05$

Item Measures

The item difficulty measures of the item bank of ICAS were shown in Table 9. Item 16, 18, 22, 26, 29 40, 43, 45, 46, 53, 58 and 65 got INFIT and OUTFIT statistics out of the range. It implied that they were not fit for the Rasch model. The item reliability index was 0.73 and the overall chi square was 48.5 with $p = 0.78$. Thus there was no association among items, or all the items were independent to each others.

Figure 6 showed both patients' cognitive ability measure and items difficult measures in the same scale.

Table 9 Item difficulty measures for the 65 items in the ICAS

Obsvd Score	Obsvd Count	Obsvd Average	Fair-M Avrage	Measure	Model S.E.	Infit MnSq	Infit ZStd	Outfit MnSq	Outfit ZStd	PtBis	Items
6	14	.4	.43	1.44	.63	1.0	0	.8	0	.50	01
8	14	.6	.62	.65	.63	1.4	1	1.2	0	.24	02
9	14	.6	.71	.24	.65	.8	0	.7	0	.61	03
14	14			(-3.66	1.89)	Minimum				.00	04
8	14	.6	.62	.65	.63	1.1	0	1.1	0	.40	05
8	14	.6	.62	.65	.63	.6	-1	.5	0	.76	06
9	14	.6	.71	.24	.65	.7	0	.6	0	.68	07
14	14			(-3.66	1.89)	Minimum				.00	08
10	14	.7	.79	-.21	.69	.7	-1	.5	0	.73	09
7	14	.5	.52	1.05	.63	.7	-1	.6	0	.69	10
10	14	.7	.79	-.21	.69	.9	0	1.0	0	.52	11
11	14	.8	.87	-.72	.75	1.2	0	.9	0	.34	12
12	14	.9	.92	-1.36	.86	.7	0	.3	0	.66	13
9	14	.6	.71	.24	.65	1.3	0	1.4	0	.22	14
13	14	.9	.97	-2.29	1.11	1.2	0	.6	0	.27	15
12	14	.9	.92	-1.36	.86	1.6	0	1.3	0	.02	16*
11	14	.8	.87	-.72	.75	.9	0	.6	0	.57	17
11	14	.8	.87	-.72	.75	.5	-1	.3	0	.79	18*
11	14	.8	.87	-.72	.75	.7	0	.5	0	.65	19
12	14	.9	.92	-1.36	.86	.7	0	.4	0	.62	20
11	14	.8	.87	-.72	.75	.9	0	.8	0	.53	21
12	14	.9	.92	-1.36	.86	.6	0	.3	0	.71	22*
11	14	.8	.87	-.72	.75	1.0	0	1.3	0	.39	23
11	14	.8	.87	-.72	.75	.8	0	.6	0	.59	24
10	14	.7	.79	-.21	.69	1.4	1	1.2	0	.22	25
12	14	.9	.92	-1.36	.86	1.7	1	3.1	0	-.20	26*
6	14	.4	.43	1.44	.63	.8	-1	.6	0	.63	27
11	14	.8	.87	-.72	.75	.9	0	.6	0	.57	28
10	14	.7	.79	-.21	.69	1.8	1	2.0	0	-.08	29*
12	14	.9	.92	-1.36	.86	1.4	0	1.1	0	.14	30
6	14	.4	.43	1.44	.63	1.3	1	1.3	0	.22	31
9	14	.6	.71	.24	.65	.7	-1	.7	0	.69	32
8	14	.6	.62	.65	.63	.8	0	.7	0	.62	33
11	14	.8	.87	-.72	.75	.9	0	.8	0	.53	34
9	14	.6	.71	.24	.65	1.1	0	1.0	0	.42	35
10	14	.7	.79	-.21	.69	1.1	0	1.2	0	.38	36
2	14	.1	.08	3.59	.97	.9	0	.8	0	.38	37
12	14	.9	.92	-1.36	.86	.7	0	.4	0	.62	38
6	14	.4	.43	1.44	.63	1.2	0	1.2	0	.32	39
9	14	.6	.71	.24	.65	2.0	2	2.3	0	-.20	40*
13	14	.9	.97	-2.29	1.11	.6	0	.2	0	.61	41
11	14	.8	.87	-.72	.75	.8	0	.7	0	.56	42
13	14	.9	.97	-2.29	1.11	.6	0	.2	0	.61	43*
10	14	.7	.79	-.21	.69	1.2	0	1.1	0	.35	44
11	14	.8	.87	-.72	.75	.5	-1	.3	0	.79	45*
12	14	.9	.92	-1.36	.86	.6	0	.3	0	.71	46*
11	14	.8	.87	-.72	.75	1.2	0	1.1	0	.31	47
12	14	.9	.92	-1.36	.86	.7	0	.3	0	.66	48
7	14	.5	.52	1.05	.63	1.1	0	1.0	0	.40	49
9	14	.6	.71	.24	.65	1.1	0	1.1	0	.45	50
12	14	.9	.92	-1.36	.86	.7	0	.3	0	.66	51
7	14	.5	.52	1.05	.63	.7	-1	.5	0	.71	52
1	14	.1	.02	4.95	1.39	.2	-1	.1	0	.42	53*
12	14	.9	.92	-1.36	.86	1.2	0	.6	0	.37	54
6	14	.4	.43	1.44	.63	1.0	0	.8	0	.48	55
11	14	.8	.87	-.72	.75	1.2	0	1.1	0	.31	56
7	14	.5	.52	1.05	.63	1.2	0	1.8	0	.24	57
2	14	.1	.08	3.59	.97	2.3	1	1.7	0	.22	58*
7	14	.5	.52	1.05	.63	1.2	0	1.1	0	.33	59
14	14			(-3.66	1.89)	Minimum				.00	60
12	14	.9	.92	-1.36	.86	1.2	0	.6	0	.37	61
14	14			(-3.66	1.89)	Minimum				.00	62
0	14			(6.84	2.05)	Maximum				.00	63
0	14			(6.84	2.05)	Maximum				.00	64
1	14	.1	.02	4.95	1.39	.2	-1	.1	0	.42	65*

Obsvd Score	Obsvd Count	Obsvd Average	Fair-M Avrage	Measure	Model S.E.	Infit MnSq	Infit ZStd	Outfit MnSq	Outfit ZStd	PtBis	Items
9.4	14.0	.7	.71	.00	.77	1.0	-.1	.9	-.1	.41	Mean (Count: 65)
3.4	.0	.2	.27	1.53	.17	.4	.9	.6	.3	.25	S.D.

RMSE (Model) .79 Adj S.D. 1.30 Separation 1.65 Reliability .73**

Random (normal) chi-square: 48.5 d.f.: 57 significance: .78^

* Items out of range and not fit for Rasch model

** Item reliability index

^ $p > 0.05$

Figure 6 Person measures and Item measures in the same scale

```
Measure   persons - MAP - items
 Logits  <More Able  |<Most Difficult Item>
           Person>
  6                 +  63     64
               10   |
                    |
                    |
                    |
                    |
  5                 +  53     65
                    |
                    |
                   T|
                    |
                    |
  4                 +
                    |
                    |  37     58
                    |
                    |
                    |
  3                 +T
                   S|
                    |
                    |
                    |
               15   |
  2                 +
    08  09  12  13  |
                    |
                    |S 01     27     31     39     55
               04   |
               14  M|
  1                 +  10     49     52     57     59
               05   |
                    |  02     05     06     33
               07   |
                    |
                    |  03     07     14     32     35     40     50
  0                 +M
               06   |  09     11     25     29     36     44
               01   |
                   S|
               02   |  12     17     18     19     21     23     24     28     34     42     45     47     56
                    |
 -1                 +
                    |
                    |  13     16     20     22     26     30     38     46     48     51     54     61
                    |S
               03   |
                    |
 -2                 +
                   T|
                    |  15     41     43
                    |
                    |
                    |
 -3                 +T 04     08     60     62
   <less Able Person>|<Most Easy Items>
```

Phase II Study – Development of the item difficult trait level

Phase II study was to investigate and validate the item difficulty measures of testing items of the ICAS by the Rasch Model and assess the stability of the item difficulty measure of the ICAS testing items.

Demographic data of subjects

In this phase, 30 subjects were successfully recruited into the study. Altogether 11 were in recruited from the sub-acute stage, 10 from the rehabilitation stage and 9 from the community phase. The age of subjects ranged from 60 to 80 years (mean = 71.7, SD = 7). Twenty-five of them suffered from cerebral infarction and five of them suffered from hemorrhagic stroke. The distribution of the affected side was equal, i.e, 15 with the right side affected and 15 with the left side affected. There were no significant differences among the demographic data across the different phases, as shown by the chi square statistics (Table 10).

Table 10 Distribution of demographics in the three phases of stroke rehabilitation

	Sub-acute phase	Rehabilitation phase	Community phase	Chi-square statistic	p value
Gender					
Male	3	5	4		
Female	6	5	7	0.644^	0.724
Type of Stroke					
Infraction	8	7	10		
Hemorrhage	1	3	1	3.297^	0.570
Side affected					
Left	4	6	5		
Right	5	4	6	0.602^	0.811

^p > 0.05

Development of item difficulty measures

The 65 items of the ICAS item bank were arranged hierarchically on a linear scale according to their item difficulty measures (Table 11). Their Cronbach's alpha was 0.88 and the item reliability index was 0.87.

Table 11 Item difficulty measures for the ICAS item bank.

Item	Measures	Item	Measures	Item	Measures	Item	Measures
1	0.64	18	1.13	35	-1.50	52	0.48
2	0.16	19	-2.31	36	0.32	53	3.26
3	1.13	20	0.48	37	2.89	54	-3.06
4	-3.06	21	0.16	38	-0.96	55	1.48
5	-0.01	22	-1.21	39	2.08	56	-0.01
6	-0.36	23	-0.01	40	1.30	57	-0.01
7	-2.31	24	1.13	41	-4.30	58	3.75
8	-4.30	25	0.97	42	-1.21	59	0.80
9	-0.18	26	0.64	43	-0.01	60	-4.30
10	-0.01	27	1.67	44	0.16	61	-0.18
11	1.30	28	-1.21	45	-0.01	62	-4.30
12	-0.75	29	-0.36	46	-0.36	63	4.54
13	-0.96	30	-0.96	47	-1.50	64	5.80
14	-0.96	31	1.67	48	-0.18	65	5.80
15	-2.31	32	-0.36	49	0.97		
16	-1.21	33	-0.18	50	-0.96		
17	-1.84	34	-0.36	51	-2.31		

Item 1 – 3 Working memory; Item 4 – 10 Orientation to time; Item 11 – 20 Semantic Memory; Item 21 – 27 Calculation; Item 28 – 30 Visual Recognition; Item 31 – 40 Abstract Thinking; Item 41 – 46 Visual Interference; Item 47 – 51 Attention Span; Item 52 – 53, 55, 59 – 62 Executive Function; Item 54 Visual Inattention; Item 56 -57 Similarity Categorization; Item 58 Sequence; Item 63 – 65 Memory

The item difficulty measures analyzed by the Rasch model revealed that all of the assessment items fit the model, with the INFIT statistics ranging from 0.5 to 1.5 (Figures 7 & 9). The OUTFIT statistics showed that 10 items fell out of the 0.5-1.5 range (Figures 8), but their OUTFIT t-statistics were within the range from -2 to +2, except for items 3, 11, 14 and 39 (Figure 10).

Figure 7 INFIT mean squares for the ICAS item bank.

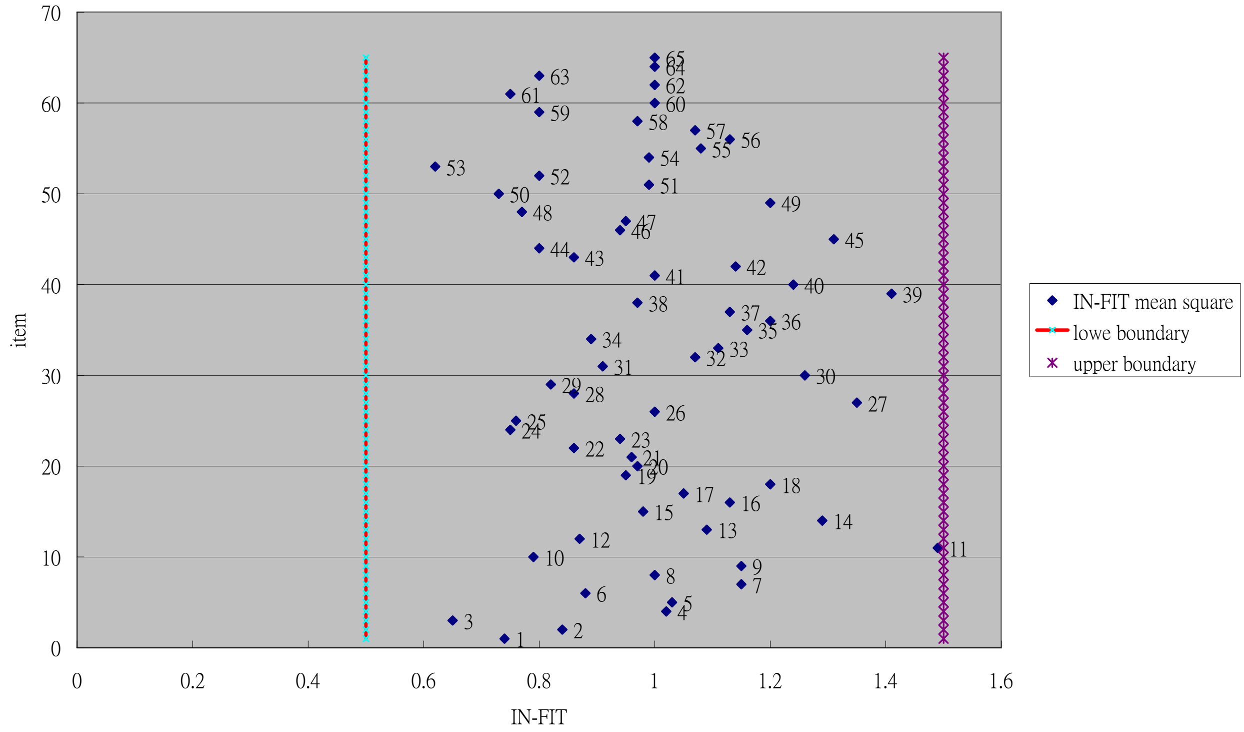


Figure 8 OUTFIT mean squares for the ICAS item bank.

