International Psychogeriatrics (2004), 16:1, 13–31 © 2004 International Psychogeriatric Association
DOI: 10.1017/S1041610204000043 Printed in the United Kingdom

2003 IPA RESEARCH AWARDS FIRST-PRIZE WINNER

The Rowland Universal Dementia Assessment Scale (RUDAS): a multicultural cognitive assessment scale

Joella E. Storey,[1] Jeffrey T. J. Rowland,[1] David A. Conforti[1] and Hugh G. Dickson[1,2]

[1] *Geriatric Medicine, Liverpool Hospital, Sydney, Australia*
[2] *Aged Care and Rehabilitation, University of New South Wales, Australia*

ABSTRACT

Objective: To develop and validate a simple method for detecting dementia that is valid across cultures, portable and easily administered by primary health care clinicians.

Design: Culture and Health Advisory Groups were used in Stage 1 to develop culturally fair cognitive items. In Stage 2, clinical testing of 42 items was conducted in a multicultural sample of consecutive new referrals to the geriatric medicine outpatient clinic at Liverpool Hospital, Sydney, Australia ($n = 166$). In Stage 3, the predictive accuracy of items was assessed in a random sample of community-dwelling elderly persons stratified by language background and cognitive diagnosis and matched for sex and age ($n = 90$).

Measurements: A research psychologist administered all cognitive items, using interpreters when needed. Each patient was comprehensively assessed by one of three geriatricians, who ordered relevant investigations, and implemented a standardized assessment of cognitive domains. The geriatricians also collected demographic information, and administered other functional and cognitive measures. DSM-IV criteria were used to assign cognitive diagnoses. Item validity and weights were assessed using frequency and logistic regression analyses. Receiver-operating characteristic (ROC) curve analysis was used to determine overall predictive accuracy of the RUDAS and the best cut-point for detecting cognitive impairment.

Results: The 6-item RUDAS assesses multiple cognitive domains including memory, praxis, language, judgement, drawing and body orientation. It appears not to be affected by gender, years of education, differential performance factors and preferred language. The area under the ROC curve for the RUDAS was 0.94

Correspondence should be addressed to: Joella E. Storey, Aged Care Research, Liverpool Hospital, Locked Bag 7103, Liverpool BC, NSW, 1871, Australia. Phone: 02 9828 6516. Fax: 02 9828 6199. E-mail joella_storey@swahs.nsw.gov.au.
Date received: 3 Jul 03; Date accepted: 28 Jul 03.

(95% CI 0.87–0.98). At a cut-point of 23 (maximum score of 30), sensitivity and specificity were 89% and 98%, respectively. Inter-rater (0.99) and test-retest (0.98) reliabilities were very high.

Conclusions: The 6-item RUDAS is portable and tests multiple cognitive domains. It is easily interpreted to other languages, and appears to be culturally fair. However, further validation is needed in other settings, and in longitudinal studies to determine its sensitivity to change in cognitive function over time.

Key words: dementia, multicultural, cognition, assessment, culturally fair, primary health care

Introduction

Projected increases in migration and the aging population will have significant demographic and social consequences for many of the world's countries during the first half of the twenty-first century (Martin, 2001; United States Department of Commerce *et al.*, 1999). In 2001, approximately 150 million people lived outside their country of birth (Martin, 2001). By 2025, the world's population 65 years of age and above will more than double, and economically, the world community will face an elderly support burden almost 50% larger than that in 1998 (United States Department of Commerce *et al.*, 1999).

Disorders such as dementia, which disproportionately affect the oldest age groups, cause considerable morbidity to patients and carers, and generate large health-care costs. Early detection of dementia is critical for the purposes of differential diagnosis, secondary prevention and psychosocial intervention (Brodaty and Moore, 1997; Sandson and Price, 1996; Small, 1998). Developing a simple method for detecting dementia that is valid across cultures, easily portable and easily administered by primary health care clinicians would provide more equitable access to health care to those already at high risk of poor outcomes.

Most previous attempts to develop instruments for this use in culturally diverse populations modified or translated scales that were originally developed in dissimilar groups. For example, the Folstein Mini-Mental State Examination (MMSE) was developed in an English-speaking population, but is nevertheless commonly used to assess cognition in persons from culturally and linguistically diverse backgrounds (Folstein *et al.*, 1975). MMSE scores are influenced by age, education, ethnicity and language of the interview (Escobar *et al.*, 1986). Many words cannot be easily translated and several concepts are not relevant to people from other cultures. On excluding items that might be culturally biased, ethnic differences in the rates of "severe" cognitive impairment disappeared. Escobar concluded that the MMSE should be revised to diminish social and education artefacts, through item selection and weighting. Furthermore, the

MMSE usually fails to detect cognitive impairment primarily involving the frontal lobes (Royall *et al.*, 1994).

The Fuld Object-Memory Evaluation (FOME) was an early attempt to develop an instrument that was culturally fair (Fuld *et al.*, 1988). This instrument is not easily portable, tests only a few cognitive domains, and is difficult to standardize across cultural settings. In a study of elderly African and European Americans, the FOME had a sensitivity of 93%, but a specificity of only 63.5% (Mast *et al.*, 2001).

The Cognitive Abilities Screening Instrument (CASI) borrowed items from the MMSE, the Modified Mini-Mental State Test and the Hasegawa Dementia Screening Scale (Folstein *et al.*, 1975; Hasegawa, 1983; Teng and Chui, 1987; Teng *et al.*, 1992). Comparisons across cultural groups are difficult, due to a lack of standardized items in multiple versions of the instrument and the significant effect for education in some cultural groups (Graves *et al.*, 1993; Shadlen *et al.*, 2001).

