

# A HINDI VERSION OF THE MMSE: THE DEVELOPMENT OF A COGNITIVE SCREENING INSTRUMENT FOR A LARGELY ILLITERATE RURAL ELDERLY POPULATION IN INDIA

MARY GANGULI\* AND GRAHAM RATCLIFF†

\*Assistant Professor of Psychiatry and Epidemiology, †Adjunct Assistant Professor of Psychiatry and Neurology, University of Pittsburgh, USA; ‡Neuropsychologist, Harmarville Rehabilitation Center, Pittsburgh, PA, USA

VIJAY CHANDRA

Director, Centre for Ageing Research in India, New Delhi, India; Adjunct Assistant Professor of Epidemiology, University of Pittsburgh, USA; Consultant Neurologist, Holy Family Hospital, New Delhi, India

SUJATHA SHARMA

Neuropsychologist, Centre for Ageing Research in India, New Delhi, India

JOANNE GILBY

Senior Research Associate, Department of Psychiatry, University of Pittsburgh, USA

RAJESH PANDAV

Medical Officer, Centre for Ageing Research in India, New Delhi, India

STEVEN BELLE\*, CHRISTOPHER RYAN†, CAROL BAKER‡, ERIC SEABERG§ AND STEVEN DEKOSKY\*\*

\*Assistant Professor of Epidemiology and Nursing, †Associate Professor of Psychiatry and Psychology, ‡Assistant Professor of Education and Director, Office of Measurement and Evaluation, §Statistician, Epidemiology Data Center,

\*\*Professor of Psychiatry and Neurology, University of Pittsburgh, USA

## SUMMARY

The Indo-US Cross-National Dementia Epidemiology Study seeks to compare two rural populations, in the US and India: the Monongahela Valley, a rural community of relatively low socioeconomic status in southwestern Pennsylvania (USA), and Ballabgarh, a rural community near New Delhi in North India. Of particular interest is the fact that the Ballabgarh elderly population is exclusively Hindi-speaking, has little or no education and is largely illiterate, rendering its cognitive screening a particular challenge. In this article we report methods and preliminary data on the development of a Hindi cognitive screening instrument suitable for the Ballabgarh elderly population. We use as an example our Hindi adaptation of the Mini-Mental State Examination (MMSE), a widely used global cognitive screening scale. Systematic, item-by-item, empirically based test development has shown that effective modifications can be made to existing tests that require reading and writing; and that culturally sensitive modifications can be made to render the test meaningful and relevant while still tapping the appropriate cognitive domains. Certain cognitive functions, particularly orientation to time, remain difficult to test adequately in this type of population. In Ballabgarh, as in the Monongahela Valley, educated individuals obtain higher test scores. Implications for dementia screening are discussed, including those relevant to the hypothesis that low education predisposes to dementia.

KEY WORDS—dementia epidemiology; ageing; neuropsychology; cross-cultural research

Cognitive impairment, characteristic of dementia, is measured objectively by standard neuropsychological (cognitive) tests. Screening for dementia is usually accomplished by means of a global cognitive

scale, such as the widely used Mini-Mental State Exam (MMSE) (Folstein *et al.*, 1975; Tombaugh and McIntyre, 1992). This scale draws its strength from the fact that it briefly taps a variety of cognitive functions known to be impaired in dementia. The MMSE, variously modified, translated into several languages, for example Chinese and Finnish (Salmon *et al.*, 1989), Korean (Park and Kwon, 1990), Japanese (Larson *et al.*, 1992), Spanish

Address for correspondence: Dr Mary Ganguli, Division of Geriatrics and Neuropsychiatry, Department of Psychiatry, University of Pittsburgh School of Medicine, Western Psychiatric Institute and Clinic, 3811 O'Hara Street, Pittsburgh, PA 15213-2593, USA. Tel: (412) 647-2078. Fax: (412) 647-5422.

Loewenstein *et al.*, 1993), Yoruba (Hendrie *et al.*, 1992) and Cree (Hall *et al.*, 1993), and sometimes incorporated into larger test batteries or interview schedules, has been used in several independent cross-national studies of dementia epidemiology. It is uncertain, however, whether or not the many studies using this instrument have employed similar approaches to its translation and cultural adaptation for their populations. Thus, the comparability of the results of these studies with one another, and with those of studies using the English MMSE, can be called into question.