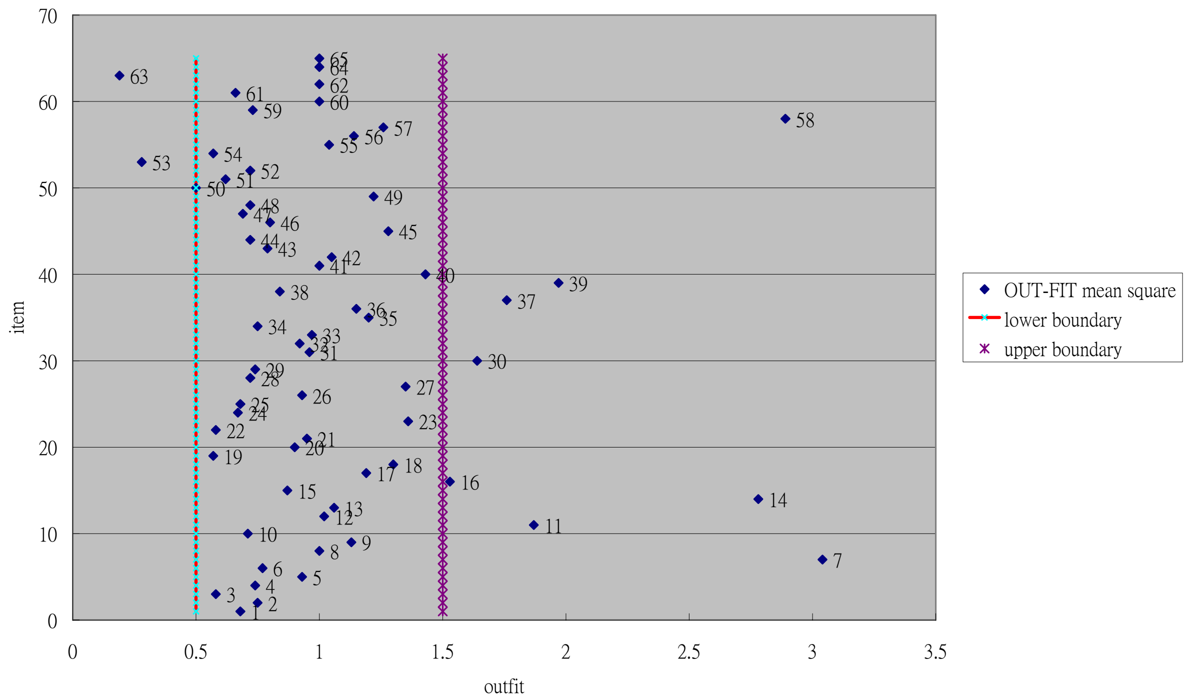


Figure 9 INFIT t-score for the ICAS item bank

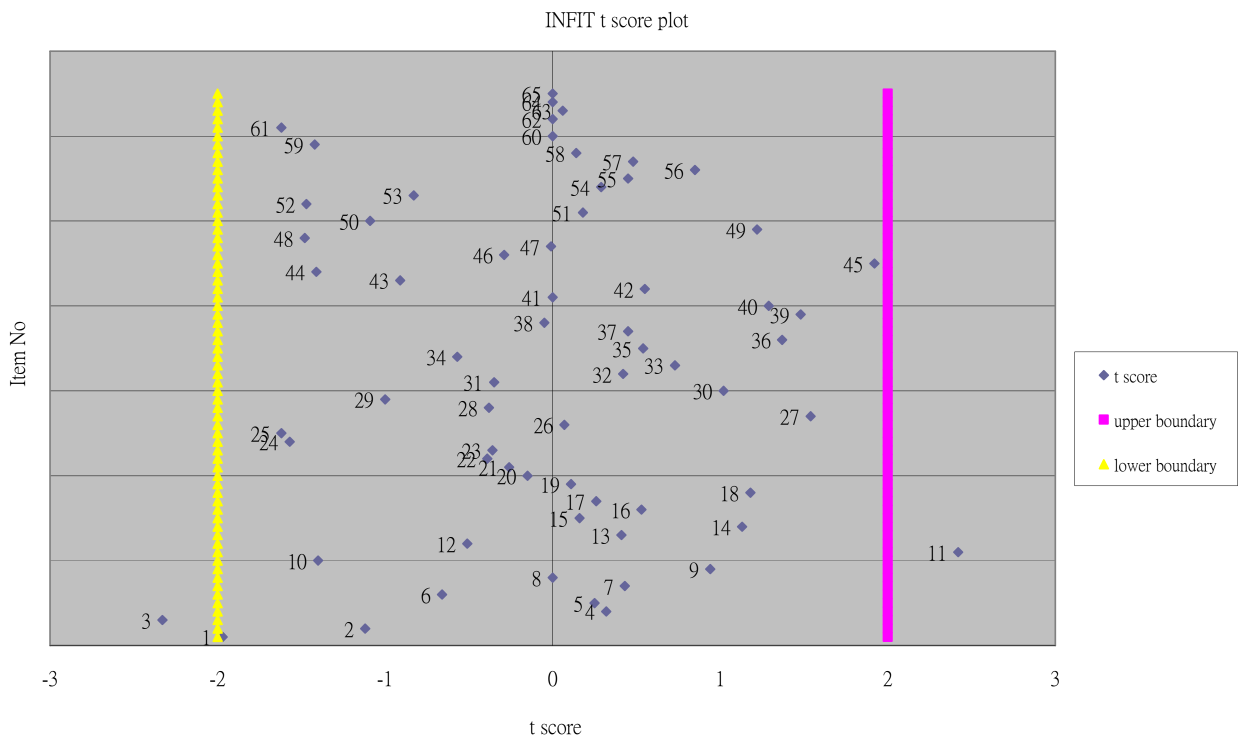
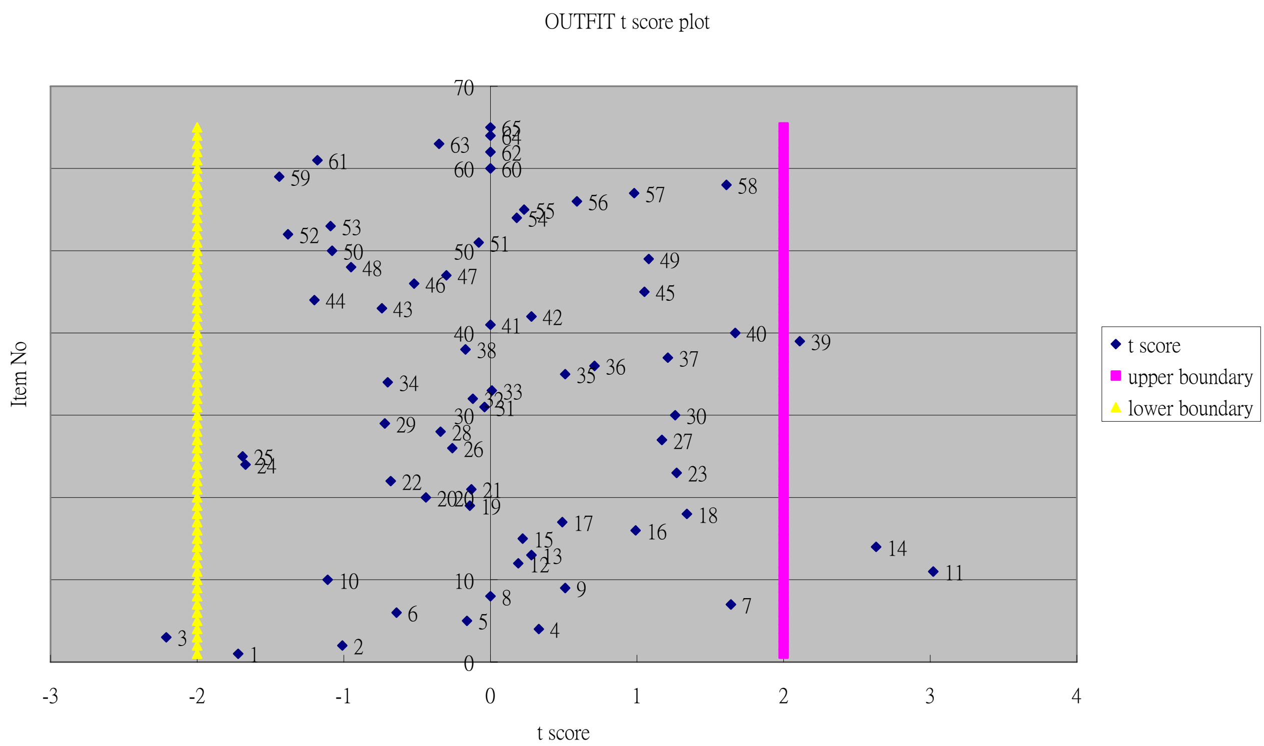


Figure 10 OUTFIT t-scores for the ICAS item bank.



Principal components analysis (PCA) of the residuals of the ICAS item bank showed that the measures explained 66.1% of the variance and the first contrast explained 3.7% of the variance (Figure 11). Moreover, the standardized residual variance plot shown that the variance explained by contrasts were relatively small when comparing with the variance explained by the measure (Figure 12). These indicators showed that the item bank of the ICAS fitted the criteria of uni-dimensionality.

Figure 11 Principal components analysis of the residuals of the ICAS item bank

```
       STANDARDIZED RESIDUAL CONTRAST 1 PLOT
Table of STANDARDIZED RESIDUAL variance (in Eigenvalue units)
                                                 Empirical       Modeled
Total variance in observations     =        174.2 100.0%         100.0%
Variance explained by measures     =        115.2  66.1%          67.8%
Unexplained variance (total)       =         59.0  33.9% 100.0%   32.2%
Unexplned variance in 1st contrast =          6.4   3.7%  10.9%

      -4     -3     -2     -1      0      1      2      3      4      5
      ++------+------+------+------+------+------+------+------+------++ COUNT
   .7 +           A        B      |                                   + 2
      |                           |C                                  | 1
   .6 +                        D  |E                                  + 2
      |                           |                                   |
C  .5 +                 F         |                                   + 1
O     |                           |G                                  | 1
N  .4 +                        I  |  H                                + 2
T     |                         K |       J                           | 2
R  .3 +           L               M                                   + 2
A     |                    O      |     P       N                     | 3
S  .2 +                         R | Q    S                            + 3
T     |                   T       U  V                                | 3
   .1 +                   X       |       W                           + 2
1     |                          YZ                      1            | 3
   .0 +---------------------------|-------1-----------------------1---+ 2
L     |           1       1       |        1                          | 3
O -.1 +      1      x     z       v       y  w                        + 6
A     |                        t  |   u                               | 2
D -.2 +      p    q     s      r  |                                   + 4
I     |                    o      |                                   | 1
N -.3 +                    n      |   m                               + 2
G     |                   k       l                     j             | 3
  -.4 +                           |     g i  h       f                + 4
      |                    b    e cd                                  | 4
  -.5 +                           |                                   +
      |                           |                                   |
  -.6 +                          a|                                   + 1
      ++------+------+------+------+------+------+------+------+------++
      -4     -3     -2     -1      0      1      2      3      4      5
                                 item MEASURE
 COUNT:      2    4  1  2 45 1  5473122 12321 2 1    1  1  1     1
```

Figure 12 Standardized residual variance plot of the ICAS items

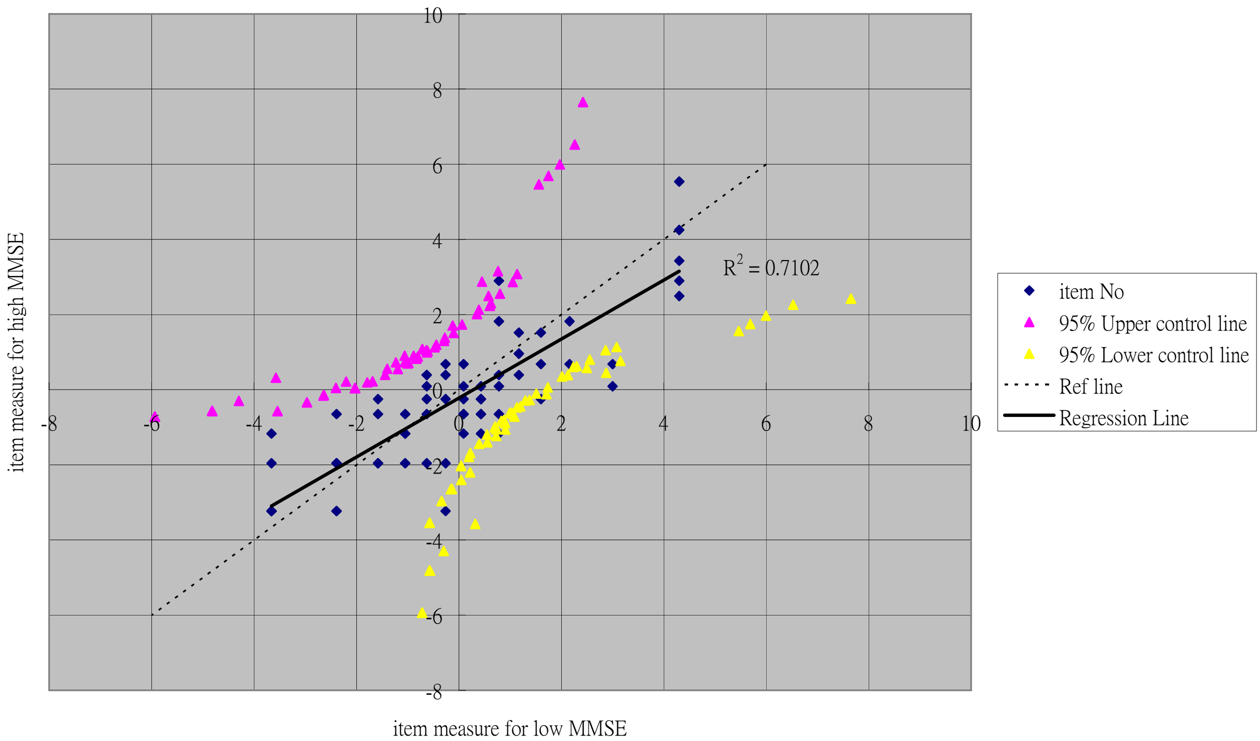
```
       STANDARDIZED RESIDUAL VARIANCE SCREE PLOT
Table of STANDARDIZED RESIDUAL variance (in Eigenvalue units)
                                                 Empirical        Modeled
Total variance in observations     =        174.2 100.0%          100.0%
Variance explained by measures     =        115.2  66.1%           67.8%
Unexplained variance (total)       =         59.0  33.9% 100.0%    32.2%
Unexplned variance in 1st contrast =          6.4   3.7%  10.9%
Unexplned variance in 2nd contrast =          5.3   3.0%   8.9%
Unexplned variance in 3rd contrast =          4.5   2.6%   7.7%
Unexplned variance in 4th contrast =          4.3   2.5%   7.3%
Unexplned variance in 5th contrast =          3.6   2.1%   6.1%

      VARIANCE COMPONENT SCREE PLOT
       +--+--+--+--+--+--+--+--+--+
   100%+  T                       +
       |                          |
  V 63%+     M                    +
  A    |                          |
  R 40%+                          +
  I    |        U                 |
  A 25%+                          +
  N    |                          |
  C 16%+                          +
  E    |                          |
    10%+                          +
  L    |                          |
  O  6%+                          +
  G    |                          |
  |  4%+                          +
  S    |           1              |
  C  3%+              2  3        +
  A    |                    4  5  |
  L  2%+                          +
  E    |                          |
  D  1%+                          +
       |                          |
   0.5%+                          +
       +--+--+--+--+--+--+--+--+--+
          TV MV UV U1 U2 U3 U4 U5
         VARIANCE COMPONENTS
```

Stability of the testing items of the ICAS

Testing the stability of the item measures between the subject groups, as classified by MMSE–CV score (≤ 21 or > 21), revealed that there were no statically significant differences among the item measures ($t = 1.997$, $p > 0.05$), and the correlation of the item measures between the two groups was 0.843. Plotting the item measures of the higher MMSE–CV group against those of the lower MMSE–CV group showed that all of the items were within a 95% confidence interval region (Figure 13).

Figure 13 Stability of item difficulty between high and low MMSE–CV groups.



Concurrent validity of the ICAS

The cognitive ability scores of patients derived from Rasch analysis were highly correlated with their MMSE–CV scores ($p < 0.001$). Correlation analysis of each subject's cognitive ability score with his or her performance in the NCSE–CV showed that the cognitive ability score correlated with all of the components of the NCSE–CV at the 5% level of significance, except for attention (which was at 5.5% level of significance). The cognitive ability score was highly correlated with higher cortical functions tested in the NCSE–CV including comprehension, calculation, construction, judgment and orientation ($p < 0.01$), and moderately correlated with repetition, naming, memory and similarity ($p < 0.01$). The correlation of the cognitive ability scores of patients with their NCSE–CV overall scores; which obtained by using Rasch analysis of NCSE—CV components; was 0.876 at the 1% level of significance (Table 12).

Table 12 The Spearman's rho correlation coefficient of ICAS with MMSE–CV and NCSE–CV in Phase II study

	ICAS	p-value
MMSE–CV *	0.760	0.000
NCSE–CV		
Orientation*	0.688	0.000
Attention[#]	0.353	0.055
Comprehension**	0.727	0.000
Repetition*	0.595	0.001
Naming*	0.534	0.002
Construction*	0.636	0.000
Memory**	0.530	0.003
Calculation*	0.786	0.000
Similarity*	0.521	0.003
Judgment*	0.709	0.000
NCSE–CV Overall (by Rasch analysis)**	0.876	0.000

*$p < 0.05$;

**$p < 0.01$;

[#]$p = 0.055$.

Phase III study – Psychometric properties of ICAS and the ANN model for outcome prediction

The Phase III study was to further confirm the psychometric properties of the ICAS with the golden standard of cognitive assessments – MMSE–CV. The cutoff point of the ICAS would be computed to see if it could be used to screen out patients with cognitive impairment. A back propagation artificial neural network (BP–ANN) was built to predict the functional outcome of the subjects by using age, sex, side of stroke affect, type of stroke, residual upper limb function, cognitive ability and residual function outcome as predictors.

Demographic data of subjects

In this Phase, 66 subjects were successfully recruited into the study. They had age ranged from 60 to 93 (mean = 72.8; SD = 8.8). Fifty-nine of them had suffered from a cerebral infarction and seven of them suffered from hemorrhagic stroke. Twenty-five subjects were classified as cognitively impaired and 41 as non-cognitively impaired by the MMSE–CV cutoff score (Chiu, et al., 1994). Borderline significant differences were found between the two groups in terms of age, sex and type of stroke but no significant difference was found in the side that stroke affected (Table 13).

Table 13 Baseline comparison of the demographic data of the cognitively and non-cognitively impaired groups

	Cognitively impaired group (n = 25)	Non-cognitively impaired group (n = 41)	*p* value
Gender			
Male	7	26	
Female	18	15	0.05
Type of stroke			
Infarction	20	39	
Hemorrhage	5	2	0.053
Side affected			
Right	10	23	
Left	15	18	0.205
Age (mean)	75.4	71.15	0.053
ICAS (mean)	0.94	3.65	0.001*
MMSE (mean)	16.32	26.20	0.001*
MMSE (range)	6-21	26-30	

*p < 0.01

Psychometric Properties of ICAS

<u>Reliability</u>

The ICAS proved to have good internal consistency (Cronbach's alpha = 0.878) and test-retest reliability (Pearson correlation coefficient = 0.789; $p < 0.01$), which indicated that it was a reliable measure. The Pearson correlation coefficient of the MMSE–CV was 0.757 ($p < 0.01$), which demonstrated the concurrent validity of the ICAS with the MMSE–CV.

<u>Sensitivity, Specificity and Cutoff point of ICAS for cognitive impairment in stroke</u>

The ROC curve of the ICAS was showed in Figure 14, and the AUC was 0.909. The sensitivity and specificity rates for various cutoff scores of the ICAS are shown in Table 14. A cutoff score of 3.02 logits yielded a high sensitivity rate of 80.5% and specificity rate of 96%. This cutoff score appeared to be the best cutoff for the stroke sample.

Figure 14 ROC curve for the ICAS.

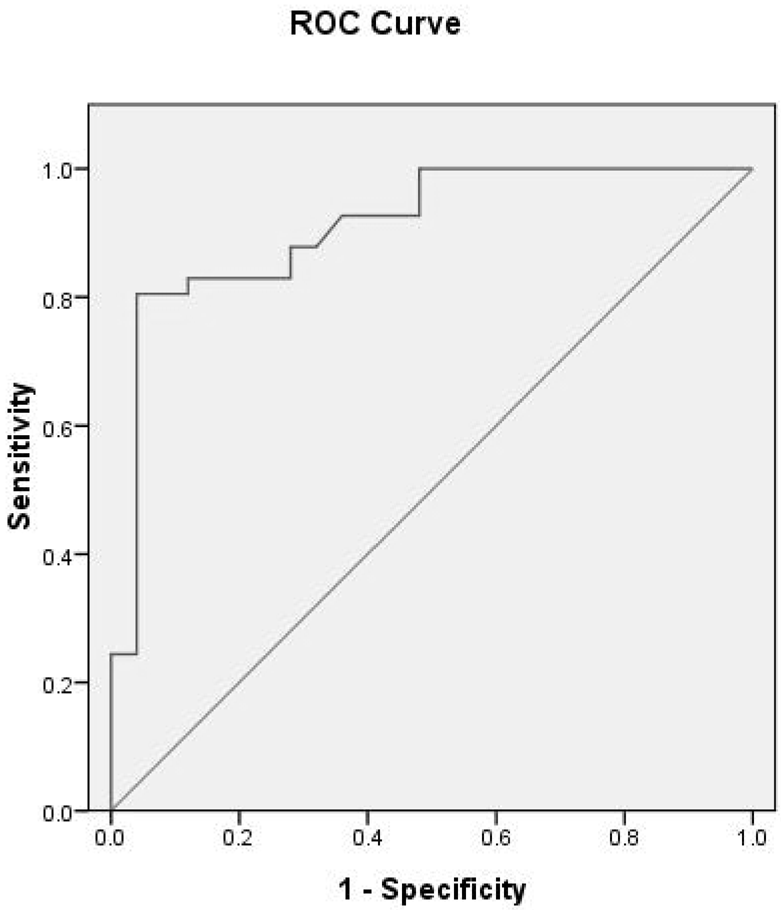


Table 14 Cutoff score and corresponding sensitivity and specificity of the ICAS

ICAS score (logits)	Sensitivity (%)	Specificity (%)
0.8	100	52
1.47	92.7	56
2.07	87.8	68
2.35	85.4	72
2.79	82.9	88
2.94	80.5	92
3.02	80.5	96
3.08	78	96
3.12	75.6	96
3.2	73.2	96
4.54	24.4	100

Note: Only some of the scores are listed.

To ensure that the results were not due to chance, a power analysis by PASS2008 was also carried out. It was found that the sample sizes of 25 and 41 subjects for the cognitively and the non-cognitively impaired group, respectively, achieved 97% power to detect a difference of 0.0900 between the AUC of the ICAS and that of the MMSE–CV (using a two-sided z test at a significance level of 0.05). Thus, the power of study was maintained, even though the sample sizes were unequal.

A further analysis was conducted to investigate how many cases were cognitive impaired for the same pool of subjects if classified by the ICAS. The result showing that classified by the ICAS, there were 32 cases were cognitive impaired and 34 were non cognitive impaired. That means 7 subjects classified by MMSE is not cognitive impaired but classify by ICAS is cognitive impaired. Their MMSE scores ranged from 22 to 25. It was hard to conclude that ICAS was more sensitive than MMSE, as this was conduct in the same group of client that used to find out the cutoff of ICAS, but provide some information on the use of ICAS as a screening tool for cognitive impairments.