The Elderly Cognitive Assessment Questionnaire (ECAQ) was developed in Singapore in a predominantly male population, and includes 10 items taken from the MMSE and the Geriatric Mental State Schedule (GMS) (Copeland *et al.*, 1976; Folstein *et al.*, 1975; Kua and Ko, 1992). The ECAQ uses a 4-digit number to test memory recall, and relies heavily on demographic information (such as date of birth) that may be difficult to confirm in patients from diverse cultural backgrounds. The ECAQ reports a sensitivity of 85.3% and specificity of 91.5%.

The 2-stage Cross-Cultural Cognitive Examination (CCCE) was developed for use in cross-cultural epidemiological dementia research (Glosser *et al.*, 1993). The CCCE was developed in subjects from Guam and the United States, and used the translation-back-translation procedure to ensure cultural fairness. The reported 100% sensitivity and 83% specificity of the 5-minute, 6-item screen and the 94% sensitivity and 99% specificity of the 20-minute mental state exam are impressive. However, the average age of the participants in the Guam validation study was 52.6 years, and only 18 of 115 were diagnosed as definitely or possibly demented. The young age of the participants is unusual in dementia research, and the results may be difficult to generalize to other older populations.

The Community Screening Interview for Dementia (CSI'D) also used the translation-back-translation procedure to develop a lengthy cognitive and informant-based instrument for use in Cree-speaking natives living on reserves in Manitoba, and English-speaking residents in Winnipeg (Hall *et al.*, 1993). The CSI'D is further limited by the need to interview a reliable informant. A revised version was recently evaluated in the original samples as well as in African Americans in Indianapolis and Jamaica and the Yoruba population in Nigeria (Hall *et al.*, 2000). The areas under the receiver-operating characteristic

(ROC) curve with and without informant data were 0.82–0.97 and 0.74–0.93, respectively. Education had a significant effect at most sites.

The Mini-Cog is the shortest and most portable of the recent cross-cultural instruments (Borson *et al.*, 2000). It combined a 3-item memory task and a clock drawing task, and reported a sensitivity of 99% and a specificity of 93% in the development sample (validation data not reported). While the 3-item recall task was the more powerful component in the Mini-Cog, our work suggests that the accuracy of clock drawing may be modest at best in multicultural samples (Storey *et al.*, 2002).

Researchers who believe that behavior can only be understood within the cultural context in which it occurs suggest that concepts and instruments should be developed for each distinct cultural group (Kim, 2000). Can we reasonably believe however, that Italians who migrated to Australia 50 years ago are similar culturally and linguistically to those who have always lived in Italy? Acculturation research suggests that the same criteria should not be used to compare natives living in their country of cultural origin to those who have relocated to a different culture (Berry, 1997; Bourhis *et al.*, 1981; Harwood *et al.*, 1994). Furthermore, the existence of several thousand different cultures and languages makes the development of distinct instruments for each group clearly impractical (Bernatzik, 1957; Swadesh, 1967). While we acknowledge that both inherited and environmental factors probably influence cognition, we believe that humans have common experiences that allow cognition to be measured independent of cultural differences (Berry *et al.*, 1992).

In this paper, we report the development and validation of the Rowland Universal Dementia Assessment Scale (RUDAS), a scale that fulfils many of the requirements needed to accurately assess cognition in culturally diverse populations.

Method

Stage 1 – Development of Items

AIMS

The aims of the item-development stage were to:

1. Identify cognitive domains important in the assessment of cognition (particularly those associated with early cognitive impairment).
2. Propose potential items to measure cognition in each of these domains.
3. Optimize the psychometric validity, and cultural and linguistic equivalence of the proposed items.
4. Develop a final list of items to test in a culturally heterogeneous population.

Table 1. Aims and outcomes of advisory groups in item development

GROUP	AIM	OUTCOME
1st Health	Identify cognitive domains important in assessment of cognition.	List of domains, including memory, learning, attention, orientation, language, gnosis, visuo-spatial, perseveration/initiation, praxis, planning, judgment, insight, function and behavior.
1st Culture	Explore the meaning of culture. Assess cultural relevance of proposed cognitive domains for conceptual relevance and importance, cultural sensitivity and specificity, and role in normal and abnormal function.	All domains endorsed as culturally relevant.
2nd Health	Propose potential items for each cognitive domain (previously published and validated items, and experiential-based items).	List of approximately 80 items.
2nd Culture	Assess proposed items for cultural and linguistic fairness, and suggest modifications where appropriate.	Refined list of items (some discarded, some modified).
3rd Health	Revise validity of culture modified items.	Validity-based revised list of items.
3rd Culture	Final cultural and linguistic item review.	List of 60 items for clinical testing.

ADVISORY GROUPS

After a comprehensive review of the literature, we formed two advisory groups to help us identify and appraise potential cognitive domains and items. The Health Group included professionals from a number of health disciplines (geriatric medicine, aged care psychiatry, neuropsychology, nursing, occupational therapy, physiotherapy, social work and speech therapy) whose role was to advise on the validity of culturally and linguistically modified items. The aim of the Culture Group (representatives from 22 cultural and linguistic groups) was to advise on the cultural and linguistic equivalence of proposed cognitive items, and to suggest modifications where relevant and appropriate. The iterative process used to conduct the two groups, and the aims and outcomes of each are described in Table 1.