Predictability of ICAS by BP–ANN model

Four BP–ANN models with 5, 10, 15 and 20 neurons in the hidden layer were built and the correlation between the observed values and the predicated values of MBI were presented in Table 15. BP–ANN with 15 hidden neurons was found to be the most forecasting network (with the correlation coefficient at 0.85 between the observed and the predicted MBI scores; see Figure 15). The observed MBI score, predicted MBI score and their differences in value were plotted in the same graph for the 7-15-1 BP–ANN network (Figure 16). The detailed plots between observed and predicted MBI score by BP–ANN model with different hidden neurons were shown separately in Appendix IV.

Table 15 The correlation coefficient between observed and predicted MBI in BP–ANNs

BP–ANN models	Correlation coefficient between observed and predicted MBI
7-5-1	0.53
7-10-1	0.78
7-15-1	0.85
7-20-1	0.67

Figure 15 The observed and predicted MBI value by 7-15-1 BP–ANN model

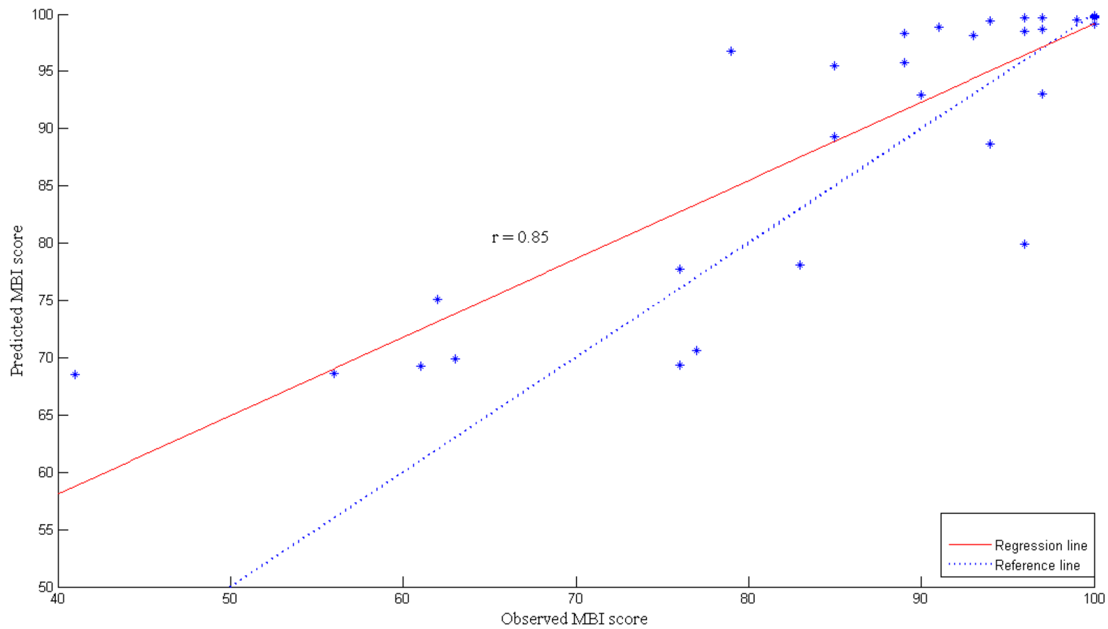
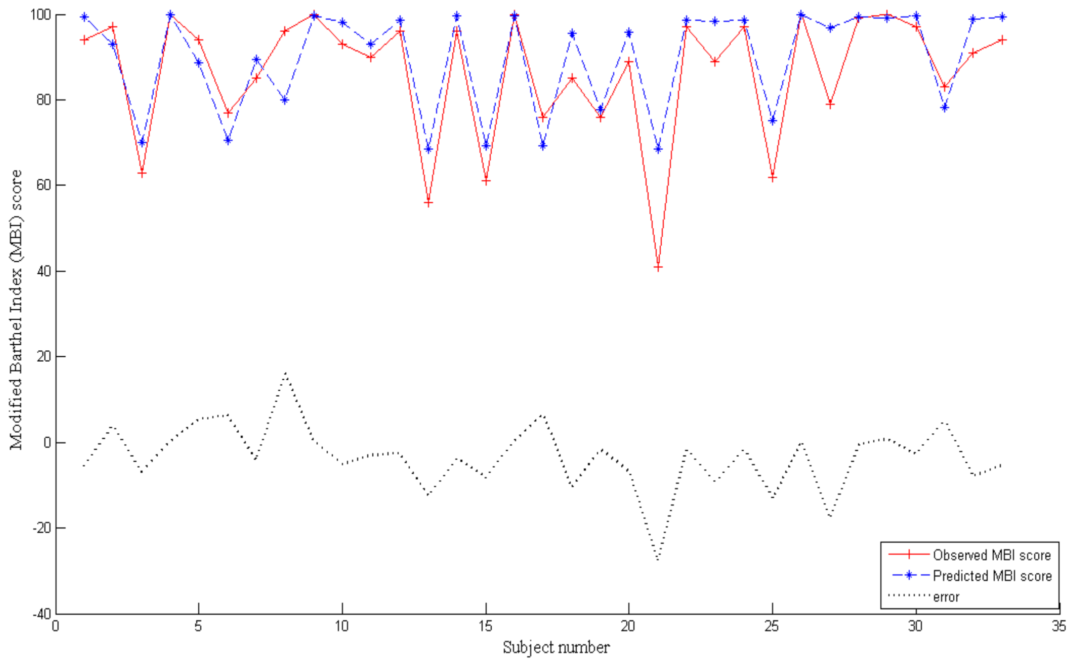


Figure 16 The observed, predicted MBI score and their difference by 7-15-1 BP–ANN model



In conclusion, the ICAS went through three phases of validation study and the results seemed to support that the ICAS could be an valid and reliability tool to assess cognitive functions of stroke survivors and might serve as a potential outcome indicator to monitor the personal progress and document the effectiveness of the treatment programme (based on one of its characteristic that it could pull different constructs of cognitive components into a linear ratio scale for comparisons). Moreover, the application of BP–ANN in forecasting the functional outcome was an innovative attempt in the field of rehabilitation and may shed light in further development in the near future.

Chapter 6 Discussion

The study successfully developed and validated an intelligent cognitive assessment system (ICAS) which can now provide an interactive testing environment which simulates the real life environment to motivate stroke patients to achieve the best participation and maximum performance. One of the key features of the ICAS is having a linear scale score which provides the advantages of precise monitoring of the progress of stroke patients' cognitive functions and may serve as an important outcome indicator for rehabilitation programmes. The ICAS score can be obtained by considering comprehensive cognitive components and easily transforming a total of 13 different cognitive domains. Moreover, based on the artificial neural network (ANN), the system is empowered to predict the functional outcomes of stroke survivors, which are based on their residual functions and demographic variables.

Thus, after presenting the key results in the last chapter, this chapter will further discuss the findings according to individual phases. The limitations, implications, and conclusions of the study will follow in the latter parts. The purposes of different phases are reiterated here again to facilitate discussion. The Phase I study developed the content validity and pilot testing of the ICAS to obtain data for estimation and planning of the Phase II and Phase III studies. The Phase II study aimed to investigate the item difficulty measures, item stability, and the unidimensionality of the testing items of the ICAS. Phase III investigated the cutoff score for the ICAS to screen out the cognitive impairment for stroke survivors and to validate the BP-ANN predictive model to predict the functional outcome for stroke survivors.

Phase I: Content validity and pilot testing of the ICAS

The Phase I study provided preliminary validity information on the ICAS supporting its utility for assessing the cognitive function of stroke survivors. This piece of information was similarly reflected by a motor study of the correlation between visual-motor skills of the affected arm evaluated with a computerized motor-skill analyser (CMSA) and clinical test of upper extremity function in patients with stroke (Yamanaka et al., 2005). The content validity of the ICAS items by the panel member was good, except 7 items got fair agreement among panel members. After discussion among the panel members, we made some amendments on these items and finally decided to keep these items in the item bank of the ICAS due to two reasons. The first one was that after amendment panel members agree that these items got its clinical significance and second reason was these items fit into the Rasch model in the Phase II of the study. The construct validity of the ICAS items and the concurrent validity with MMSE-CV were demonstrated and they provided supportive information to implement Phase II and Phase III studies. The regression model in Phase I showed that the cognitive ability measure was found to be a significant predictor of the functional outcome and explained 73.65% of the variance of the MBI in the regression model (after adjustment for the sample size and number of predictors) with the upper limb function test. The cognitive ability measures thus represented the cognitive function while the upper limb function represented the physical components. Therefore, this model seems to suggest that the self-care function could be an integration of the cognitive and physical components. The results also indicated that the ICAS could be an alternate way to assess the cognitive function of stroke survivors.

We also targeted stroke survivors aged 60 or above, who were a bit younger than in the common definition of elderly as people aged 65 or above. This selection was due to the fact that the prevalence rate of vascular dementia after stroke is reported to be 24% in people aged 60 to 69 and 23% in those aged 70–79 years (Lowery et al, 2002; Sachdev et al., 2006). So the results may not fully generalize to younger adults with stroke. Another reason for targeting stroke survivors was that cognitive decline is associated with stroke. Therefore, our study recruited subjects aged 60 or above in order to cover a larger group of subjects who possibly had a higher chance of suffering from cognitive impairment.

In terms of fluency of administration of the ICAS to stroke survivors, most of them could respond according to the instructions given by the ICAS. A few of them needed supervision by an assistant but none required physical assistance. For those elderly with limited reading ability, the use of graphic were much easier for them to understand (Lewis, 1996). Therefore, some adjustments were made to the ICAS in terms of the graphic presentation and the layout of materials presented in each item to facilitate the fluency of administration and user-friendly interface of the ICAS.

Phase II: Investigate the item difficulty measures, item stability, and the unidimensionality of the testing items of the ICAS

As reported in the Results section, Phase II of the study successfully established the item difficulty measures and investigated the unidimensionality of the 65-item test bank of the newly developed ICAS. The concurrent validity of the ICAS with both the MMSE-CV and NCSE-CV was established. It was also revealed that the ICAS was able to evaluate higher cortical functions better, as its measurement of cognitive ability was highly correlated with relevant higher thinking components (comprehension, calculation, construction, and judgement of the NCSE-CV), as stated by Radomski (2008), rather than with primary cognitive capacity (orientation, repetition, naming, memory, and similarity with the NCSE-CV) (Radomski, 2008). The reason may be due to the fact that ICAS adapted an integrative functional approach to assessment where some testing items required the stroke patients to perform some functional tasks (for example, keying in the password at the door entrance gate and using a mobile phone) rather than just assessing the basic skills for specific cognitive components.

In analysis of fit of the item bank of the ICAS, we found that some items did not fit the criteria of the OUTFIT statistics, but we still kept them in the item bank for two reasons. First, all of these items fit the INFIT statistics, and second, they had relevant clinical meaning in the assessment. For example, item 3 assesses the use of a mobile phone, a common activity in daily life. Although the operation procedures may be novel for some older adults, the use of this item may justify further observation. Items 11 and 14, which assess the semantic memory of a traditional Chinese festival, are overtly meaningful and ecologically valid in the Chinese culture. Item 39 assesses the response to an abstract sign of a wheelchair, and it is considered to be important to test abstract thinking. Also, we paid more attention to the INFIT than the OUTFIT statistics because the latter were influenced by outliers, which could easily be remedied and were less of a threat to measurement (Bond & Fox, 2007). In addition, the INFIT and OUTFIT statistics adopt slightly different techniques to assess the item fit to the Rasch model. The former give more weight to the performance of persons closer to the item value whereas the latter are not weighted. Therefore, the OUTFIT statistics are more sensitive to the influence of outlying scores (Bond & Fox, 2007).

With regard to the structure of the ICAS item bank, it assesses 13 cognitive domains. The findings of the hierarchy of the item difficulty measures showed that orientation, attention, and semantic memory were relatively easier than calculation, abstract thinking, or sequencing. Our findings were consistent with Ben-Yishay's cognitive interaction hierarchy hypothesis that attention, orientation, and memory are higher cognitive functions (Goldstein & Levin, 1987), and supported Radomski's contention that orientation, attention, and memory are components of primary cognitive capacity and prerequisites to higher-level thinking ability and meta-processing such as reasoning, concept formation, and problem solving (Radomski, 2008).

We also found that there were different levels of difficulty measures in the same cognitive domain. For example, for orientation, different kinds of responses resulted in different levels of item difficulty measures. Higher-level difficulty items required the stroke patients to input answers by themselves, whereas they could just pick one answer from several possible ones on the screen for the lower-level difficulty items. Therefore, the orientation aspect of the ICAS contains a spectrum of items that are capable of distinguishing patients' abilities. Similarly, each cognitive component of the ICAS had different item difficulty measures (see page 86, Table 11, Chapter 5). As a result, the item bank coverage of each aspect facilitated the CAT procedure in the

ICAS.

In the aspects of stability of the ICAS item bank, theoretically, test item difficulty measures should remain the same among different groups of stroke patients, so that they can distinguish their different abilities. The present study revealed that the item difficulty measures of patients who had passed or failed the MMSE-CV were within the 95% confidence region, which was an acceptable range of invariance. Therefore, the item difficulty measure was stable for the different groups of patients. The results also showed that there were no significant differences among the item difficulty measures between the two groups (high and low MMSE-CV group), and their correlations were high, which also indicated the stability of the ICAS item bank. Therefore, when applying the ICAS to different cognitive function groups, we could ensure that the item difficulty measure of each item in the ICAS did not differ.

Phase III: Investigation of the cutoff score for the ICAS and validation of the BP-ANN predictive model to predict the functional outcome for stroke survivors

The cutoff point of the ICAS was found and the prediction model driven by ANN was developed. The results revealed that the ICAS was a reliable and valid instrument to detect cognitive impairment in stroke survivors. The ICAS score was a linear continuous ratio scale that transformed 13 domains of cognitive function. With this scale, we can better monitor the progress of stroke patients and directly evaluate the effectiveness and outcome of the cognitive training programme. For example, if a patient's MMSE-CV score improves from 10 to 20, then it is hard to conclude that there is a corresponding 100% improvement in cognitive function because the distances between intervals are not equal. However, if on applying the ICAS a patient is found to improve from 2 to 4 logits, then we can conclude that there is a corresponding 100% improvement in cognitive function, as the distances from 2 to 3 and from 3 to 4 are equal on the ICAS scale. Therefore, the ICAS serves as a good clinical tool for monitoring the progress of the patients and at the same time could serve as an accurate outcome indicator. In addition, the ICAS covers such a wide range of cognitive functions that the resultant score can reflect cognitive function more accurately. The adaptive testing procedure of ICAS, which enhances each presented test item, helps us to obtain maximal information on the proficiency level of a patient. As a result, the testing time can be greatly shortened to minimize the chance that fatigue or loss of attention and motivation will affect the accuracy and interpretation of the results. The CAT feature may thus enhance the ICAS's possible usage in different neuro-cognitive patient groups. Finally, the ICAS is a computer-assisted assessment and it can readily be programmed to use another language (for instructions) and other culturally relevant content for wider application too.

For the fluency of ICAS administration in stroke survivors, in general, most of the stroke patients could respond according to instructions given by the ICAS. Few of them (5 out of 110 subjects) required verbal supervision that could be offered by a stand-by assistant and none of them required physical assistance. Patient with low literacy skills appear to benefit from individualized pace of instruction and the non-threatening learning that occur with a computer based program (Lewis, 1996). Therefore, when applying the ICAS to totally illiterate subjects, we need more verbal explanation and let them know the location of answer at this moment. Further development will improve to have audio prompting from the system automatically. Moreover, it was observed that stroke patients interacted well with the system through the touch monitor (rather than using a keyboard or mouse). Therefore, it is envisaged that this type of computer input equipment could be effectively used even by the elderly to interact with a computer during cognitive testing. The average time for them to complete the ICAS was 25 minutes with range from 10 minutes to 45 minutes due to different cognitive abilities level needed different number of test items. The study finding was match with the finding from Fredrickson et al in 2010 that computerized test was shown to have good acceptability, efficiency and stability for the repeated assessment of cognitive function in older people.

The existing computerized cognitive assessment such as CANTAB, CDR and

MindStream, their test items were more laboratory orientated and were just added up the raw score to form a total score. Therefore the ICAS got its advantages in providing score in linear scale, applying the CAT test procedures and the test items were more daily living orientated.

Future development of ICAS

Obviously, the study is not the end of ICAS development but the beginning. The ICAS now has only 65 items in the item bank and its coverage of 13 cognitive domains can be expanded such that more testing items can be included in each of the cognitive domains and these items were more in activities of daily living and leisure activities orientated items. Moreover, based on the advantages of the computer programme and adaptive testing method, ICAS could extend its usage to a variety of applications. It is most likely that ICAS can be adapted to assess cognitive functions of clients suffering from mild cognitive impairment (MCI), dementia, and traumatic brain injury, as ICAS has been equipped with a basic comprehensive cognitive testing item bank that is suitable for assessing these kinds of clients. In order to extend ICAS's usage, another study with different sampling populations is needed to investigate the cutoff score for specific populations. Moreover, the CAT procedures could shorten the time needed for test-taking and the computer could provide an assessment environment with different controlled multimedia simulation. These two features are especially suitable for assessing children with attention deficit who may have low tolerance to long-duration paper-pencil tests.

In addition to extending the possible usage of ICAS to different clients, another aspect is to develop the norm reference for the performance of specific disease in different age, gender, and education-level groups. Once the norm references are developed, they could provide information on the cognitive performance within that population for comparison in addition to self-comparison or comparison with another client. Besides, further study to explore the correlation between the ICAS with other more extensive neuropsychological battery, performance in instrumental activities of daily living and leisure activities were needed.

Furthermore, as the ICAS is Chinese culture-related, it could provide culture relevance when applied to other Chinese societies. In the same vein, it can also be modified for application in other societies by considering cultural relevancy, proper translation of text, changing the instruction language and the photos/images of the ICAS programme, and conducting another validation study. Then all this information could be stored inside the ICAS system and the user could choose the language version of the test at the beginning of the assessment. This helps to solve the common problems of cultural difference of most paper-pencil tests.

Implications of the study

The ICAS is a new cognitive assessment in CAT format applicable in the rehabilitation field. It assesses major cognitive functions in 13 different domains and transforms them into an overall linear score that represents the underlying cognitive ability of stroke survivors. From the test administration point of view, conducting the assessment is easy and does not require any special training in administration procedure and the interpretation of results. The administration procedures are wholly automatically controlled by the system, which can enhance the standardization of administration procedure and scoring accuracy by reducing human recording errors. The results are generated immediately after the test and there is no need for complex calculation or interpretation procedures. In addition, the CAT characteristic of the ICAS can shorten the testing period, which reduces the chance of patient fatigue.

Clinically, the results of this study have positive implications for therapists working with stroke survivors. The ICAS has the potential to be utilized in assessing stroke survivors' cognitive abilities in relevant tasks and familiar environments. They can benefit from a valid and reliable instrument that can generate linear estimation of cognitive ability so that comparisons can be made regarding their performance on cognitive test items. With the linear cognitive score from the ICAS, therapists could plan their treatment, monitor their progress, and conduct discharge planning according to the information provided by the ICAS.

In addition to clinical implications, the ICAS can also benefit cognitive research by serving as an outcome measure. The ICAS score is a linear, unidimensional ratio score reflecting the underlying cognitive trait of test-takers. Therefore, it can be used as an outcome measure tool for research.

As discussed before, the ICAS could be used cross-culturally, as it can be translated into other language versions and adapted for use in different cultural contexts through simple programming techniques. Moreover, the mode of delivery of the ICAS also has an impact on the rehabilitation field. Traditionally, paper-pencil testing requires face-to-face contact for administration of the test. As the ICAS is delivered by computer and can be accessed over the Internet, patients who have difficulty going to hospital or clinic for assessment could access the ICAS easily and may be assessed equally well.

Limitations of the study

The findings of the study should be understood in the context of having some limitations. The sample size of the study was small, totally we recruited 110 subjects but compared with the total admission of stroke in 2008, which were 26,176 cases (Hospital Authority, 2008), our study only contributed to 0.4% in the total stroke admission. However the demographic distribution (percentages of types of strokes, side of stroke affect, and age) in subjects recruited in the three phases were similar to those of the seven-year stroke study done by Roth and Lovell in 2003. In Phase III, the sample size was not large enough to enable the results to be generalized to the whole stroke population. The sample sizes of the two groups were unequal, but the power analysis of the area under the curve (AUC) showed that the study sample had 97% power to determine the AUC difference at 0.09 between the ICAS and MMSE-CV. Also, the results of the study may refer only to persons aged over 60; no information is available on younger stroke survivors. In future studies, a wider age range and more comparable sample sizes of the two groups should be considered. This can increase the generalization ability and predictive ability of the scores and make it possible to obtain norms for stroke survivors in different age groups, including the group of ever younger stroke patients.

Moreover, this study was not a multi-centre study even though we recruited stroke survivors from different rehabilitation phases and from different wards and day hospitals of a single centre. All these limited the generalization ability of the study results.

In this study, we did not capture the data on the educational level which influenced the cognitive performance after stroke (Tang, et al 2006) and according to the Hong Kong Census and Statistic department in 2006, 30.4% of elderly have no schooling and the rest got elementary level or above. Therefore in phase III study, we use the MMSE cutoff score for elementary educational level to indicate whether they had cognitive impairment or not, in order to prevent underestimate the subjects' cognitive function influenced by their educational level.

Moreover, as the first step to explore the application of ICAS in stroke subjects, we did not capture the post stroke depression situation, as depression will influence the motivation and also affect the performance, this kind of data will be capture in future studies and this factor affecting our study may not be so significant as an post interview with the therapists, they report that subjects recruited did not got significant poor motivation.

Although the BP-ANN network predicting the self-care outcome measure was significantly correlated with the observed one, the validation samples were also small in size. A larger-scale study to validate this BP-ANN model should be conducted before launching for clinical application. Moreover, in our prediction model, we used only cerebral infraction or haemorrhage and did not use a more standardized scale such as the National Institutes of Health Stroke Scale (NIHSS) and Bamford Classification which could better describe the severity of a stroke. As a result, further study should include these two scales in the BP-ANN model.

Conclusion

This study achieved the objective of developing and validating the ICAS in the Phase I study, the second objective that investigated the item difficulty measures and unidimensionality of the item bank of ICAS for stroke survivors in Phase II, and the third and fourth objectives in Phase III.

The Phase I study showed the content validity of the ICAS and preliminary information supporting the concurrent validity and prediction ability of ICAS. The Phase II study showed that the difficulty measures of the ICAS item bank fit the Rasch model, and the overall linear cognitive score correlated well with standardized cognitive assessments, namely, the MMSE-CV and NCSE-CV. The Phase III study showed that the cutoff point of the ICAS was sensitive and specific for screening the cognitive impairment of stroke survivors and built up the ANN model for predicting the functional outcome of stroke survivors.

The results of the study provide preliminary evidence of the internal scale validity, person response validity, and reliability of the ICAS to serve as a clinical tool to assess the cognitive function of stroke survivors. These initial findings affirm that the ICAS has the potential to provide a linear numerical estimation of stroke survivors' cognitive ability. Internal scale validity and stability of the ICAS testing items were demonstrated through the goodness of fit to the Rasch model. The cutoff point of the ICAS was sensitive and specific to screen out the cognitive impairment for stroke survivors, and the ANN prediction model successfully predicted the functional outcome of the stroke survivors.

In conclusion, the study showed that the ICAS can serve as a quick and valid cognitive assessment tool for stroke survivors in our daily clinical practice, and further validation may further improve its applicability to stroke and persons with other neuro-cognitive disabilities. The potential for developing ICAS in other language versions is envisaged in future studies.

Reference

Abreu, B., Duval, M., Gerber, D. & Wood, W. (1994). Occupational performance an the functional approach. In C.B. Royeen (Ed) *AOTA self-study series: Cognitive rehabilitation.* Rockville, MD, American Occupational Therapy Association.

Allard, G., Butler, J., Faust, D., Shea, M.T. (1995). Errors in hand scoring objective personality tests: The case of the Personality Diagnostic Questionnaire – Revised (PDQ-R). *Professional psychology, research and practice*, 26, 304-308.

Anastasi, A. (1982). *Psychological testing* (5th ed.). New York: Macmillan.

Anderson ND, Winocur G, Palmer H (2003). *Principles of cognitive rehabilitation. In Handbook of clinical neuropsychology*. Oxford : Oxford University Press.

Arnadottir, G. (1990). *The brain and behavior: Assessing cortical dysfunction through activities of daily living (ADL).* St. Louis: Mosby.

Baddeley, A.D. & Hitch, G.J. (1974). Working memory. In G.A. Bower (ed.), *Recent Advances in Learning and Motivation*, Vol. 8 (pp. 47–89). New York: Academic Press.

Baker, R. (1997). Classical test theory and item response theory in test analysis. Extracts from: *An investigation of the rash model in its application to foreign language proficiency testing*. SDS Supplies Limited, Preston, Lancashire.

Barba R., Martinez-Espinosa S., Rodriguez-Garcia E., Pondal M., Vivancos J. & Del Ser T. (2000). Post stroke dementia: Clinical features and risk factors. *Stroke*, 31, 1494-1501.

Baum, C.M. & Edwards, D.F. (1993). Cognitive performance in senile dementia of the Alzheimer's type: A kitchen task assessment. *American Journal of Occupational Therapy*, 47, 431-436.

Berg, E.A. (1948). A simple objective technique for measuring flexibility in thinking. *Journal of General Psychology*, 39, 15-22.

Bolt, R.A. (1981). Gaze-orchestrated dynamic windows. *Computer Graph*, 15(3), 109-119.

Bond, G.T. & Fox, M.C. (2007). *Applying the Rasch model: fundamental measurement in the human sciences.* Mahwah, N.J.:Lawrence Erlbaum.

Bradley, V. & Kapur, N. (2003). Neuropsychological assessment of memory disorders. In Halligan, P.W., Kischka, U & Marshall, J.C. (ed.) *Handbook of Clinical Neuropsychology.* Oxford University Press.

Bruce, M.A.G. (1994). Cognitive rehabilitation: Intelligence, insight and knowledge. In Royeen, C.B. (Ed.), *AOTA self-study series: cognitive rehabilitation*. Rockville, MD: American Occupational Therapy Association.

Burns, T. (1990). The cognitive performance test: A new tool for assessing Alzheimer's disease. *OT week,* December 27. Rockville, MD: American Occupational Therapy Assoication.

Butcher, J.N. (1987). The use of computers in psychological assessment: An overview of practices and issues. In J.N. Butch (Ed.), *Computerized psychological assessment: A practitioner's guide*. New York: Basic Books.

Butcher, J.N., Perry, J.N. & Atlis, M.M. (2000). Validity and Utility of computer-based test interpretation. *Psychological Assessment*, 12(1), 6-18.

Caty, G.D., Arnould, C., Stoquart, G.G., Thonnard, J.L., & Lejeune, T.M. (2008). ABILOCO: A Rasch built 13-item questionnaire to assess locomotion ability in stroke patients. *Archives of Physical Medicine and Rehabilitation*, 89, 284-290.

Chan, C.C.H., Lee, T.M.C., Fong, K.N.K., Lee,C. & Wong, V. (2002). Cognitive profile for Chinese patients with stroke. *Brain Injury,* 16(10), 873-884.

Chan, C.C.H., Lee, T.M.C., Wong, V., Fong, K., & Lee, C. (1999). Validation of Chinese version Neurobehavioral Cognitive Status Examination (NCSE). *Archives of Clinical Neuropsychology*, 14, 71.

Cheng, B. & Titterington, D.M. (1994). Neural Networks: A review from a statistical perspective. Statistical Science, 9(1), 2-30.

Chiu, H., Lee, H.C., Chung, W.S., & Kwong, P.K. (1994). Reliability and validity of the Cantonese version of Mini-Mental State Examination: A preliminary study. *Journal of Hong Kong College Psychiatrists*, 4, 425-28.

Chute, D.L. (2002). Neuropsychological Technologies in Rehabilitation. *Journal of Head Trauma Rehabilitation*, 17, 369-377.

Collie, A., Barby, D. & Maruff, P. (2001). Computerized cognitive assessment of athletes with sports related head injury. *British Journal of Sports Medicine,* 35, 297-302.

Coster, W., Ludlow, L. & Mancini, M. (1999). Using IRT variable maps to enrich understanding of rehabilitation data. *Journal of Outcome Measure*, 3(2), 123-133.

Courtenay, B.C. & Weidemann, C. (1985). The effects of a "don't know" response on palmore's facts on aging quizzes. *The Gerontologist*, 25(2), 177-181.

Cross, S.S., Harrison, R.F. & Sanders, D.S. (2003). Supporting decisions in clinical medicine: neural networks in lower gastrointestinal haemorrhage. The *Lancet,* 362(9392), 1250-1251.

Cullum, S., Huppert, F.A., McGee, M., Dening, T., Ahmed, A., Paykel, E.S. & Brayne, C. (2000). Decline across different domains of cognitive function in normal ageing: results of a longitudinal population-based study using CAMCOG. *International Journal of Geriatric Psychiatry*, 15(9), 853-862.

de Jager, C.A, Milwain, E. & Budge, M. (2002). Early detection of isolated memory deficits in the elderly: the need for more sensitive neuropsychological tests. *Psychological medicine*, 32, 483-491.

de Jager, C.A, Budge, M.M. & Clarke, R. (2003). Utility of TICS-M for the assessment of cognitive function in older adults. *International Journal of Geriatric Psychiatry,* 18(4), 318-324.

Desmond, D.W., Moroney,J.T., Sano,M & Stern,Y. (1996). Recovery of cognitive function after stroke. *Stroke*, 27(10), 1798-1803.

Donovan, N.J., Kendell, D.L., Heaton, S.C., Kwon, S., Velozo, C.A. & Duncan, P.W.

(2008). Conceptualizing functional cognition in stroke. *Neurorehabilitation and Neural Repair*, 22, 122-135.

Duchek, J.M. & Abreu, B.C. (1997). Meeting the challenges of cognitive disabilities. In Christiansen, C. & Baum, C. (Eds.), *Occupational therapy: Enabling function and well-being* (2nd ed.). Thorofare, NJ: SLACK Incorporated.

Elkins, J.S., O'Meara,E.S., Longstreth Jr., W.T., Carlson, M.C., Manolio,T.A. & Johnston, S.C. (2004). Stroke risk factors and loss of high cognitive function. *Neurology*, 63,793-799

Embretson, S.E., (1999). Issues in the measurement of cognitive abilities. In Embretson, S.E. & Hershberger, S.L. (Ed.). *The new rules of measurement: what every psychologist and educator should know*. Lawrence Erlbaum Associates.

Emond, B. & West, R.L. (2003). Cyberpsychology : A human interaction perspective based on cognitive modeling. *Cyberpsychology and behavior,* 6(5), 527-536.

Epstein, J. & Rotunda, R.J. (2000). The utility of computer versus clinician-authored assessments in aiding the prediction of patient symptomatology. *Computers in Human Behavior,* 16, 519-536.

Epstein, J., Klinkenberg, W.D., Wiley, D. & McKinley, L. (2001). Insuring sample equivalence across internet and paper-and-pencil assessments. *Computers in Human Behavior,* 17, 339-346.

Evans, J.J. (2003). Basic concepts and principles of neuropsychological assessment. In Halligan, P.W., Kischka, U & Marshall, J.C. (ed.) *Handbook of Clinical Neuropsychology*. Oxford University Press.

Finlay, J. & Dix, A. (1996). *An introduction to artificial intelligence*. UCL Press.
Fisher, A.G. (1992). *Assessment of motor and processing skills*. Unpublished test manual. Fort Collins, CO: Colorado State University.

Fisher, G.H., & Molenaar, I.W. (1995). Rasch models: Foundations, recent development and applications. Berlin, Germany: Springer-Verlag.

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

Fong, K., Ng, B., Chan, D., Chan, E., Ma, D., Au, B., Chiu, V., Chang, A., Wan, K., Chan, A, & Chan, V. (2004). Development of the Hong Kong version of the functiuonal test for the hemiplegic upper extremity (FTHUE-HK). *Hong Kong journal of occupational therapy*, 14, 21-29.

Franzen, M.D. (2000). *Reliability and validity in neuropsychological assessment* (2nd ed.). KluwerAcademic/Plenum Publishers.

Franzen,M.D. & Wilhelm, K.L. (1996). Conceptual foundations of ecological validity in neuropsychological assessment. In Sbordone, R.J. & Long, C.J. (eds.) *Ecological validity of neuropsychological testing*. GR Press/ St. Lucie Press.

Fredrickson, J., Maruff, P., Woodward, M., Moore, L., Fredrickson, A., Sach, J. & Darby, D. (2010). Evaluation of the usability of a brief computerized cognitive screening test in older people for epidemiological studies. *Neuroepidemiology*, 34, 65-75

Goldstein, F.C., & Levin, H.S. (1987). Neuropsychological assessment for rehabilitation: Fixed batteries, automated systems and non-psychometric methods. In M.J. Meir, A.L. Benton & L. Diller L. (Eds). *Neurological Rehabilitation (pp.18-40).* New York: Guilford Press.

Gourlay, D., Lun, K.C., Lee, Y.N., Tay, J. (2000). Virtual reality for relearning daily living skills. *International Journal of Medical Informatics,* 60, 255-261.

Graf, P. & Schacter, D. L. (1985). Implicit and explicit memory for new associations in normal and amnesic subjects. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 11, 501-518.

Gray, W.D. & Altmann, E.M. (2001) Cognitive modeling and human computer interaction. In Karwowski, W. (Ed.), *International encyclopedia of ergonomics and human factors*, Vol. 1, 387-391.

Grieve, J. & Gnanasekaran, L. (2008). *Neuropsychology for occupational therapists: Cognition in occupational performance*. Blackwell Publishing.

Gronwall, D. & Sampson, H. (1974). *The psychological effects of concussion*. Auckland University Press, Auckland.

Grudin, J. (1990). The computer reaches out: the historical continuity of interface design. *Conference on Human Factors in Computing Systems, Proceedings of the SIGCHI conference on Human factors in computing systems: Empowering people*, 261-268.

Gulliksen, H. (1987). *Theory of mental tests*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Hahn, A.E., Cella, D., Bode, K.R., Gershon, R., Lai, J.S. (2006). Item banks and their potential applications to health status assessment in diverse populations. *Medical Care*, 44(11), S189-S197.

Halkitis, P.N. (1993) Computer-adaptive testing algorithm. *Rasch Measurement Transactions*,6(4), 254-5.

Hambleton, R.K. & Cook, L. L.(1977). Latent trait models and their use in the analysis of educational test data. *Journal of educational measurement*, 14(2), 75-96.

Hambleton, R.K., Swaminathan, H., Cook, L.L., Eignor, D.R. & Giffora, J.A. (1978). Developments in latent trait theory: models, technical issues, and applications. *Review of educational research*, 48(4), 467-510.

Hancox, P.J., Mills, W.J. & Reid, B.J. (1990). *Key guide to information sources in artificial intelligence / expert systems*. Mansell, London.

Handel, R.W., Ben-Porath, Y.S. & Watt, M. (1999). Computerized adaptive assessment with the MMPI-2 in a clinical setting. *Psychological Assessment*, 11, 369-380.

Hart, T., & Hayden, M. E. (1986). The ecological validity of neuropsychological assessment and remediation. In Uzzell, B. & Gross, Y. (Eds.), *Clinical neuropsychology.* New York: Martinys Nijhoff.

Hartingsveld, F.V., Lucas, C., Kwakkel, G., & Lindeboom, R. (2006). Improved interpretation of stroke trial results using empirical Barthel item weights. *Stroke*, 37, 162-166.

Heruti, R.J., Lusky, A., Dankner, R., Ring, H., Dolgopiat, M., Barell, V., Levenkrohn, S. & Adunsky, A. (2002). Rehabilitation Outcome of elderly patients after a first stroke: effect of cognitive status at admission on functional outcome. *Archives of Physical Medicine and Rehabilitation*, 83, 742-749.

Hobart, J. (2002). Measuring disease impact in disabling neurological conditions: Are patient's perspectives and scientific rigor compatible? *Current Opinion in Neurology*, 15, 721-724.

Hofer, P.J. (1985). Developing standards for computerized psychological testing. *Computers in Human Behavior,*1, 301-315.

Hoffmann, M. (2001). Higher cortical function deficits after stroke: An analysis of 1,000 patients from a dedicated cognitive stroke registry. *Neurorehabilitation and Neural Repair*, 15, 113-127.

Hol, A.M., Vorst, H.C.M. & Mellenbergh, G. (2008). Computerized adaptive testing of personality traits. *Journal of Psychology*, 216(1), 12-21.

Hollen, J., Hutchins, E. & Kirsh, D. (2000). Distributed Cognition: Toward a new foundation for human computer interaction research. *ACM Transactions on Computer Human Interaction,* 7(2), 174-196.

Hospital Authority. (2008). *Hospital authority statistical report 2007-2008*. Hospital Authority.

Hulin, C.L., Drasgow, F. & Parsions,C.K., (1983). *Item response theory*. Homewood, Illinosi: Dow-Jones Irwin.

Jacob, R.J.K. (1991). The use of eye movements in human computer interaction techniques: What you look at is what you get. *ACM transactions on information and system security,* 9(3), 152-169.

Jacob, R.J.K. (1996). Human computer interaction: Input devices. *ACM Computing Surveys*, 28(1), 177-179.

Kalra, L. & Crome, P. (1993). The role of prognostic scores in targeting stroke rehabilitation in elderly patient. *Journal of American Geriatric Society*, 41, 396-400.