ITEM-INCLUSION CRITERIA AND SCORING

The advisory groups endorsed 60 potential items for clinical testing. These items were reviewed by a cross-section of professional language interpreters, and finally

Table 2. Pilot items and cognitive domains

DOMAIN	ITEM – BRIEF DESCRIPTION
Orientation, memory and learning	Full name, country of birth, country now, city or town, month, day of week, 4-item animal recall, 4-item grocery recall, boy-big-dog sentence recall, circle pointing, children-fire story recall, design recognition and recall
Visuo-spatial and gnosis (including construction and drawing)	Body orientation, body placement, tell the time, pattern discrimination, copy (cube, lemniscates, intersecting pentagons, draw clock
Language and judgement	Animal generation, food generation, naming objects, 3-stage command, naturalistic conversation on childhood games, crossing the road, bathroom fall, kitchen fire, locked-out
Attention and perseveration/ initiation	Symbol cancellation, days of week reversed, counting backwards from 20, finger tapping, ramparts
Planning and praxis	Simple maze, simple commands, fist-palm alternation task, fist-palm-side task
Insight, behavior and function	Self-rated insight, assessor-rated insight, behavioral rating by assessor, functional rating by assessor

by the study authors. The item-inclusion criteria were:

1. Good construct validity.
2. Independent of factors rich in cultural and linguistic variance or specific to a small number of cultures and languages.
3. Endorsed by interpreters.
4. Relatively short and easy to administer, without the need for complex stimuli or cue cards.

The 42 items which met the inclusion criteria were operationalized and retained for clinical testing in an outpatient setting (Table 2).

Stage 2 – Clinical testing of items

CLINIC PROCEDURE

The 42 items were tested on consecutive new referrals to the Geriatric Medicine outpatient clinic at Liverpool Hospital, in Sydney, Australia (January to December 2000). At the beginning of the clinic appointment, a research psychologist (JES) administered the items, blinded to patient diagnosis and clinical history. The items were randomly ordered to control for order effects, and professional health interpreters were used for all patients who did not speak English. Each patient was then comprehensively assessed by one of three geriatricians (JTR, DB and DAC), who ordered relevant investigations (including cerebral CT scans and laboratory tests), and implemented a standardized assessment of cognitive domains that excluded the 42 items for testing, so as to avoid

contamination. The geriatricians also collected demographic information, and administered the Modified Barthel Index (MBI) (Wade and Collin, 1988), the Instrumental Activities of Daily Living (IADL) Scale (Lawton and Brody, 1969), the Geriatric Depression Scale (GDS) (Yeasavage *et al.*, 1983), and the Clinical Dementia Rating (CDR) Scale (Berg, 1988). DSM-IV criteria for Dementia and Other Amnestic Disorders were used to assign cognitive diagnoses (American Psychiatric Association, 1994). Each patient was classified as normal, cognitively impaired but not demented (including those with age-associated cognitive decline), or demented. The study protocol was approved by the institutional review committee of the South Western Sydney Area Health Service, and all patients (or their proxies) gave informed consent.

SAMPLE SIZE

The planned sample size of 150 for clinical testing was calculated on the basis of several factors. High sensitivity is more important than high specificity for a scale primarily intended for targeted screening. An estimated sensitivity of 80% (95% confidence interval (CI of $\pm$ 10%) in a non-English speaking background (NESB) population needs a minimum of 60 NESB patients. Forty per cent of the patients were predicted to be from NESB countries (based on demographic data from our outpatient clinic). A minimum of 35 patients with normal cognition (not including those with age-associated cognitive decline) were required. The sample has to be sufficiently large to evaluate the independent contribution of approximately 15 items in a logistic regression model, with 8–12 subjects recommended for each item (Hosmer and Lemeshow, 2000).

DATA ANALYSIS

We measured the strength of the association between each of the 42 items and the cognitive diagnosis using the Spearman's rank-order correlation coefficient. Items with correlation coefficients less than 0.35 were excluded from further analysis. Where several of the remaining items measured the same domain, the item with the strongest correlation was retained, provided that it had the best construct validity, cultural appropriateness and clinical utility as judged by unanimous agreement of the authors. Multivariate logistic regression modeling was used to select the set of items that best predicted dementia, after adjustment for age.

Stage 3 – Validation of the RUDAS in a community setting

PROCEDURE

The predictive accuracy of the RUDAS was validated in 90 elderly, community-dwelling persons, selected at random from a large database of all patients referred to a large, multidisciplinary, community-based aged-care team between

1997 and 1999. Potential subjects were stratified into six groups according to language background (English-speaking background, Asian non-English speaking background, and non-Asian non-English speaking background) and cognitive diagnosis (demented and non-demented). Subjects from each cognitive group were then matched for age and sex, and invited by telephone to participate in the validation study. Invitations to participate continued until 15 participants were recruited to each of the six groups.

The research psychologist administered the RUDAS at the participant's home. Within several days, a geriatrician (JTR), blinded to the results of the RUDAS scale, assessed each subject at home or in the outpatient department. Both clinicians were blinded to previous cognitive diagnosis. To evaluate inter-rater reliability, each subject was assessed at home by both the research psychologist and a trained member of the aged care team, by coin toss. Approximately one week later, the psychologist administered the RUDAS again to measure test-retest reliability. Professional language interpreters were used where necessary (over the telephone and face-to-face) and written consent was obtained from all participants or their proxies.

SAMPLE SIZE

The sample size for Stage 3 was based primarily on showing good inter-rater and test-retest reliability in randomly selected community-dwelling persons. At least 65 subjects were required for an estimated correlation coefficient of 0.8 (with 95% CI $\pm$ 15%). We recruited at least 90 subjects to allow preliminary evaluation of the RUDAS in each of the language background subgroups.

DATA ANALYSIS

Logistic regression analysis was used to assess the independent contribution of the RUDAS components in predicting dementia. Frequency response analysis and receiver-operating characteristic (ROC) curve analysis were used to determine the extent to which item responses segregated normal and demented patients. As the items used different measurement scales, standardized regression coefficients were calculated to compare the strength of the relationships between the items and the cognitive diagnosis. The standardized coefficients were used to assign scores to each of the items, which were then added together to yield a total score for the RUDAS. The overall accuracy of the RUDAS and the optimal cut-point for the diagnosis of dementia were assessed using ROC curve analysis (Hanley and McNeil, 1982). Multivariate logistic regression modeling was used to assess the effect of age, gender, education, preferred language and other factors which may differentially affect performance, on the relation between the RUDAS scores and cognitive diagnoses. Inter-rater and test-retest reliability of the RUDAS was measured using the intraclass correlation coefficient (ICC) (Shrout

Table 3. Characteristics of 166 patients evaluated in the geriatric medicine clinic

CHARACTERISTIC	MEASURE
Age in years (mean ± SD)	77.9 ± 7.0
Female (%)	68.7
Years in Australia (median, Q1–Q3)*	52.5, 40–76
Preferred language other than English (%)	34.9
Interpreter used (%)	29.5
Years of education (median, Q1–Q3)†	9, 6–10
MBI score (median, Q1–Q3)	19, 17–20
Lawton IADL score (median, Q1–Q3)	5, 3–7
GDS score (median, Q1–Q3)	5, 3–8
CDR‡	
No dementia (%)	21.7
Questionable dementia (%)	30.7
Mild dementia (%)	22.3
Moderate dementia (%)	19.3
Severe dementia (%)	4.8

Note. SD = standard deviation; Q1–Q3 = interquartile range; MBI = Modified Barthel Index; Lawton IADL = Lawton Instrumental Activities of Daily Living Scale; GDS = 15-point Geriatric Depression Scale; CDR = Clinical Dementia Rating Scale.
* 84 (50.6%) patients were born in ESB countries; 61 (36.7%) were born in one of 18 European NESB countries; 13 (7.8%) were born in one of 7 Asian countries; 4 (2.4%) patients were born in Central or South America, 3 (1.8%) in Africa, and 1 (0.6%) in the Middle East.
† Education unknown for one patient.
‡ CDR unknown for two patients.

and Fleiss, 1979). Differences between patients were tested using *t* tests for continuous, normally distributed variables, chi-square tests for dichotomous variables, and Wilcoxon rank-sum tests for ordinal variables. All statistical analyses were performed using the SPSS 11 statistical software system (SPSS Inc., 2001).

Results

Stage 2 – Clinical testing of items

PATIENT CHARACTERISTICS

The 42 items endorsed for clinical testing were evaluated in 166 consecutive new referrals to the geriatric medicine outpatient department (Table 3). The patients were elderly (mean age of 77.9 years) and moderately well educated (median of 9 years of education). Similar proportions of patients were born in English-speaking background (ESB) and non-English-speaking background (NESB) countries. Twenty-five language groups were represented.

Table 4. Item-diagnosis correlation coefficients

DOMAIN	ITEM	SPEARMAN'S COEFFICIENT
Memory	4-item grocery recall	0.501
Language	Animal generation	0.487
Memory	4-item animal recall	0.472
Memory	Children-fire story recall	0.461
Language	Food generation	0.422
Insight	Assessor-rated insight	0.418
Praxis	Fist-palm alternation task	0.417
Visuo-spatial (drawing)	Draw clock (Shulman)*	−0.393
Judgement	Crossing the road	0.388
Language	Naturalistic conversation on childhood games	0.387
Visuo-spatial (drawing)	Cube copying	0.379
Planning	Simple maze	0.377
Gnosis	Body orientation	0.374
Visuo-spatial (drawing)	Copy lemniscates	0.360
Judgement	Locked-out	0.354
Memory	Boy-big-dog sentence recall	0.354
Visuo-spatial (drawing)	Draw clock (Mendez)†	0.351

* Shulman clock scoring method {Shulman, 1986 #376; Shulman, 1993 #36}.
† Mendez clock scoring method {Mendez, 1992 #41}.

ITEM-DIAGNOSIS CORRELATIONS

Table 4 shows the Spearman's correlation coefficients between each of the items tested and the cognitive diagnosis. Items with correlation coefficients less than 0.35 were excluded from further analysis, and are not shown in Table 4. For simplicity, only the primary cognitive domain is shown for each item, although many items clearly measure more than one domain.

LOGISTIC REGRESSION

Table 5 shows the multivariate logistic regression model that best predicted dementia, after adjustment for age. Items significantly associated with a diagnosis of dementia included "4-item grocery recall", "crossing the road", "cube copying" and animal generation". Although "fist-palm alternation task" ($p = 0.065$) and "body orientation" ($p = 0.094$) failed to reach statistical significance, both were retained, as no other item (with a Spearman's coefficient of at least 0.35) measured praxis or gnosis. Age ($p = 0.17$) was not a significant predictor of dementia (after adjustment).