Katz, N., Itzkovich, M., Averbuch, S. & Elazar, B. (1989). Loewenstein occupational therapy cognitive assessment (LOTCA) battery for brain injured patients: Reliability and validity. *American Journal of Occupational Therapy*, 43, 184-192.

Khateb, A. Annoni, J.M., Lopez, U., Bernasconi, F. Lavanchy, L. & Bogousslavsky, J. (2007). Evaluation of cognitive and behavioral disorders in stroke unit. In Godefroy, O & Bogousslavsky, J. (Eds.) The Behavioral and cognitive neurology of stroke (pp 1-14). Cambridge University Press.

Kay, A. (2001). Artifical neural networks. *Computerworld*, Feb (12),35(7), 60.

Laver, A.J. & Powell, G.E. (1995). *The Structured Observational Test of Function (SOTOF)*. Windsor, England: NFER Nelson.

Lee, J.H., Ku, J., Cho, W., Hahn, W.Y., Kim, I.Y., Lee, S.M., Kang, Y., Deog, Y.K., Yu, T., Wiederhold, B.K., Wiederhold, M.D. & Kim, S.I. (2003). A virtual reality system for the assessment and rehabilitation of the activities of daily living. *Cyberpsychology and behavior,* 6(4), 383-388.

Leirer, V.O., Morrow, D.G., Pariante, G.M., Sheikh, J.I. (1988). Elders' non adherence, its assessment, and computer-assisted instruction for medication recall training. *Journal of American Geriatric Society*, 36(10), 877–884.

Lesniak, M., Bak, T., Czepiel, W., Seniow, J. & Czlonkowska, A. (2008). Frequency and prognostic value of cognivitive disorders in stroke patients. *Dementia and Geriatric Cognitive Disorders*, 26, 356-363.

Lewis, D. (1996). Compute-based approaches to patient education: A review of the literature. *Journal of the American Medical Informatics Association*, 6, 272-282.

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

Linacre, J.M. (1994). Sample size and item calibration stability. *Rasch Measurement*

Transactions, 7, 328.

Linacre, J.M. (2006). *A user's guide to Winsteps, Ministep, Rasch-model computer programs*. Retrieved May 6, 2009, from WINSTEPS & Facets Rasch Software Web site: http://www.winsteps.com

Linden, T., Skoog, I., Fagerberg, B., Steen, B. & Blomstrand, C. (2004). Cognitive impairment and dementia 20 months after stroke. *Neuro-epidemiology*, 23, 45-52.

Lippmann, R.P. (1987). An introduction to computing with neural nets. *IEEE ASSP Magazine*, 4(2), 4-22.

Lord, F.M. (1952). *A theory of test scores*. Psychometric Monographs, 7, 1-84.

Lord, F. M. & Novick, M. R. (1968). *Statistical Theories and Mental Test Scores*. Reading, Mass.: Addison-Wesley.

Lowery, K., Ballard, C., Rodgers, H., McLaren, A., O'Brien, J., Rowan, E. & Stephens, S. (2002). Cognitive decline in a prospectively studied group of stroke survivors, with a particular emphasis on the >75s. *Age and Ageing*, 31(Suppl. 1), 24-27.

Lucila, O.M., & Todd, R., (1999). Neural network applications in physical medicine and rehabilitation. *American Journal of Physical Medicine & Rehabilitation,* 78(4), 392-398.

Luk, J.K., Chiu, P.K. & Chu, L.W. (2008). Rehabilitation of older Chinese patients with different cognitive functions: how do they differ in outcome? *Archives of Physical Medicine and Rehabilitation*, 89, 1714-1719

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

Manly, M. & Robertson, I.H. (1997). Sustained attention and the frontal lobes. In Rabbit, P. (Ed.), *Methodology of Frontal and Executive function*. Hove, U.K.: Psychology Press.

Mayer, R.E. (1992). *The Thinking, Problem Solving, Cognition*. New York: Freeman.

McGraw, K.O. & Wong, S.P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, 1, 30-46.

Miller, R.B. (1968). Response time in man –computer conversation transactions, in *Proceedings of Fall Joint Computer Conference*, 267-277.

Morris, R.G., Evenden, J.L., Sahakian, B.J. & Robbins, T.W. (1987). Computer-aided assessment of dementia: comparative studies of neuropsychological deficits in Alzheimer-type dementia and Parkinson's disease. *Cognitive Neurochemistry*, 21-36.

Myers, B., Hollen, J. & Cruz, I., Bryson, S., Bulterman, D., Catarci, T., Citrin, W., Glinert, E., Grudin, J. & Ioannidis, Y. (1996). Strategic directions in human computer interaction. *ACM Computing Surveys*, 28(4), 794-809.

Neisser, U. (1967). *Cognitive Psychology*. New York : Appleton-Century-Crofts.

Neistadt, M.E. (1992). The Rabideau Kitchen Evaluation Revised: An assessment of meal preparation skills. *Occupational Therapy Journal of Research*, 12, 242-255.

Neistadt, M.E. (1994). A meal preparation treatment protocol for adults with brain injury. *American Journal of Occupational Therapy,* 48, 431-438.

Nilsson. N.J. (1998). *Artificial intelligence: A new synthesis*. Morgan Kanfmann.

Northern California Neurobehavioral Group (1995). *Cognitstat: The Neurobehavioural Cognitive Status Examination in an older psychiatric population*. Fairfax, CA: Northern California Neurobehavioral Group.

Northern California Neurobehavior Group. (1995). *Manual for Cognistat: The Neurobehavioral Cognitive Status Examination*. Fairfax, CA: The Norhtern California Neurobehavior Group.

O'Brien, J.T.O., Erkinjuntti, T., Reisberg, B., Roman, G., Sawada, T., Pantoni, L., Bowler, J.V., Ballard, C., DeCarli, C., Gorelick, B.P., Rockwood, K., Burns, A., Gauthier, S., & DeKosky, S.T. (2003). Vascular cognitive impairment. *Lancet Neurology*, 2, 89-98.

Ogozalek, .VZ. (1993). The “automated pharmacist”: comparing the use of leaflets, text-based computers, and multimedia computers to provide medication information to the elderly. *Journal of Medical Education Technology*, 6–11.

Olson-Buchanan, J.B., Drasgow, F., Moberg, P.J., Mead, A.D., Keenan, P.A. & Donovan, M. (1998). Interactive video assessment of conflict resolution skills. *Personnel Psychology,* 51, 1-24.

Pallant, J.F., & Tennant, A. (2007). An introduction to the Rasch measurement model: An example using the Hospital Anxiety and Depression Scale (HADS). *British Journal of Clinical Psychology*, 46, 1-18.

Pasquier, F. (1999). Early diagnosis of dementia: neuropsychology. *Journal of neurology*, 246(1), 6-15.

Paolucci, S., Antonucci, G., Gialloretti, L., Traballesi, M., Lubich, S., Pratesi, L. & Palombi, L. (1996). Predicting stroke inpatient rehabilitation outcome: the predominant role of neuropsychological disorders. *European Neurology,* 36, 385-390.

Patel,M., Coshall, C., Rudd, A.G. & Wolfe, C.D. (2003). Natural history of cognitive impairment after stroke and factors associated with its recovery. *Clinical rehabilitation*, 17, 158-166.

Pedersen, P.M., Jorgensen, H.S., Nakayama, H., Raaschou, H.O. & Olsen, T.S. (1996). Orientation in the acute and chronic stroke patient: impact on ADL and social activities. The Copenhagen Stroke Study. *Archives of Physical Medicine and Rehabilitation*, 77(4), 336-339.

Perez-Quinones, M.A. & Sibert, J.L. (1996). A collaborative model of feedback in human computer interaction. *Computer Human Interaction,* 4, 13-18.

Ponser, M.I. & Boies, S.J. (1971). Components of attention. *Psychological review,* 78(5), 391-408.

Portney, L.G. & Waltkins, M.P. (1993). *Foundations of clinical research application to practice.* Appleton & Lange.

Preece, J., Rogers, Y. & Sharp, H. (2002). *Interaction design: beyond human* *computer interaction,* NewYork.

Prigatano, G.P. & Fordyce, D.J. (1986). Cognitive dysfunction and psychological adjustment after brain injury. In G.P. Prigatano (Ed.), *Neuropsychological rehabilitation after brain injury* (pp1-17). Baltimore: JohnHopkins University Press.

Radomski, M.V. (2008). Assessing abilities and Capacities: Cognition. In Trombly, C.A. & Radomski M.V. (Eds.). *Occupational therapy for physical dysfunction (6th ed.*)(pp 260-283). Lippincott Williams & Wilkins.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Chicago, IL: University of Chicago Press.

Richman-Hirsch, W.L., Kiesler, S., Weisband, S. & Drasgow, F. (1999). A meta-analytic study of social desirability response distortion in computer-administrated questionnaires, traditional questionnaires and interviews. *Journal of Applied Psychology*, 84, 754-775.

Richman-Hirsch, W.L., Olson-Buchanan, J.B. & Dragow, F. (2000). Examining the impact of administration medium on examinee perceptions and attitudes. *Journal of Applied Phycology*, 85, 880-887.

Rippey, .RM., Bill, D., Abeles, M., et al. (1987). Computer-based patient education for older persons with osteoarthritis. *Arthritis Rheumatology*, 30(8), 932–935.

Ritter, F.E., Bazter, G.D. & Jones, G & Young, R.M. (2000). Supporting cognitive models as users. *ACM Transactions on computer human interaction,* 7, 141-173.

Rizzo, A.A. & Buckwalter, J.G. (1997). Virtual reality and cognitive assessment and rehabilitation: The state of the art. In Riva, G. (Ed.) *Virtual reality in neuro-psycho-physiology*. Amsterdam: Ios Press.

Robertson, I.H., Ridgeway, V., Greenfield, E. & Parr, A. (1997). Motor recovery after stroke depends on intact sustained attention: a 2 year follow up study. *Neuropsychology*, 11(2), 290-295.

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

Rosamond, W., Flegal, K., Friday, G., Furie, K., Go, A., Greenlund, K., Haase, N., Ho, M., Howard, V., Kissela, B., Kittner, S., Lloyd-Jones, D., McDermott, M., Meigs, J., Moy, C., Nichol, G., O'Donnell, C.J., Roger, V., Rumsfeld, J., Sorlie, P., Steinberger, J., Thom, T., Wasserthiel-Smoller, S. & Hong, Y. (2007). Heart disease and stroke statistics—2007 updates: A report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. *Circulation*, 115, 69-171.

Sachdev, P.S., Brodaty, H., Valenzuela, M.J., Lorentz, L., Looi, J.C.L., Wen, W. & Zagami, A.S. (2004). The neuropsychological profile of vascular cognitive impairment in stroke and TIA patients. *Neurology*, 62, 912-919.

Schatz, P. & Browndyke, J. (2002). Applications of computer-based neuropsychological assessment. *Journal of Head Trauma Rehabilitation,* 17(5), 395-410.

Schmandt, C. (1993). From desktop audio to mobile access: Opportunities for voice in computing. In Hartson, H.R. & Hix, D. (Eds.). *Advances in Human Computer Interaction,* Ablex Publishing Corporation, Northwood, NJ. 251-283.

Schulenberg, S.E. & Yutrzenka, B.A. (1999). The equivalence of computerized and paper-and-pencil psychological instruments: implications for measures of negative affect. *Behavior Research Methods Instruments & Computers,* 31, 315-321.

Shah, S., Vanclay, F. & Cooper, B. (1989). Improving the sensitivity of the Barthel Index for stroke rehabilitation. *Journal of Clinical Epidemiology*, 42(8), 703-709.

Shah, S. & Copper, B. (1992). The Barthel index and ADL evaluation in stroke rehabilitation in Australia, Japan, the UK and the USA. *Australian Occupational Therapy Journal*, 39, 5-13.

Shallice, T. (1988). *From neuropsychology to mental structure*. Cambridge University Press, Cambridge.

Shneiderman, B. (1987). *Designing the user interface: Strategies for effective human computer interaction*. Addison-Wesley Publishing Co., Reading, Masachusettts.

Simes, D.K. & Sirsky, P.A. (1988). Human factors: An exploration of the psychology of human computer dialogues, In Hartson, H.R. & Hix., D. (Eds.). *Advance in Human-Computer Interaction,* Ablex Publishing Corporation, Northwood, New Jersey.

Simon, E.S., Doniger, G.M., Dimant, L., & Dwolatzky, T. (2007). Validity of a 15 minute multi-domain computerized cognitive battery for moderate to severe impairment. *Alzheimer's and Dementia*, 3, S137-S137.

Simpson, P., Surmon, D., Wesnes, K. & Wilcock, G. (1991). The cognitive drug research computerized assessment system for demented patients: a validation study. *International Journal of Geriatric Psychiatry*, 6, 95–102.

Sohlberg, M.M. & Mateer, C.A. (1989). *Introduction to cognitive rehabilitation theory and practice*. New York: Guilford.

Srikanth, V.K., Thrift, A.G., Saling, M.M., Anderson, J.F.I., Dewey, H.M., Macdonell, R.A.L. & Donnan, G.A. (2003). Increase risk of cognitive impairment 3 months after mild to moderate first ever stroke: A community based prospective study of nonaphasic English speaking survivors. *Stroke*, 34, 1136-1143.

Stephens, S., Kenny, R.A., Rowan, E., Allan, L., Kalaria, R.N., Bradbury, M. & Ballard, C.G. (2004). Neuropsychological characteristics of mild vascular cognitive impairment and dementia after stroke. *Internal Journal of Geriatric Psychiatry*, 19, 1053-1057.

Strub, R.L. & Black, F.W. (1977). *The mental state examination in neurology*. Philadelphia: F.A.Davis.

Tang, W.K., Chan, S.S.M., Chiu, H.F.K., Ungvari, G.S., Wong, K.S., Kwok, T.C.Y., Mok, V.,Wong, K.T., Richards, P.S. & Ahuja, A.T. (2004). Frequency and determinants of poststroke dementia in Chinese. *Stroke*, 35,930-935.

Tang, W.K., Chan, S.S.M., Chiu, H.F.K., Ungvari, G.S., Wong, K.S., Kwok, T.C.Y., Mok, V.,Wong, K.T., Richards, P.S. & Ahuja, A.T. (2006). Frequency and clinical determinants of post stroke cognitive impairment in nondemented stroke patients. *Journal of Geriatric Psychiatry and Neurology*, 19, 65-71.

Tatemichi,T.K., Desmond,D.W., Stern,Y., Sano,M. & Bagiella,E. (1994). Cognitive impairment after stroke: frequency, patterns, and relationship to functional abilities. *Journal of Neurology, Neurosurgery and Psychiatry*, 57,202-207.

Thorndike, R. L. (1982). *Applied Psychometrics*. Bonston: Houghton Mifflin.

Thorndike, R.M. (1999). IRT and intelligence testing: Past, present and future. In Embretson, S. E. & Hershberger, S. L. (Eds.) *The new rules of measurement: What every psychologist and educator should know*. Mahwah, NJ: Lawrence Erlbaum Associates.

Traub, R.E. & Wolfe, R.G. (1981). Latent trait theories and the assessment of educational achievement. *Review of research in Education,* 9, 377-435.

Tyerman, R., Tyerman, A., Howard, P. & Hadfield, C. (1986). *The Chessington Occupational Therapy Neurological Assessment Battery introductory manual*. Nottingham: Nottingham Rehab.

Van der Linden, J.W. & Glas, A.W.C. (2000). *Computerized adaptive testing theory and practice*. London: Kluwer Academic Publishers.

Wainer, H. (Ed). (2000). *Computerized adaptive testing: A primer (2nd Ed.).* Mahwah, NJ: Lawrence Erlbaum Associates.

Wasserman, P.D. (1989). *Neural computing: Theory and Practice*. VanNostrand Reinhold: New York, NY.

Webb, P.M., Zimet, G.D., Fortenberry, J.D., Blythe, M.J. (1999). Comparability of a computer-assisted versus written method for collecting health behavior information from adolescent patients. *Journal of Adolescent Health,* 24, 383-388.

Wellman, B., Haase, Q.A., Witte, J. & Hampton, K. (2001). Does the Internet Increase, Decrease, or Supplement Social Capital? *American Behavioral Scientist*, 45(3), 436-455.

Wenzel, G.R. (2002). Creating an interactive interdisciplinary electronic assessment. *Computers, Informatics, Nursing,* 20, 251-260.

Wheatly, C.J. (2001). Evaluation and treatment of cognitive dysfunction. In. Pedretti, L.W. & Early, M.B. (Eds.). *Occupational therapy: Practice skills for physical dysfunction*. St. Louis: Mosby Elsevier.

Willmes, K. (2003). The methodological and statistical foundations of neuropsychological assessment. In Halligan, P.W., Kischka, U & Marshall, J.C. (eds.) *Handbook of Clinical Neuropsychology*. Oxford University Press.

Wilson, B.A., Alderman, N., Burgess, P.W., Emslie, H. & Evans, J.J. (1996). *BADS – Behavioral Assessment of the Dysexecutives Syndrome*. Thames Valley Test Company, Bury, St. Edmunds.

Wilson, D.J., Baker, L.L. & Craddock, J.A. (1984). *Functional test for the hemiparetic upper extremity: test manual.* CA: Los Amigos Research and Education Institute, Inc.

Wright BD, Douglas GA. (1996). Estimating measures with known item difficulties. *Rasch Measurement Transactions*, 10(2), 499.

Wright, B.D, & Stone, M.H. (1999).*Measurement essentials*. Wilmington, Delaware: Wide Range, Inc.

Wright. B.D., & Linacre, M.J. (1987). Rasch model derived from objectivity. *Rasch Measurement Transactions*, 1, 15.

Yamanaka, H., Kawahira, K., Arima, M., Shimodozono, M., Etoh, S., Tanaka, N. & Tsujio, S. (2005). Evaluation of skilled arm movements in patients with stroke using a computerized motor-skill analyser for the arm. *International journal of rehabilitation research,* 28, 277-283.

Yip, C.K. & Man, D.W. (2009). Validation of a computerized cognitive assessment system for persons with stroke: a pilot study. *International journal of rehabilitation research*, 32(3), 270-278.

Zandvoort, M.J.E., Kessels, R.P.C., Nys, G.M.S., Haan, E.H.F. & Kappelle, L.J. (2005). Early neuropsychological evaluation in patients with ischaemic stroke provides valid information. *Clinical Neurology and Neurosurgery*, 107, 385-392.

Zhang, L., Abreu, B.C., Masel, B., Scheibel, R.S., Christiansen, C.H., Huddleston, N., Ottenbacher, K.J. (2001). Virtual reality in the assessment of selected cognitive function after brain injury. *American Journal of Physical Medical Rehabilitation,* 80(8), 597-604.

Zomeren, E.V. & Spikman, J. (2003). Assessment of attention. In Halligan, P.W., Kischka, U & Marshall, J.C. (ed.) *Handbook of Clinical Neuropsychology*. Oxford University Press.