RUDAS ITEMS

The research psychologist administered the six RUDAS items. All items were administered sitting opposite the patient to control for level of difficulty. The

Table 5. Logistic regression model: likelihood of dementia ($N = 152$)

VARIABLE	PARAMETER ESTIMATE	STANDARD ERROR	*P* VALUE	ODDS RATIO (95% CI)
4-item grocery recall	−0.97	0.31	0.002	0.38 (0.21−0.69)
Animal generation	−0.12	0.063	0.051	0.88 (0.78−1.00)
Fist-palm alternation task			0.065	
Fist-palm (1)*	3.55	1.53	0.020	34.83 (1.75−694.66)
Fist-palm (2)*	0.38	0.65	0.56	1.46 (0.41−5.18)
Crossing the road	−0.88	0.36	0.016	0.42 (0.20−0.85)
Cube copying	−0.50	0.24	0.039	0.61 (0.38−0.98)
Body orientation	−0.58	0.35	0.094	0.56 (0.29−1.10)
Age	−0.059	0.043	0.17	0.94 (0.87−1.03)

* These refer to dummy variables for the fist-palm alternation task.

order of items in the RUDAS validation instrument was determined on the basis of item characteristics and minimization of those factors which would be likely to increase test anxiety and make the questions confrontational and threatening for respondents. The final order of items was Memory (4-item grocery list), Gnosis (Body orientation), Praxis (Fist/palm alternating task), Visuo-spatial drawing (Cube copying), Judgement (Crossing the road), Memory Recall (Grocery list recall), and Language (Animal generation).

4-item grocery recall: The patient was required to remember four grocery items (tea, cooking oil, eggs and soap), after a maximum of five learning trials (to ensure item registration). If the patient was unable to recall any of the groceries, "tea" was used as a prompt.

Body orientation: The patient was asked to respond to the following commands (in the order stated):

1. "Show me your right foot".
2. "Show me your left hand".
3. "With your right hand, touch your left shoulder".
4. "With your left hand, touch your right ear".
5. "Point to or indicate my left knee".
6. "Point to or indicate my right elbow".
7. "With your right hand, point to or indicate my left eye".
8. "With your left hand, point to or indicate my left foot".

Fist-palm alternation task: The patient was asked to imitate a motor task, beginning by placing both hands palm down on the table (or the lap). One hand was then placed in a fist (in the vertical position) while the other remained palm down. Both hands were then simultaneously alternated between the two positions. Having learned the task, the patient was asked to maintain it at a moderate (walking) pace for approximately 10 seconds.

Table 6. Characteristics of 90 validation study patients by cognitive diagnosis

CHARACTERISTIC	NORMAL ($n = 45$)	DEMENTED ($n = 45$)	p VALUE
Age in years (mean ± SD)	78.1 ± 8.4	81.4 ± 8.1	.082
Female (%)	82.2	73.3	.310
Years in Australia (median, Q1–Q3)*	42.0, 19.5–53.5	30.0, 18.5–69.0	.728
Preferred language other than English (%)	51.1	66.7	.134
Interpreter used (%)	44.4	66.7	.034
More than 6 years of education (%)†	51.1	35.0	.135
MBI score (median, Q1–Q3)	20.0, 18.0–20.0	11.0, 4.0–16.0	.0001
Lawton IADL score (median, Q1–Q3)	6.0, 3.5–8.0	0.0, 0.0–1.0	.0001
Factor potentially affecting performance (%)‡	42.2	52.3	.342
CDR			.0001
No dementia (%)	91.1		
Questionable dementia (%)	8.9	8.9	
Mild dementia (%)		20.0	
Moderate dementia (%)		31.1	
Severe dementia (%)		40.0	

Note. SD = standard deviation; Q1–Q3 = interquartile range; MBI = Modified Barthel Index; Lawton IADL = Lawton Instrumental Activities of Daily Living Scale; GDS = 15-point Geriatric Depression Scale; CDR = Clinical Dementia Rating scale.

* 30 (33.3%) patients were born in ESB countries; 20 (22.2%) were born in one of 11 European NESB countries; 30 (33.3%) were born in one of 12 Asian countries; 2 (2.2%) patients were born in Central or South America, 3 (3.3%) in Africa, and 5 (5.5%) in the Middle East.

† Education unknown for 5 subjects.

‡ All subjects were assessed for vision and hearing impairment, psychiatric disease (including depression), dysarthria and dysphasia, and musculoskeletal, and neurological diseases than might affect performance on the RUDAS.

Cube copying: This is a drawing task that requires the patient to copy a large line drawing of a cube.

Crossing the road: The patient is asked to describe how he or she would go about safely crossing a very busy street or similar thoroughfare where there is no pedestrian crossing or traffic lights. If the patient did not address two necessary components (looking for traffic and safety), a prompt – "is there anything else you would do?" was used.

Animal generation: The patient is asked to name as many new animals as possible in one minute.

Stage 3 – Validation of the RUDAS in a community setting

PATIENT CHARACTERISTICS

The 6-item RUDAS scale was validated in 90 elderly, community-dwelling persons. More than half of subjects had six or less years of education. Table 6 shows the characteristics of the participants by cognitive diagnosis (demented or non-demented). Those with dementia were similar to normal patients,

except for MBI score ($p = .0001$) and Lawton IADL score ($p = .0001$), functional measures known to correlate strongly with dementia. By default, NESB participants with dementia were assessed using a professional language interpreter.

NON-WEIGHTED ITEM SCORING

4-item grocery recall: Each correct response scored one point (maximum of four points). If a prompt was used, the maximum score was three. Registration of grocery items was not scored.

Body orientation: Response frequencies showed that normal participants were able to complete at least five instructions without difficulty. No subject with dementia performed more than four correctly, therefore it was decided that commands should continue only until five were completed correctly (for a maximum score of five).

Cube copying: Initially, one point for each of five "cube characteristics" was assigned to the patient's drawing (3-dimensional drawing, based on a square, all internal lines drawn, all external lines drawn, perfect copy). However, on the basis of multivariate logistic regression, "3-dimensional drawing" and "perfect copy" were not significant predictors of cognitive diagnosis. The scoring method was simplified, allocating one point for each of the remaining three characteristics.

Crossing the road: Each of the two necessary components (looking for traffic and safety) is scored out of two points. Any component which is prompted scores a maximum of one point.

Animal generation: Based on ROC curve analysis, the ability to name eight animals separated most normal patients from those with dementia. It was therefore decided that once eight animals were named, this item could be concluded without testing for the entire one minute. There was a maximum score of eight points for the number of new animals named in one minute.

WEIGHTED ITEM SCORING

The standardized logistic regression coefficients used to compare the strength of the relationships between the items and the cognitive diagnosis, and the item scores, are shown in Table 7. Items with the largest standardized coefficients (most strongly associated with the cognitive diagnosis) were weighted and scored accordingly, for a total RUDAS score of 30.

RELIABILITY OF THE RUDAS

Both the inter-rater (ICC = 0.99) and the test-retest (ICC = 0.98) reliabilities of the RUDAS were very high.

Table 7. Item weights and scores using standardized coefficients

ITEM	STANDARDIZED COEFFICIENT (STANDARD ERROR)	ITEM SCORE
4-item grocery recall	1.8 (0.28)	8
Animal generation	1.6 (0.25)	8
Body orientation	1.3 (0.20)	5
Crossing the road	1.0 (0.16)	4
Fist-palm alternation task	0.5 (0.08)	2
Cube copying	0.2 (0.03)	3
Total		30

DIAGNOSTIC ACCURACY OF THE RUDAS

In our population, the diagnostic accuracy of the RUDAS (for detecting dementia based on DSM-IV criteria) was excellent, with an area under the ROC curve of 0.95 (95% CI 0.88–0.98) (Figure 1). Scores lower than 23 (the optimal cut-point based on the ROC curve) detected dementia with a sensitivity of 89% (95% CI 76%–96%) and a specificity of 98% (95% CI 88%–97%).

EFFECT OF GENDER, EDUCATION, PREFERRED LANGUAGE, AGE AND PERFORMANCE FACTORS

In the validation sample, gender ($p = 0.18$), years of education ($p = 0.20$), preferred language (English or otherwise) ($p = 0.33$), and factors which may

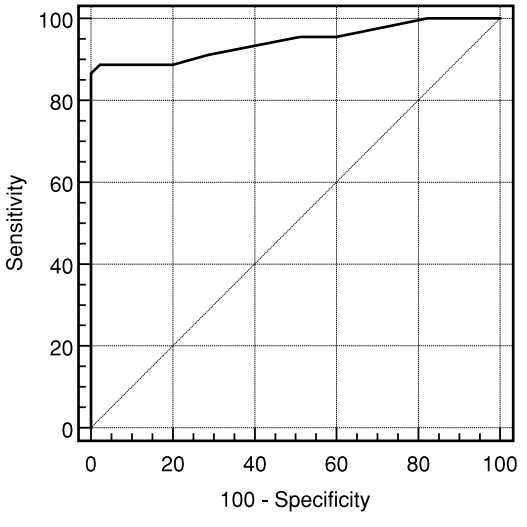

Figure 1. ROC Curve for RUDAS

affect performance on a cognitive test ($p = 0.42$) were not independent predictors of dementia. Age ($p = 0.04$) was a significant predictor in the same model.

Discussion

This study is the first to develop a multicultural cognitive test in a culturally heterogeneous sample of elderly patients and addresses the gap in previous work that has focused specifically on the development of cognitive screening tests for use within or across discrete cultural groups. Most previous attempts to develop instruments for use in culturally homogenous or heterogeneous populations modified or translated scales (or their components) that were originally developed in dissimilar groups.

The 6-item RUDAS is portable and requires no special equipment. It can be administered by health care workers after approximately 40 minutes of training (using videotape). Based on our experience, the RUDAS can be directly translated into at least 30 other languages, without the need to change the structure or the format of any item. Our study populations and advisory groups were culturally diverse, allowing us to develop items we believe to be culturally fair, thus avoiding the common approach of superimposing "western" ideas on people from other backgrounds.

The RUDAS takes about 10 minutes to complete and tests multiple cognitive domains. In particular, items relevant to frontal lobe function, such as "crossing the road", "animal generation", and "cube copying", evaluate executive functioning (the ability to initiate, plan and execute tasks relevant to daily living), both directly and indirectly. Lack of evaluation of executive function is a major limitation of the MMSE, which usually fails to detect meaningful deficits involving the frontal lobes. The diverse response formats of the RUDAS (verbal, non-verbal, written and praxic) allow more comprehensive assessment of a subject's overall cognitive ability. Impairment in a domain necessary to communicate a response (but not necessarily causing important cognitive deficits) is not over-emphasized, thereby reducing misclassification of cognitive capacity.

Gender, years of education, preferred language (English or otherwise), and differential performance factors were not significant predictors of dementia in the multivariate logistic model. This represents a major advantage over several earlier instruments, where education, in particular, exerted a substantial influence on instrument scores. Age was a significant predictor in our validation sample ($p = 0.035$), but not in the development sample ($p = 0.170$). The effect of age therefore remains uncertain and needs further evaluation. We computed standardized logistic regression coefficients to compare the strength of the relationships between the items and the cognitive diagnosis. Items most strongly associated with the cognitive diagnosis were weighted and scored accordingly.