LIST OF APPENDICES

Appendices

Appendix I -- Definition of cognitive components

Cognition can be classified into three levels: primary cognitive capacity, higher level thinking ability and meta processes ability or executive function (Radomski, 2008). The definitions of each level were as following:

Primary cognitive capacity

Orientation is an individual's ongoing awareness of their current situation, their environment and the passage of time. It is relates to an individual's memory capacity, as an individual must be able to remember past occurrences in order to orientate to time, place and person and hence orientation is included as a subset of most standardized assessments of cognitive function. Orientation can be assessed by simply asking the patient what year, month, day and time is it; where the patient is; and who their relatives or attending staffs are.

Attention is the ability to focus on a particular stimulus and to maintain that stimulus in mind, sometimes over an extended period of time before further processing can take place. Hence, it is the first critical step of more complex cognitive skills, especially memory (Duchek & Abreu, 1997; Evans 2003). Specifically, attention can be further classified into sustained attention; focused attention; divided attention and alternating attention. Sustained attention is the capacity to maintain attention performance over time (Lezak, Howieson, Loring, Hannay & Fischer 2004).

Focus attention is the ability of an individual to focus on relevant stimuli whilst irrelevant stimuli are present and is acting as a distracter. (Lezak et al, 2004), and is also known as selective attention. While divided attention refers to in tasks situation, subtasks can be distinguished and more than one type of response is required (Zomeren & Spikman, 2003). Alternating attention is the flexibility of an individual to make necessary shifts of attention between tasks of differing cognitive requirements. (Lezak et al, 2004).

Memory is the dynamic continuation of the attention process which includes the factor of time. It requires an individual to maintain focus on a task for information to be stored (Wheatly, 1996). Memory is complex and multifaceted, it included short term (STM) and long term memory (LTM); episodic and semantic memory; retrospective and prospective memory; explicit and implicit memory; declarative and procedural memory. Memory theorists have provided distinctions among various types of memory. In the following, we will briefly discuss about different type of memories.

STM is a temporary information store that holds a limited amount of information for a short period of time (Bradley & Kapur, 2003). It is also referred to as working memory (Baddeley & Hitch, 1974), as it serves the important function of holding information in consciousness for further processing (Duchek & Abreu, 1997).

LTM is a more permanent memory store of information. It has been argued that LTM has an infinite storage capacity and that information stored in LTM is never lost. However, it is believed that information stored in LTM has the potential to become inaccessible. (Duchek & Abreu, 1997). In order to specify the nature of LTM, episodic and semantic memory; retrospective and prospective memory; explicit and implicit memory; declarative and procedural memory were purposed to describe the LTM.

Semantic memory refers to organized body knowledge about words and concepts and culturally and educationally acquired facts. It includes general knowledge and covers a wide range of materials and modalities (Bradley & Kapur, 2003). It could be describe as the memory of general knowledge. Semantic memory is organized as a complex network and retrieval can be guided by several dimensions including meaning, association or rules (Duchek & Abreu, 1997).

Episodic memory refers to the encoding, store and utilization of memory for personally experienced events that can be related to specific spatial and temporal contexts (Bradley & Kapur, 2003). Retrieval of episodic memory is guided primarily by the temporal or contextual tags stored with the relevant information (Duchek & Abreu, 1997).

Implicit memory is memory that is expressed through behavioral or physiological changes, where the individual has no or limited conscious awareness of the information that has been stored. It refers to the facilitation of performance due to previous experience with the task, regardless if there is any conscious recollection of the task (Graf & Schacter, 1985). It usually includes tasks such as priming, conditioning skill learning (Bradley & Kapur, 2003).

Explicit memory is memory that is consciously accessed and covers most standard memory tasks (Bradley & Kapur, 2003). It refers to memory which involves the deliberate and conscious retrieval of information, which is typical of recall or recognition tests (Graf & Schacter, 1985).

Retrospective memory is the recollection of information or events that have occurred in the past (Duchek & Abreu, 1997).

Prospective memory refers to remembering to carry out some action in the future, “remembering to remember” (Duchek & Abreu, 1997). For example, a patient needs to remember to take pills in the afternoon.

Declarative memory represents our knowledge for factual information, this type of knowledge appears to be represented as a series of related statements and can be easily described verbally (Duchek & Abreu, 1997). It could be included the memory of facts, concepts and principles.

Procedural memory is the knowledge required to carry out the necessary procedures to perform a certain activities. For example, the knowledge required for riding a bike. This type of knowledge appears to be represented in memory of a set of procedures and cannot be easily described verbally (Duchek & Abreu, 1997). This includes the memory required to carry out actions, make decisions and execute procedures.

The interaction that occurs between declarative and procedural memory enhances the acquisition of new skills.

Higher Level Thinking Components

Reasoning is the drawing of inferences or conclusions from known or assumed facts. The reasoning process involves sequencing, classification, deduction and induction (Sohlberg & Mateer, 1989). Sequencing is the process of organizing information into its correct order. Classification is the grouping of information according to specific characteristics (Bruce, 1994). Deductive reasoning is the use of general information to identify specific facts and principles (Bruce, 1994), whereas induction reasoning is the generation of general rules from a given set of information or examples (Mayer, 1992).

Concept formation is the ability to analyze relationships between objects and their properties (Sohlberg & Mateer, 1989). This requires firstly the ability to recognize critical features of objects and their properties and secondly the ability to determine how these features is interrelated.

Problem solving is processing the necessary information related to facts and procedures contributes to one's ability in daily life At a basic level, all human responses that are not routine or habitual can be construed as problem solving (Radomski, 2008). As a general rule, problem solving involves identifying the problem, generating possible solutions, implementing the chosen solution and finally evaluating the outcome of the process.

Executive function

Executive functions refer to the ability to plan and problem solve, self monitor and regulate behavior (Evans, 2003). They can be divided into four components: volition, planning, purposeful action and effective performance (Lezak et al, 2004).

Volition is the recognition of a need or want and the subsequent formulation of a goal or an intention to act (Leazk et al, 2004). At a basic level it can be described as the motive for acting or the ability of initiation.

Planning is the identification and sequencing of steps to progress towards a goal or end point (Leazk et al, 2004).

Purposeful action is the translation of an intention or plan into productive, self-serving activity.

Effective performance is the ability to monitor and self correct while regulating the intensity, speed and strategies used during the task (Leazk et al, 2004).

To conclude, executive functions are a meta-process that provides the basis for initiating, continuing and completing a task at an appropriate time or stage. They act as a coordinator between primary cognitive capacities and higher level thinking when completing complex tasks (Ramdoski 2002).

Appendix II -- Supplement to Item Response Theory (IRT)

The essential feature of an item response theory (IRT) approach is that a relationship is specified between observable performance on test items and the unobservable characteristics or abilities assumed to underlie this performance (Hambleton, Swaminathan, Cook, Eignor & Giffora, 1978). The characteristic measured by a given set of items is conceived of as an underlying continuum, often referred to as a latent trait or latent variable. Although the trait is usually viewed as being continuously distributed, no specific form of distribution (such as a normal distribution) needs to be assumed (Hulin, et al, 1983). This underlying continuum is represented by a numerical scale, upon which a person's standing can be estimated using his/her responses to suitable test items (Hulin, et al, 1983). Items measuring the trait are seen as being located on the same scale, according to the trait level they require of test-takers (Baker, 1997).

The IRT model shows latent variables relates to behavior. The observed behavior is individual item response. The latent variable influences the probabilities of the response to the items. The probability that a person will pass or endorse a particular item depends on their trait level on the item difficult as following:

PROB(Item Passed) = Function [(trait level) – (Item Difficult)]

The main assumptions made in IRT are those relating to the form of the item characteristic curve (ICC), test uni dimensionality and the local statistical independence (Baker, 1997). The ICC states the relationship between the probability of a correct answer to item and trait level. Various response models differ in respect of the ICCs with which they operate. However, the particular ICC adopted is assumed to provide a plausible representation of the relationship between performance on test items and ability. The assumption is thus made that the form of ICC is correct for the data set in question (Baker, 1997).The second assumption is that the item set is uni dimensional, that is the items measure a single ability or trait (Hambleton & Cook, 1977). The third assumption is that of local independence. The principle of local independence states that for persons located at any given point on the ability scale, the probability of a person answering any one item correctly is not affected by information regarding that person's success or failure on any other items (Lord, 1952, Thorndike, 1982). In general terms, all information concerning the probability of a correct or incorrect response is contained in the ability parameter, and that if this parameter is known, then observing a person's responses to one or more of the items in a test provides no additional information about his/her responses to any other (Baker, 1997 Hulin et al 1983).

To check the data fulfill the above model assumption, the evaluation of fit and misfit of person and items from the IRT programme should be performed, otherwise, Traub & Wolfe in 1981 warn that if the model is inherently and grossly wrong for the data, then applications of the analytic results, which most programmes will produce regardless of fit, can be nonsensical. Thus, the goodness of fit between model and data must be investigated before any use of the statistics obtained from an IRT based analysis of test data can be contemplated (Baker, 1997).

Advantage of IRT

There were several advantages of IRT and presented as following:

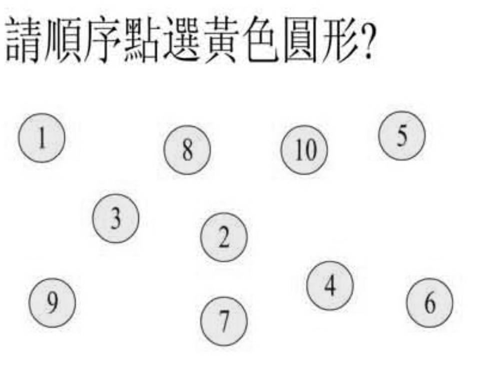
1. Person trait level estimates are controlled for the properties of the item that were administered. Item difficulty estimates are controlled for the trait levels of the particular persons in the calibration sample. In this sense, item free person estimates and population free item estimates are obtained.
2. Determining the person’s trait is not a equation of how to add up the item responses. In a sense, the IRT process of estimating trait levels is analogous to the clinical inference process. Latent variables must be inferred from presenting behavior.
3. The standard error (SE) of measurement in IRT is specific to each trait level, minimizing these errors leads to greater reliability for a group as a whole. This may be readily accomplished by adaptive testing, in which the most appropriate items are selected from the item bank for each examinee. Because the SE is smallest when the most appropriate items are administered, short tests can be

quite reliable (Embretson, 1999). Embretson in 1995 show that an adaptive test of 20 items can be more reliable, on average than a fixed content test of 30 items.

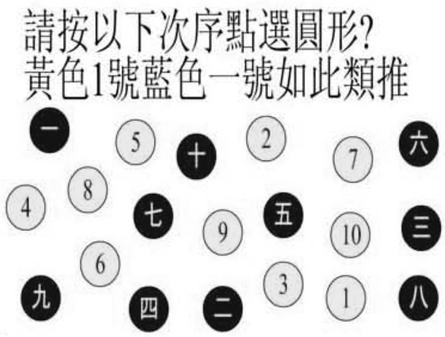
4. Comparing test score across multiple forms is optimal when test difficulty levels vary between persons (Embretson, 1999).
5. Unbiased estimates of item properties may be obtained from unrepresentative sample (Embretson, 1999).

With reviewing the basic assumptions and advantages of IRT, in conclusion, the benefits of IRT applications in scale development is the greater transparency of resulting score (Coster, Ludlow & Mancini, 1999).

Appendix III – Screen shot of some ICAS items



i



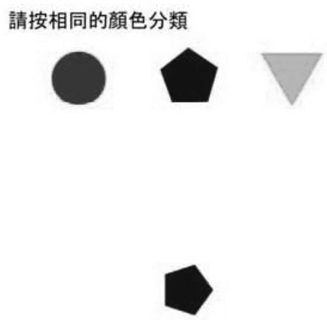
ii

iii

iv

v

vi



vii

viii

x

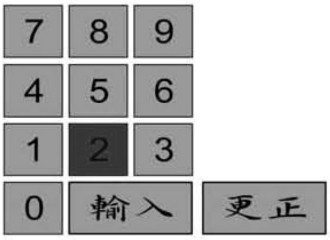
ix

請按密碼鎖後輸入密碼,密碼是5872
請記着此密碼,遲D會再用多一次

請輸入櫃員機密碼,密碼是542875
請記着此密碼,遲D會再用多一次?

xi

請輸入現在的年份?



xii

現在是什麼年份??

2004 2005

不知道

2006 2007

xiii

端午節是幾月幾號?

五月初五 六月初六

五月初四

龍舟競渡 不知道

xiv

圖中所示是什麼?

不知道 教堂

清真寺

佛寺 觀音廟

圖中標誌是什麼意思?

男人 不知道

紅色

圓形 行人止步

xv

圖中所示是什麼?

絲帶 不知道

筆

棍 筆蓋

xvi

Note: Items were not presented in order and computerized instructions of individual testing items were given.

i In this box, there are yellow circles with numbers in them. Try to connect the circles, by touching the computer screen, from 1 to 2, then to 3 and so on, until you reach the end.

ii Connecting the yellow circle with number 1 and then blue circle with Chinese number 1 (一), then yellow circle with number 2, followed by blue circle with Chinese number (二) and so on.

iii Please select by pressing on the stars.

iv Select the one which differs from the rest.

v Indicate the right sequence in boiling a kettle of water.

vi Use store-value card to buy a can of coca-cola from a vending machine.

vii Please match with the same color.

viii Please match with the same kind of clothing.

ix. Enter the password of "5872" for entrance to a building. Try to remember this number which will be recalled later.

x. Enter the password of "542875" for operating an ATM machine. Try to remember this number which will be recalled later.

xi. Please enter the number of year by keying in.

xii. Which year is now?

xiii. Date of Dragon Boat Festival (4^{th} May, 5^{th},May, 6^{th} May, Dragon Boat racing, or do not know)

xiv. Picture showing what? (a chapel, a Guanyin temple, a mosque, a budda temple, or do not know)

xv. The symbol represents what? (man, red color, circle, no pedestrian, do not know)

xvi. The picture showing what? (a ribbon, a pen, a rod, a cap of a ball pen, do not know)

Appendix IV -- Figural result by BP–ANN model with different hidden neurons

Figure 17 Observed MBI score and Predicted MBI score and their difference by 7-5-1 ANN model

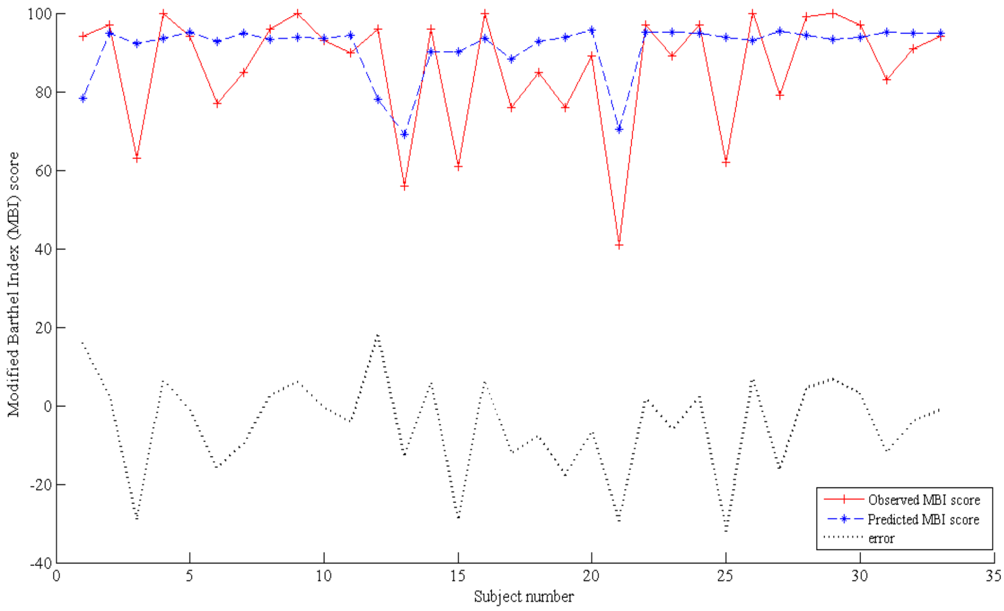


Figure 18 Observed MBI score against Predicted MBI score by 7-5-1 ANN model

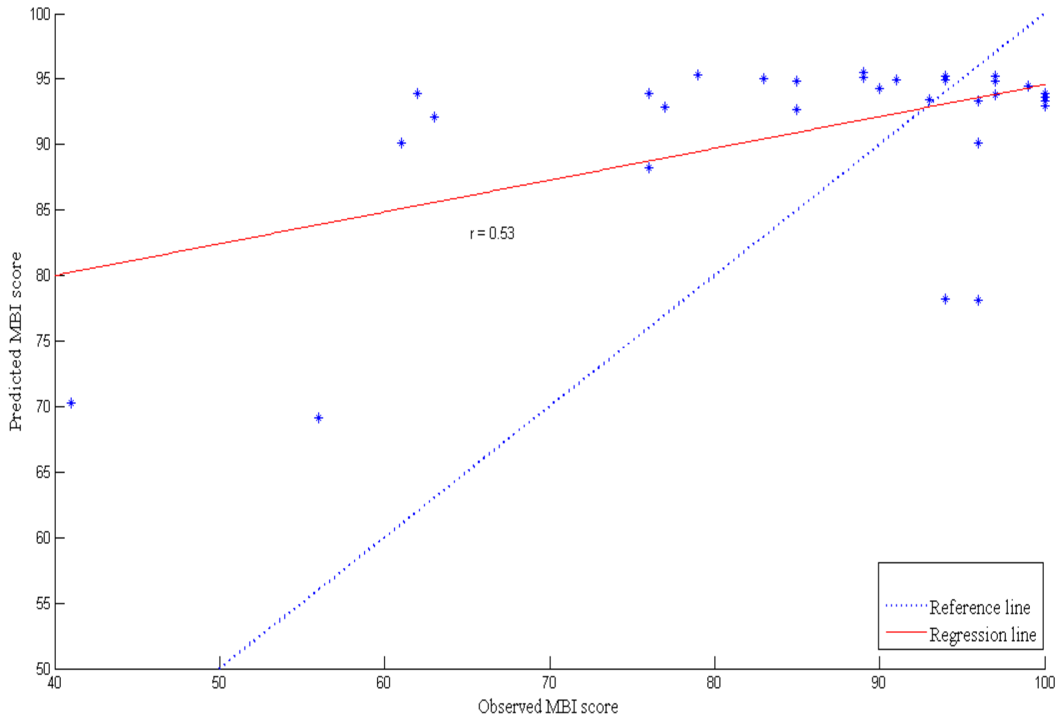


Figure 19 Observed MBI score and Predicted MBI score and their difference by 7-10-1 ANN model