In our validation sample, the RUDAS had excellent reliability (both inter-rater and test-retest) and diagnostic accuracy for detecting dementia based on DSM-IV criteria. Importantly, the RUDAS does not depend on informant history. Demented persons without informants are less likely to have carers, and are therefore at higher risk of the adverse consequences of dementia. The RUDAS items can be directly translated and are relevant to most cultures, strengths that should facilitate the use of the instrument in cross-cultural research.

Our study has several limitations. First we cannot exclude the possibility that we misclassified participants as demented or non-demented. Accurate diagnosis can be difficult, particularly in culturally heterogeneous populations (Erkinjuntti *et al.*, 1997; Prince, 2000). While there is no gold standard for dementia diagnosis, decisions based on DSM-IV criteria have the advantage of broad acceptance and good reproducibility (Baldereschi *et al.*, 1994; O'Connor *et al.*, 1996). Furthermore, each participant was comprehensively assessed by one of three geriatricians, who used a variety of information sources to make a diagnosis. The development sample was not random, and used patients attending a geriatric outpatient clinic. The possibility of bias within the sample is therefore far stronger than desirable. However, our validation sample had an even distribution of participants with normal cognition and those with dementia (50% in each group). Re-validation of the RUDAS in further community-based populations is desirable.

In summary, the 6-item RUDAS was developed and validated in culturally diverse samples. It is portable and easy to administer, with excellent reliability and diagnostic accuracy. It appears not to be affected by gender, years of education, differential performance factors and preferred language. It tests multiple cognitive domains, including those related to function of the frontal lobes. The RUDAS can be simply translated into other languages, without the need to change the structure or the format of any item. While designed primarily to detect dementia in the primary care setting, further work is needed in other settings, and in longitudinal studies to determine its sensitivity to change in cognitive function over time.

Acknowledgement

The authors would like to acknowledge and express their thanks to the Health Department of the Government of New South Wales, Australia for their funding and support of this project.

References

American Psychiatric Association (1994). DSM-IV: *Diagnostic and Statistical Manual of Mental Disorders*. (4th ed.). Washington, DC: American Psychiatric Association.

Baldereschi, M., Amato, M. P., Nencini, P., Pracucci, G., *et al.* (1994). Cross-national interrater agreement on the clinical diagnostic criteria for dementia. *Neurology*, 44, 239–242.

Berg, L. (1988). Clinical Dementia Rating (CDR). *Psychopharmacology Bulletin*, 24, 637–639.

Bernatzik, H. (1957). *Razas y Pueblos del Mundo. (World Races and People)*. Barcelone: Ediciones Ave.

Berry, J. (1997). Immigration, acculturation and adaptation. *Applied Psychology: An International Review*, 46, 5–33.

Berry, J., Poortinga, Y., Segall, M., and Dasen, P. (1992). *Cross-cultural Psychology: Research and Applications*. Cambridge: Cambridge University Press.

Borson, S., Scanlan, J., Brush, M., Vitaliano, P., and Dokmak, A. (2000). The mini-cog: a cognitive 'vital signs' measure for dementia screening in multi-lingual elderly. *International Journal of Geriatric Psychiatry*, 15, 1021–1027.

Bourhis, R., Giles, H., and Rosenthal, D. (1981). Notes on the construction of a "subjective vitality questionnaire" for ethnolinguistic groups. *Journal of Multicultural and Multilingual Development*, 2, 145–155.

Brodaty, H., and Moore, C. M. (1997). The Clock Drawing Test for dementia of the Alzheimer's type: A comparison of three scoring methods in a memory disorders *clinic*. *International Journal of Geriatric Psychiatry*, 12, 619–627.

Copeland, J. R., Kelleher, M. J., Kellett, J. M., Gourlay, A. J., Gurland, B. J., *et al.* (1976). A semi-structured clinical interview for the assessment of diagnosis and mental state in the elderly: the Geriatric Mental State Schedule. I. Development and reliability. *Psychological Medicine*, 6, 439–449.

Erkinjuntti, T., Ostbye, T., Steenhuis, R., and Hachinski, V. (1997). The effect of different diagnostic criteria on the prevalence of dementia. *New England Journal of Medicine*, 337, 1667–1674.

Escobar, J. I., Burnam, A., Karno, M., Forsythe, A., Landsverk, J., *et al.* (1986). Use of the Mini-Mental State Examination (MMSE) in a community population of mixed ethnicity. Cultural and linguistic artefacts. *Journal of Nervous and Mental Disease*, 174, 607–614.

Folstein, M., Folstein, S., and McHugh, P. (1975). Mini-mental state: A practical method for grading the mental state of patients for the clinician. *Journal of Psychiatry Research*, 12, 189–198.

Fuld, P. A., Muramoto, O., Blau, A., Westbrook, L., and Katzman, R. (1988). Cross-cultural and multi-ethnic dementia evaluation by mental status and memory testing. *Cortex*, 24, 511–519.

Glosser, G., Wolfe, N., Albert, M. L., Lavine, L., *et al.* (1993). Cross-cultural cognitive examination: validation of a dementia screening instrument for neuroepidemiological research. *Journal of the American Geriatrics Society*, 41, 931–939.

Graves, A. B., Larson, E. B., Kukull, W. A., White, L. R., and Teng, E. L. (1993). Screening for dementia in the community in cross-national studies: comparison between the Cognitive Abilities Screening Instrument and the Mini-Mental State Examination. In *Alzheimer's Disease: Advances in Clinical and Basic Research.*(Corain, B., Iqbal, K. M., Nicolini, *et al.*, eds.), (pp. 113–119). Chichester: John Wiley & Sons.