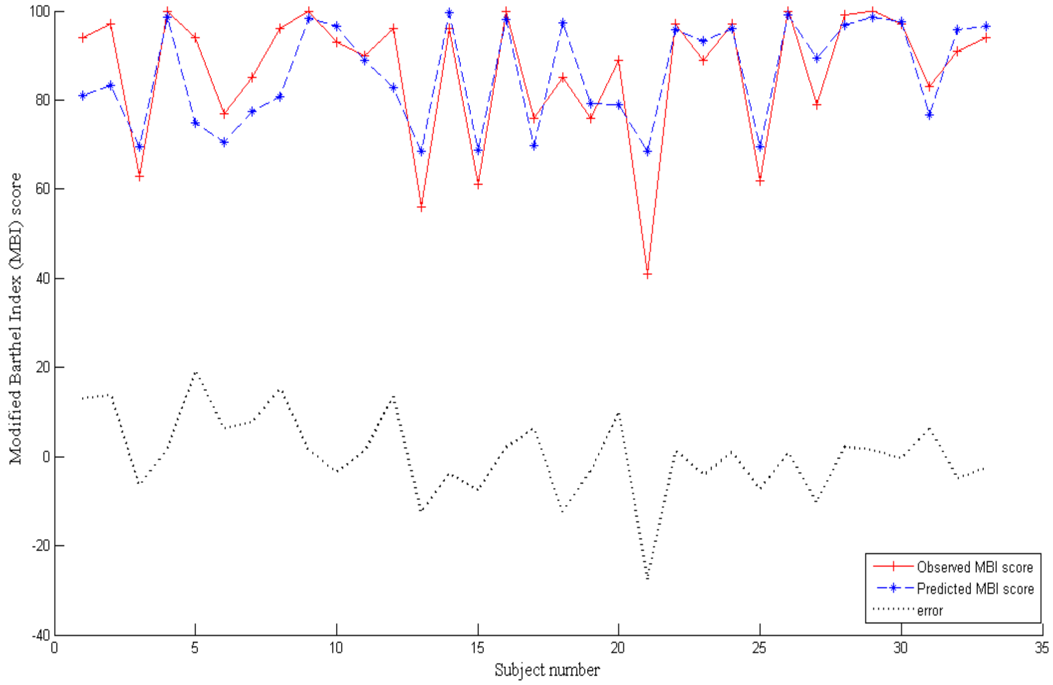


Figure 20 Observed MBI score against Predicted MBI score by 7-10-1 ANN model

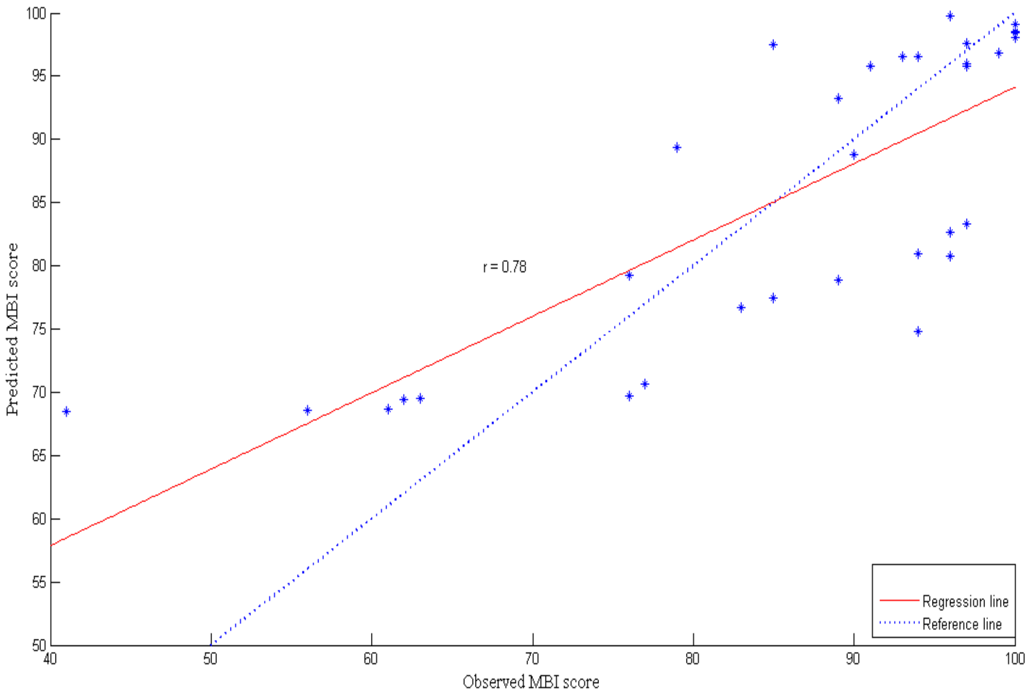


Figure 21 Observed MBI score and Predicted MBI score and their difference by 7-20-1 ANN model

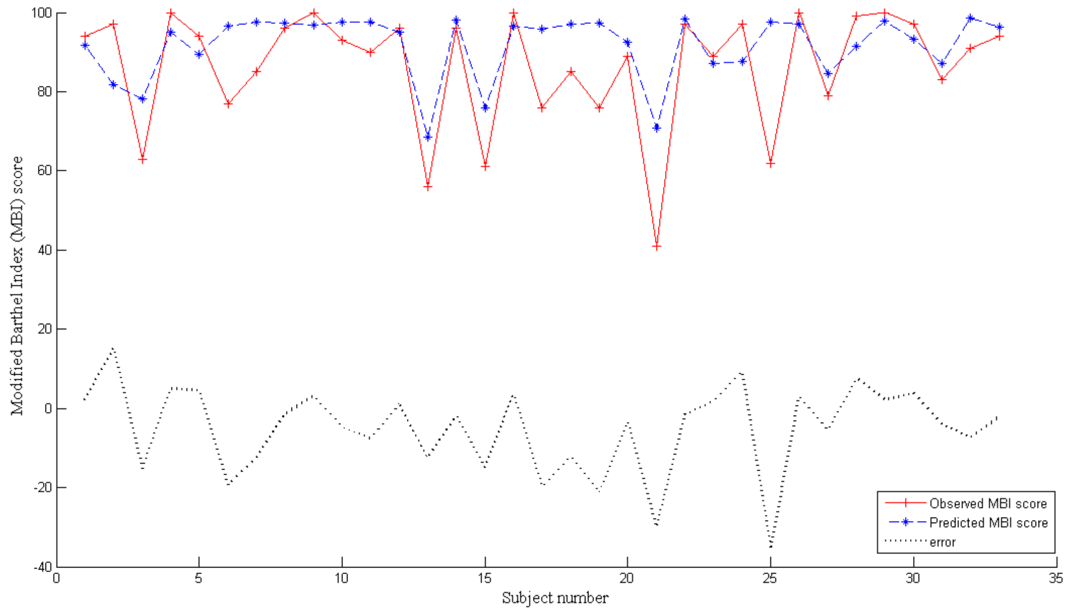
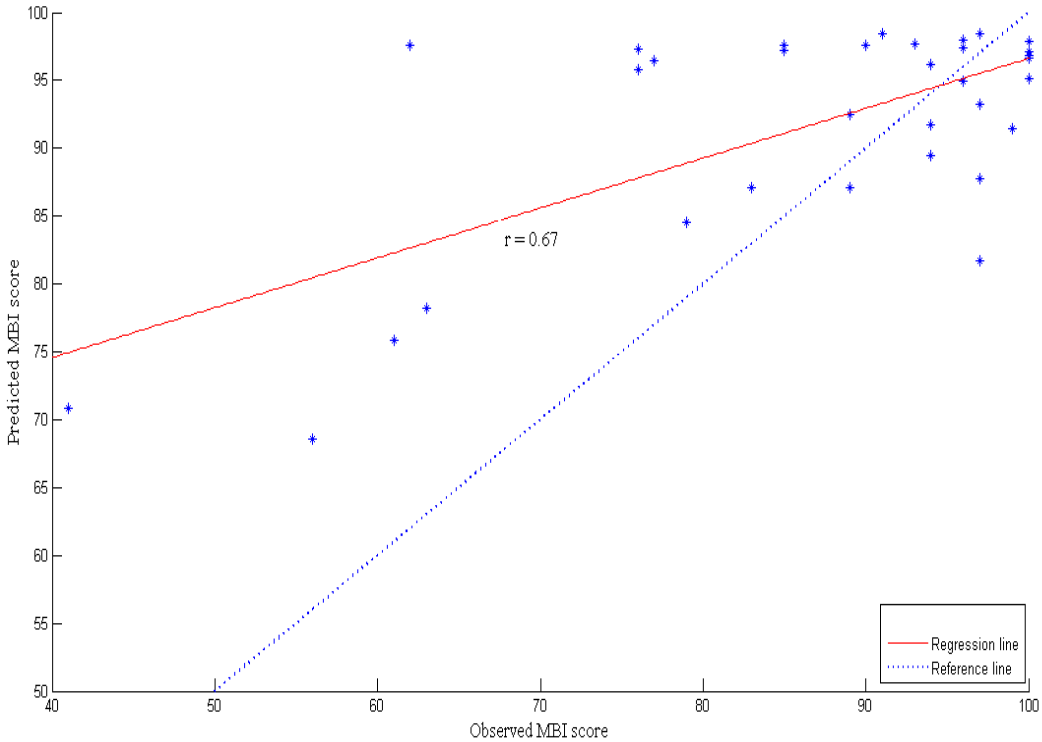


Figure 22 Observed MBI score against Predicted MBI score by 7-20-1 ANN model



Appendix V -- Ethical Approval Form

醫院管理局
HOSPITAL AUTHORITY

群策群力為病人，優質服務滿杏林
Quality Patient-Centred Care Through Teamwork

22 December 2006

Mr YIP Chi-kong
Occupational Therapist II
Occupational Therapy Department
Wong Tai Sin Hospital

Dear Mr YIP

KWC-CREC Reference: KW/EX/06-099

Development and validation of an intelligent cognitive assessment system (ICAS) for persons with cerebral vascular accident (CVA)

The Kowloon West Cluster Clinical Research Ethics Committee (KWC-CREC) is authorized by the Cluster Chief Executive to review and monitor clinical research. It serves to ensure that research complies with the Declaration of Helsinki, local regulations and HA policy. It has the authority to approve, require modifications in (to secure approval), or disapprove research. This Committee has power to terminate / suspend a research at any time if there is evidence to indicate that the above principles and requirements have been violated.

KWC-CREC has approved your research application on 12 December 2006 by expedited review process, and reached the following decision on the documents submitted as shown below. You are required to adhere to the attached conditions.

Study site(s)	**Wong Tai Sin Hospital**
Document(s) approved	I. Clinical Research Ethics Approval Application Form II. Research Proposal (June 2006) III. Research Project Informed Consent Form (Chinese and English versions)
Document(s) reviewed	I. CV of Principal Investigator
Conditions	1. Do not deviate from, or make changes to the study protocol without prior written REC approval, except when it is necessary to eliminate immediate hazards to research subjects or when the change involves only logistical or administrative issues. 2. Apply a clinical trial certificate from Department of Health if indicated. 3. Report the followings to KWC-CREC* : (i) study protocol or consent document changes, (ii) serious adverse event, (iii) study progress (iv) new information that may be relevant to a subject's willingness to continue participation in the study. 4. Report first study progress to KWC-CREC at <u>12-monthly intervals</u> until study closure. [*Forms are available from KWC-CREC intranet webpage]

Please quote the CREC Reference (<u>KW/EX/06-099</u>) in all your future correspondence with the KWC-CREC, including submission of progress reports and requesting for amendments to the research protocol.

If you have any inquiry, please feel free to contact Mr Lewis LI, Secretary of the KWC-CREC, at 2990 3749. Thank you for your attentions.

Yours sincerely,

(Dr TSAO Yen-chow)
Chairperson
Clinical Research Ethics Committee
Kowloon West Cluster

c.c. DM(OT), WTSH

Secretariat of Clinical Research Ethics Committee, Kowloon West Cluster
Room 133, Block J, Princess Margaret Hospital, Lai Chi Kok, Kowloon, Hong Kong Tel (852) 2990 3749 Fax (852) 2990 1059

Appendix VI -- Consent Form (English Version)

Research Project Informed Consent Form

Project title: Development and validation of an intelligence cognitive assessment system (ICAS) for patient with cerebral vascular accident (CVA)

Investigator(s): Dr David Man, PhD

Co-investigator(s): Mr. Yip Chi Kong (Occupational Therapist, Tung Wah Group of Hospitals Wong Tai Sin Hospital)

Purpose of the study:

The objectives of the study were to investigate concurrent validity of the newly developed ICAS with MMSE–CV and NCSE–CV

Who are to be recruited ?

1. Age 60 or above
2. Suffered from Stroke and confirmed by CT brain
3. Hemorrhagic stroke or infraction stroke
4. Medically stable
5. Able to follow instructions

What to do as a participant ?

Participants will be assessed by the ICAS, MMSE–CV and NCSE–CV, the demographic information including age, sex, type of stroke, side affected, upper limb function and score of MBI will be collected for analysis.

Consent

I, ____________________, have been explained the details of this study. I voluntarily consent to participate in this study. I understand that I can withdraw from this study at any time without giving reasons, and my withdrawal will not lead to any punishment or prejudice against me. I am aware of little potential risk in joining this study. I also understand that my personal information will not be disclosed to people who are not related to this study and my name or photograph will not appear on any publications resulted from this study.

I can contact the chief investigator, Dr David Man at telephone 27666711 for any questions about this study. I know I will be given a signed copy of this consent form.

Signature (subject): ____________________ Date: ______________

Signature (witness): ____________________ Date: ______________

Signature (Co-investigator): ________________ Date: ______________

Appendix VII -- Consent Form (Chinese version)

參予研究同意書

研究題目：建立智能認知評核系統及對其有較度及可信度之檢定

總研究員： 文偉光博士

副研究員： 葉志剛（東華三院黃大仙醫院，職業治療部）

研究目的： 本研究旨在檢定智能認知評核系統的有效度及可信度。除此之外，測試智能認知評核系统之運作流暢性為本研究之第二目的。

研究對象：

1. 經 CT 掃描，被主診醫生斷定患上中風或腦血管病之病人
2. 不論出血性中風或缺血性中風
3. 不論男性及女性
4. 年齡在 60 歲以上

參加者須知事項：

參與者需接受簡短智能測驗、智能認知評核系统及腦神經行為認知測驗之考核，然後接受自我照顧功能評核。此外參與者的年齡性別患肢功能及中風的類別亦會一拼搜集並連同認知功能一同作資料分析。

同意聲明

本人, ____________________, 已透過研究員解釋是項研究的目的，並自願參與。本人明白有權於任何時侯退出參與研究而不會被懲罰或控訴，而本人亦明白是項研究並不存有危害性，而有關本人的姓名，相片或病歷資料均會絕對保密，並不會刊登。

本人如對是項研究有任何問題，可致電 27666711 聯絡總研究員文偉光博士。本人並已取得此同意書之副本。

簽署（參加者）： _______________ 日期 ： ____________________

簽署（見証人）： _______________ 日期 ： ____________________

簽署(副研究員）： _______________ 日期 ： ____________________

Appendix VIII -- Data record Form (Phase I)

Data Record Sheet

Name: ☐

Sex : ☐ Male(0)
☐ Female (1)

Age: ☐☐

Type of Stroke:
☐ Infract (0)
☐ Haemorrhage (1)

Affected Side:
☐ Right (0)
☐ Left (1)

Date of Onset: ☐

Date of Assessment: ☐

MMSE: ☐☐

MBI: ☐☐☐

FIM Motor: ☐☐

ULFT ☐

Cognitive Ability: N A

NCSE:

Orientation: ☐☐

Attention: ☐☐

Comprehension: ☐☐

Repetition: ☐☐

Naming: ☐☐

Construction: ☐☐

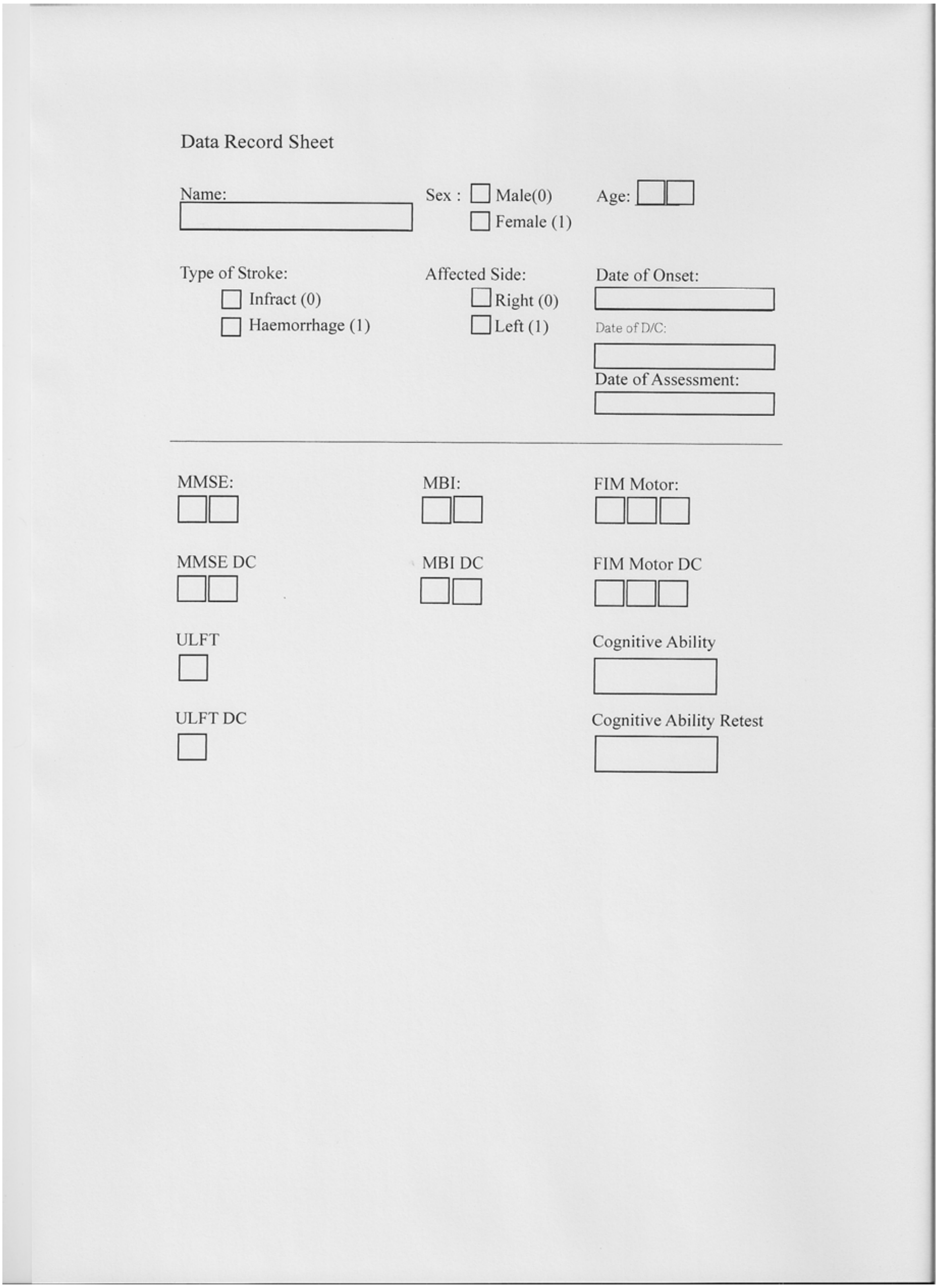
Memory: ☐☐

Calculation: ☐☐

Similarity: ☐☐

Judgment: ☐☐

Appendix IX -- Data Record Sheet (Phase III)