Hall, K. S., Gao, S., Emsley, C. L., Ogunniyi, A. O., Morgan, O., *et al.* (2000). Community screening interview for dementia (CSI 'D'); performance in five disparate study sites. *International Journal of Geriatric Psychiatry*, 15, 521–531.

Hall, K. S., Hendrie, H. C., Brittain, H. M., Norton, J. A., *et al.* (1993). The development of a dementia screening interview in two distinct languages. *International Journal of Methods in Psychiatric Research*, 3, 1–28.

Hanley, J., and McNeil, B. (1982). The meaning and use of the area under a Receiver Operating Characteristic (ROC) Curve. *Radiology*, 143, 29–36.

Harwood, J., Giles, H., and Bourhis, R. (1994). The genesis of vitality theory: Historical patterns and discoursal dimensions. *International Journal of the Sociology of Language*, 108, 167–206.

Hasegawa, K. (1983). The clinical assessment of dementia in the aged: A dementia screening scale for psychogeriatric patients. In M. Bergener, U. Lehr, E. Lang *et al.* (Eds.) *Aging in the Eighties and Beyond: Highlights of the Twelfth International Congress of Gerontology*. New York: Springer Publishing Company.

Hosmer, D. W., and Lemeshow, S. (2000). *Applied Logistic Regression*. (2nd ed.): John Wiley & Sons, Inc.

Kim, U. (2000). Indigenous, cultural, and cross-cultural psychology: A theoretical, conceptual, and epistemological analysis. *Asian Journal of Social Psychology*, 3, 265–287.

Kua, E. H., and Ko, S. M. (1992). A questionnaire to screen for cognitive impairment among elderly people in developing countries. *Acta Psychiatrica Scandinavica*, 85, 119–122.

Lawton, M. P., and Brody, E. M. (1969). Assessment of older people: self-maintaining and instrumental activities of daily living. *Gerontologist*, 9, 179–186.

Martin, S. F. (2001). Global migration trends and asylum. *The Journal of Humanitarian Assistance*, 23.

Mast, B. T., Fitzgerald, J., Steinberg, J., MacNeill, S. E., and Lichtenberg, P. A. (2001). Effective screening for Alzheimer's disease among older African Americans. *Clinical Neuropsychologist*, 15, 196–202.

O'Connor, D.W., Blessed, G., Cooper, B., Jonker, C., Morris, J. C., *et al.* (1996). Cross-national interrater reliability of dementia diagnosis in the elderly and factors associated with disagreement. *Neurology*, 47, 1194–1199.

Prince, M. (2000). Methodological issues for population-based research into dementia in developing countries. A position paper from the 10/66 Dementia Research Group. *International Journal of Geriatric Psychiatry*, 15, 21–30.

Royall, D. R., Mahurin, R. K., and Cornell, J. (1994). Bedside assessment of frontal degeneration: distinguishing Alzheimer's disease from non-Alzheimer's cortical dementia. *Experimental Aging Research*, 20, 95–103.

Sandson, T., and Price, B. (1996). Diagnostic testing and dementia. *Neurologic Clinics*, 14, 45–59.

Shadlen, M. F., Larson, E. B., Gibbons, L. E., Rice, M. M., McCormick, W. C., *et al.* (2001). Ethnicity and cognitive performance among older African Americans, Japanese Americans, and Caucasians: the role of education. *Journal of the American Geriatrics Society*, 49, 1371–1378.

Shrout, P., and Fleiss, J. (1979). Intraclass correlation: Uses in assessing rater reliability. *Psychological Bulletin*, 86, 420.

Small, G. W. (1998). Differential diagnosis and early detection of dementia. *American Journal of Geriatric Psychiatry*, 6, S26–S33.

SPSS Inc. (2001). SPSS for Windows Release 11.0.1 (15 Nov 2001). Chicago: SPSS Inc.

Storey, J. E., Rowland, J. T. J., Basic, D., and Conforti, D. (2002). Accuracy of the clock drawing test for detecting dementia in a multicultural sample of elderly Australian patients. *International Psychogeriatrics*, 14, 259– 272.

Swadesh, M. (1967). *El Lenguaje y la Vida Humana. (Language and Human Life)*. Mexico: Fondo de Cultural Economica.

Teng, E. L., and Chui, H. C. (1987). The Modified Mini-Mental State (3MS) examination. *Journal of Clinical Psychiatry*, 48, 314–318.

Teng, E. L., Hasegawa, K., Homma, A., Imai, Y., Larson, E. B., *et al.* (1992). A practical test for cross-cultural epidemiological studies of dementia: the Cognitive Abilities Screening Instrument (CASI). In H. Orimo, Y. Fukuchi, K. Kuramoto *et al.* (Eds.) *New Horizons in Aging Science*. (pp. 326–327). Tokyo: University of Tokyo Press.

United States Department of Commerce, Economics and Statistics Administration, and Bureau of the Census (1999). *International brief: world population at a glance: 1998 and beyond*. (pp. 4). United States Department of Commerce, Economics and Statistics Administration, Bureau of the Census.

Wade, D. T., and Collin, C. (1988). The Barthel ADL Index: a standard measure of physical disability? *International Disability Studies*, 10, 64–67.

Yeasavage, J., Brink, T., Rose, T., Lum, O., Huang, V., *et al.* (1983). Development and validation of a geriatric depression scale: A preliminary report. *Journal of Psychiatry Research*, 17, 37–49.