Data Record Sheet

Name:

Sex : ☐ Male(0) ☐ Female (1)

Age: ☐☐

Type of Stroke:
☐ Infract (0)
☐ Haemorrhage (1)

Affected Side:
☐ Right (0)
☐ Left (1)

Date of Onset:

Date of D/C:

Date of Assessment:

MMSE: ☐☐

MBI: ☐☐

FIM Motor: ☐☐☐

MMSE DC ☐☐

MBI DC ☐☐

FIM Motor DC ☐☐☐

ULFT ☐

Cognitive Ability

ULFT DC ☐

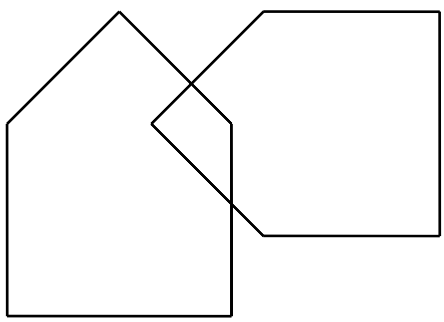
Cognitive Ability Retest

Appendix X -- Mini-Mental State Examination (Chinese Version)

職業治療部

Mini-Mental State Examination

	日期	/ /	/ /
<u>I. Orientation</u>			
1	依家係乜野日子 --- 年份、季節、月份、日期、星期幾？	0 1 2 3 4 5	0 1 2 3 4 5
2.	我地依家係邊喥？-- 城市、地區、區域、醫院名稱、幾多樓？	0 1 2 3 4 5	0 1 2 3 4 5
	（九龍/新界/香港）（九龍/新界/香港既邊度）（醫院）（邊層樓)(病房）		
或	（九龍/新界/香港）（九龍/新界/香港既邊度）（邊一科診所）（診所名字）（邊層樓）		
或	（九龍/新界/香港）（九龍/新界/香港既邊度）（邊條街）（邊一座）（邊層樓）		
或	（九龍/新界/香港）（九龍/新界/香港既邊度）（邊個屋村）（中心名字）（邊層樓）		
<u>II. Registration</u>			
1.	依家我會講三樣野既名，講完之後，請你重複一次。 請記住佢地 ，因為幾分鐘後，我會叫你再講番俾我聽。(蘋果)、(報紙)、(火車)。 依家請你講番哩三樣野俾我聽。（以第一次講的計分，一個一分；然後重複物件，直至全部三樣都記住。）	0 1 2 3	0 1 2 3
<u>III. Attention & Calculation</u>			
1.	請你用一百減七，然後再減七，一路減落去，直至我叫你停為止。(減五次後便停）	0 1 2 3 4 5	0 1 2 3 4 5
或	依家我讀幾個數目俾你聽，請你倒轉頭講番出黎。（ 4 2 7 3 1 ）		
<u>IV. Recall</u>			
1.	我頭先叫你記住既三樣野係乜野呀？	0 1 2 3	0 1 2 3
<u>V. Language</u>			
1.	哩樣係乜野？（手錶）(鉛筆）。	0 1 2	0 1 2
2.	請你跟我講句說話（ 姨丈買魚腸 ）	0 1	0 1
3.	依家檯上面有一張紙。用你既右手拿起張紙，用兩隻手一齊將紙摺成一半，然後放番張紙係檯上面。	0 1 2 3	0 1 2 3
4.	請讀出哩張紙上面既字，然後照住去做。	0 1	0 1
	# 拍　　手		
5.	請你講任何一句完整既句子俾我聽。 例如：(我係一個人)、(今日天氣好好)。	0 1	0 1
6.	哩處有幅圖，請你照住黎畫啦。	0 1	0 1
	總分	/30	/30
	簽名		
	姓名		
	職級		



Appendix XI -- Neurobehavioral Cognitive Status Examination (Chinese Version)

神經行為認知狀況測試〔簡稱 NCSE〕

(Gum label)

____

==

姓名：__________________　　　　日期：________________________

職業：__________________　　　　時間：________________________

母語：__________________

偏手傾向〔圈出來〕：左　右　　　主考人：_____________________

教育程度：____________________　測驗地點：___________________

==

認知程度概況

	意識程度	定向能力	專注能力	語言能力：理解能力	語言能力：覆述能力	語言能力：命名能力	結構組織能力	記憶能力	計算能力	推理能力：類似性	推理能力：判斷能力
							-6-			-8-	-6-
正常	清醒	-12-	-(S)8-	-(S)6-	-(S)-	-(S)-	-(S)5-	-12-	-(S)4-	-(S)6-	-(S)5-
					-12-	-8-					
		-10-	-6-	-5-	-11-	-7-	-4-	-10-	-3-	-5-	-4-
輕微	受損	-8-	-5-	-4-	-9-	-5-	-3-	-8-	-2-	-4-	-3-
中度		-6-	-3-	-3-	-7-	-3-	-2-	-6-	-1-	-3-	-2-
嚴重		-4-	-1-	-2-	-5-	-2-	-0-	-4-	-0-	-2-	-1-
寫下最低的分數											

(S)：甄別試合格

* 此項測試的準確性取決於是否嚴格地依照 NCSE 手冊執行。

※ 病人如果超過 65 歲，在測試其組織能力、記憶力及類似性時，若分數等同「輕微受損程度」一級，仍屬正常。

注意：並非所有因腦部受損而導致的認知缺陷都可從 NCSE 測試出來。故此，表示正常的分數不足以証明腦部*沒有*毛病。同樣地，表示輕微、中度或嚴重受損的分數也*不一定*反映出腦部出現機能障礙。〔參閱 NCSE 手冊中的「闡明須知」〕

腦神經行為認知狀況測試〔簡稱 NCSE〕

清楚正確地記錄病人的回應

===

一・ 意識程度： 清醒 _____ 呆滯 _____ 不穩定 _____

描述病人的情況：

____

____

____

二・ 定向能力〔分數為 2, 1 或 0〕

回應 分數

甲．人物 1.姓名〔0 分〕 _________________________ _____

2.年齡〔2 分〕 _________________________ _____

乙．地點 1.現時位置〔2 分〕_________________________ _____

2.區域名稱〔2 分〕_________________________ _____

丙．時間 1.日期：月〔1 分〕_____ 日〔1 分〕____

年〔2 分〕______ _____

2.星期〔1 分〕 _________________________ _____

3.一小時內的當時時間 (1 分)________________ _____

總分 _____

三・ 專注能力

甲．數字覆述

1. 甄別試： 8-3-5-2-9-1 合格 _____ 不合格 _____

2. 等級試： 數字分組覆述〔分數為 1 或 0；若在覆述一組數字時出現兩次錯誤，則停止此項測試。〕

	分數		分數		分數		分數
3-7-2	____	5-1-4-9	____	8-3-5-2-9	____	2-8-5-1-6-4	____
4-9-5	____	9-2-7-4	____	6-1-7-3-8	____	9-1-7-5-8-2	____

總分 ____

乙．四詞記憶測試

從第六節中選出四個不相關的詞語：鸚鵡、蘿蔔、鋼琴、綠色。
〔其他選擇：桌子、獅子、蘋果、手套〕病人必須正確地把這四個詞語覆述兩次〔參閱手冊〕並把病人所須的練習次數記錄下來：______

四．語言能力

甲．看圖描述

1. 釣魚圖畫〔清楚正確地記錄病人的回應〕

____

____

____

____

乙．理解能力〔進行此項測試時，必須最少把三件其他物件同時放於病人的面前〕假如 (i), (ii) 和 (iii) 能順利完成，此項測試的習題可設定為標準。〕

1. 甄別試：三步指令:「翻轉張紙，把原子筆遞給我跟著指著自己的鼻子。」

合格 _____ 不合格 _____

2. 等級試：[分數為 1 或 0]如果不正確，請描述病人的表現。

	反應	分數
(i) 拾起原子筆	______________________________	_____
(ii) 指向地板	______________________________	_____
(iii) 把鑰匙交給我	______________________________	_____
(iv) 指著原子筆跟著拾起鑰匙	______________________	_____
(v) 把張紙遞給我跟著指著硬幣	____________________	_____
(vi) 指著鑰匙、把原子筆遞給我跟著拾起硬幣	______________________________________	_____

總分 _____

丙．覆述能力

1. 甄別試：第一個動作顯示了作曲家的意圖

____

合格 _____ 不合格 _____

2. 等級試：〔第一次答對得 2 分，第二次答對得 1 分，答錯則 0 分〕

		回應	分數
(i)	在窗外面	______________	_____
(ii)	他游過那個湖	______________	_____
(iii)	那彎曲的道路通往那條村莊	______________	_____
(iv)	他讓門半開著	______________	_____
(v)	那蝙蝠洞吸引了一群遊歷的人	______________	_____
(vi)	不是如果、和或但是	______________	_____

總分 _____

丁．命名能力

1. 甄別試 (i)原子筆 ____ (ii)筆蓋 / 筆套 ____

(iii) 筆夾 ____ (iv)筆尖 / 筆咀 ____

合格 ____ 不合格 ____

2. 等級試 〔分數為 1 或 0〕

	回應	分數		回應	分數
(i) 鞋	________	_____	(v) 楊桃	________	_____
(ii) 巴士	________	_____	(vi) 錨	________	_____
(iii) 梯子	________	_____	(vii) 魔鬼魚	________	_____
(iv) 風箏	________	_____	(viii) 楊琴	________	_____

總分 ______

五・結構組織能力

甲. 甄別試：視覺記憶測試〔讓病人觀察測試用的圖案板，限時10秒，然後要求病人憑記憶畫出板上的圖案，所畫的圖案必須與板上的完全相同才算合格，如病人不能畫出相同的圖案，測試者可要求病人依照板上的圖案抄畫出來。〕

合格 ____　　不合格 ____

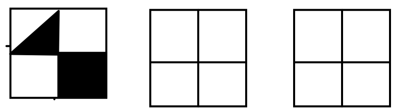
乙.等級試：組合圖案〔能夠在0-30秒內正確地完成得2分，31-60秒內才完成得1分、超過60秒才完成或仍然不正確則得0分。〕

請把方塊如下圖所示放於病人的面前

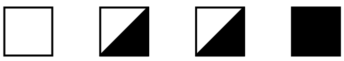
把不正確的圖案記錄在下面的方格中

時間　分數

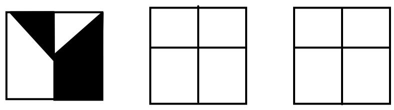
1. 圖案一：

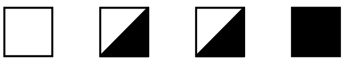


____　____



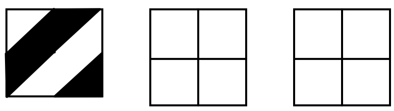
2. 圖案二：





____　____

3. 圖案三：



____　____

總分 _____

六．記憶能力〔如不需要提示下記起得3分，如需要類別提示才記起得2分，從目錄中選出正確答案得1分、選擇錯誤得0分。〕核對是否正確。

<u>詞語</u>	*核對*	<u>類別提示</u>	*核對病人的答案*
鸚鵡	____	雀鳥	________________
蘿蔔	____	蔬菜	________________
鋼　琴	____	樂器	________________
綠　色	____	顏色	________________

<u>目錄</u>〔圈出來〕	分數
麻雀、鸚鵡、了哥	_____
蘿蔔、薯仔、洋蔥	_____
小提琴、結他、鋼琴	_____
紅色、綠色、黃色	_____

第一個錯誤的回應：____

總分 ____

七．計算能力

甲．甄別試：5 x 13　答案：_____　時間：_____

〔病人必須在20秒內答對〕

合格 _____　不合格 _____

乙．等級試：〔20秒內答對得1分〕可重覆問題，但不會停止計時。

	答案	時間	分數
1.　5 + 3 等如幾多？	_____	_____	_____
2.　15 + 7 等如幾多？	_____	_____	_____
3.　39 ÷ 3 等如幾多？	_____	_____	_____
4.　31 - 8 等如幾多？	_____	_____	_____

總分 ______

八・推理能力

甲.類似性〔解釋：「帽子和外套相似的原因是它們都是衣服的種類。」假如病人不作答，必須鼓勵病人作答；如果病人所答的原因與標準答案不符合，則0分〕

1. 甄別試：一幅畫、音樂〔原因必須是抽象的；答案只可以是「藝術」、「藝術性」或「藝術的一種」

____

合格 _____　　不合格 _____

2. 等級試：〔抽象的答案得2分；答案若是部份正確的得1分;答錯則0分〕例子可參閱手冊，核對答案是否正確。

	核對	抽象概念	其他答案	分數
(i) 玫瑰、劍蘭	____	花		____________________
(ii) 的士、地鐵	____	交通工具		____________________
(iii) 手錶、間尺	____	量度工具		____________________
(iv) 罐頭刀、鎚	____	工　具		____________________

總分 _____

乙．判斷能力

1. 甄別試：假如你流落在大嶼山的梅窩碼頭，但是口袋裏只有 $1，你會怎樣做？

____

____

____

合格 ______ 不合格 ______

2. 等級試：〔答對得 2 分；部分答對得 1 分；答錯 0 分〕

(i) 假如你在早上 8:00 前一分鐘起床，記得自己要在 8:00 到市區出席一個重要的約會，你會怎樣做？

____

____

____

分數 ______

(ii) 假設你在海邊散步，看見一個 2 歲的小孩獨自在碼頭的盡頭玩耍，你會怎樣做？

____

____

分數 ______

(iii)假如當你回家的時候，發現一條水管爆裂，廚房被水浸，你會怎樣做？

____

____

____

分數 ______

總分 ______

九・服用藥物

列舉所有目前服用的藥物和份量

1. ____________ 2. ____________ 3. ____________ 4. ____________

5. ____________ 6. ____________ 7. ____________ 8. ____________

十・概括意見

記下任何已知或從觀察得知那些可以影響此項測試的缺陷，不論在肢体運動、感官或知覺各方面〔例如：視覺或聽覺受損、顫抖、活動組織能力失控、發音困難〕

____

____

____

記下「測試過程的特點」如分心、不耐煩、疲乏和合作程度必須 同時記下病人對自己表現的印象

____

____

____

==

視覺記憶測試用的空位

Appendix XII -- Modified Barthel Index

Occupational Therapy Department
Modified Barthel Index

Patient Label

Item	Score	Rating Description	Initial		D/C
		Date:			
Personnel Hygiene	0	Dependent in all aspects, unable to attend to hygiene			
	1	Assistance in all steps of hygiene			
	3	Some assistance in one or more steps of hygiene			
	4	Able to conduct but needs minimal assistance			
	5	Independent			
Feeding	0	Dependent in all aspects and needs to be fed			
	2	Can manipulate eating device, but need active assistance			
	5	Able to feed self with supervision			
	8	Independent in feeding with prepared tray			
	10	Can feed self and put on device of needed			
Dressing	0	Dependent in all aspects of dressing			
	2	Able to participate to some degree			
	5	Assistance is needed in putting on, and/or removing clothes			
	8	Minimal assistance is required with fastening clothes			
	10	Independent in all aspects			
Toileting	0	Fully dependent in toileting			
	2	Assistance is required in all aspects of toileting			
	5	Assistance is required in clothing, transferring or washing hands			
	8	Supervision required for safety with normal toileting			
	10	Independent (transferring, managing clothes and hygiene)			
Bathing	0	Total dependent in bathing self			
	1	Assistance is required in all aspects of bathing			
	3	Assistance is required with either transfer, wash or dry			
	4	Supervision on safety in adjusting water temp. or in transfer			
	5	Independent to do all step			
Bed-Chair	0	Unable to participate in transfer & 2 assistants required			
	3	Able to participate but max. assistance of 2 person required			
	8	Assistance of one person in any aspect of transfer			
	12	Supervision for safety required			
	15	Independent in all phases of transfer			
Ambulating	0	Dependent in ambulation			
	3	Constant presence of 1 or more assistants in ambulation			
	8	One person is required to offer assistance			
	12	Independent in less than 50 meters supervision for safety			
	15	Independent in ambulation>50m and using aids if necessary			
* Wheelchair	0	Dependent in wheelchair ambulation			
	1	Can propel in short distances on flat surface, not need help in all steps			
	3	Presence of 1 person for assistance in chair to bed, table, etc			
	4	Can propel self for a reasonable duration & need minimal assistance around corner			
	5	Can propel wheelchair independently for at least 50 meters			
Stair Climbing	0	Unable to climb stairs			
	2	Assistance in all aspects of stair climbing including aids			
	5	Able to ascend and descend but is unable to carry aids			
	8	Generally need supervision for safety			
	10	Able to go up and down stairs without help			
Bladder Control	0	Dependent in bladder management, incontinent or has indwelling catheter			
	2	Incontinent but is able to assist with the application of an internal or external device			
	5	Generally dry by day, but not at night, and needs some assistance with the device			
	8	Generally dry by day and night, but may have an occasional accident, or need minimal assistance with internal or external devices			
	10	Able to control bladder day and night, and/or is independent with internal or external devices			
Bowel Control	0	Bowel incontinent			
	2	Needs help to assume appropriate positions and with bowel movement facilitatory techniques			
	5	Assume appropriate position, but cannot use facilitatory techniques, or clean self without assistance and has frequent accidents			
	8	Requires supervision with the use of the suppository or enema and has occasional accidents			
	10	Control bowels and has no accidents, can use suppository, or take an enema when necessary			
		Total score:			
		Signature:			
		Therapist Name:			

* Score Wheelchair only when Ambulating rate 0

#Shah et al (1989)

0-20 total dependent
21-60 severe dependence
61-90 moderate dependence
91-99 slight dependence
100 independent of assistance from others

Appendix XIII -- Upper Extremity Functional Test

Occupational Therapy Department
Upper Limb Functional Assessment Form

Name:
Sex/Age:
Diagnosis:

Upper Extremity Functional Test (For stroke patient only) Grade:

(--) unable to complete; (+) complete the task

	Date	/ /		/ /		/ /	
Level	Task	Grade	Time	Grade	Time	Grade	Time
1	Nil						
2	A. Associated reaction						
	B. Hand into lap						
3	C. Arm clearance during shirt tuck						
	D. Hold a pouch		15 sec		15 sec		15 sec
4	E. Stabilize a jar						
	F. Simulate "Wringing a rag"						
5	G. "Blocks and box"						
	H. Eat with a Spoon						
6	I. Box on shelf						
	J. Drink from Glass						
7	K. Key Turning						
	L1.Use Chopsticks (dominant hand)						
	L2.Clip cloth peg (non-dominant hand)						
	O.T. signature						
	Name & Rank						