In this article, we provide a detailed report of the methods and preliminary results of a systematic and empirical approach we have employed in developing a Hindi version of the MMSE. The resulting instrument is being used in an epidemiological investigation of dementia in a largely illiterate rural elderly population in India and for its comparison with a rural elderly population in the USA. While our modification of the MMSE will not be equally appropriate for screening another population, our methods of test development should be useful to others conducting similar studies, particularly those involving comparisons between diverse populations.

#### STUDY POPULATIONS

The Indo-US Cross-National Dementia Epidemiology Study represents a collaboration between (a) the University of Pittsburgh School of Medicine and Graduate School of Public Health, Pittsburgh, PA (USA) and (b) the Centre for Ageing Research in India (CARI), with the cooperation of the Centre for Community Medicine (CCM) of the All-India Institute of Medical Sciences (AIIMS), New Delhi, India. Comparisons are being made between elderly populations at two sites:

1. *A rural US site, the Monongahela Valley.* The Monongahela Valley Independent Elders Survey (MoVIES) project is an ongoing prospective study of dementia epidemiology in a rural southwestern Pennsylvanian blue-collar community approximately 25 miles south of Pittsburgh. At baseline, the study population was an age-stratified random population sample of 1366 English-speaking, community-dwelling, individuals aged 65 years or older with a median educational level of high school graduate. Further details have been reported elsewhere (Ganguli *et al.*, 1991, 1993a,b). Since the MoVIES project had already

been in progress for some years when the cross-national study was funded, methods for the study in India were based on appropriate modifications of MoVIES methods.

2. *A rural Indian site, Ballabgarh.* The community of Ballabgarh, 22 miles from New Delhi, India, is the intensive field practice site for the CCM. Its census and health status have been continuously monitored for over 25 years by community health workers, for public health service and research purposes, through the Comprehensive Rural Health Services Project of AIIMS. The language spoken is the Haryanvi dialect of Hindi, an Indo-European language written in Devanagiri script and pronounced phonetically. In accordance with the lower life expectancy of the Indian population, we selected age 55 as our lower bound, as has been done by other investigators in developing countries (Zhang *et al.*, 1990). The majority of elderly individuals in this community have had no formal education and are illiterate.

Our overall goals are to determine the prevalence and incidence of, and risk factors for, dementia in the Ballabgarh population; to characterize normal and abnormal cognition in this population; and to make cross-national comparisons of these data with those obtained in the MoVIES project. The project's first objective was to develop for the elderly Ballabgarh population, systematically and empirically, a suitable cognitive and functional screening instrument which (a) was culturally fair, psychometrically sound and valid for a population with no education; (b) was optimally sensitive and specific for dementia; (c) would allow not only the identification but also the more detailed characterization of dementia, and of normal and abnormal cognitive ageing, in this cohort; and (d) would allow meaningful cross-national comparisons with MoVIES project data. Accordingly, we required the instruments used at the two sites to be analogous, ie to the extent possible, tap similar cognitive domains with tests of similar relative difficulty, presented in similar format.

The MoVIES project employs a cognitive screening battery which includes a global scale, the MMSE, as well as other tests tapping a variety of cognitive domains known to be affected in the dementing disorders (Ganguli *et al.*, 1991, 1993a). In this report, we focus on the translation and adaptation of the MMSE as an illustration of our general method of cognitive instrument development for

the Ballabgarh elderly population. As our adaptation represents certain major changes to the MMSE, we have named it the Hindi Mental State Examination (HMSE) to distinguish it from the original.

## METHODS

### *Rationale*

If our goal in developing a Hindi version of the MMSE, and of the other tests in the MoVIES battery, had been solely to identify those subjects in Ballabgarh most likely to be currently demented, the instrument would have needed to be reliable, valid, sensitive and specific, but not necessarily analogous to that used in the MoVIES study. However, if the scales are made analogous, and substantive differences are then found between the Indian and American populations, we would be on stronger ground in attributing those differences to population characteristics rather than to testing artifacts. Second, if cognitive profiles of the two samples (both demented and non-demented subjects) are to be compared, we would have to know whether the tests were equivalent. Third, to be able to use cognitive findings from the MoVIES study to test hypotheses in the Ballabgarh sample, the tests would have to be similar. This approach assumes that the 'normal' Ballabgarh subjects' difficulty with certain cognitive tasks is attributable to lack of practice, lack of education, etc, and not due to any essential difference in brain function; this assumption appeared to us to be safer, at this stage, than any of the alternatives.

### *Phases of instrument development*

1. *Initial selection, by consensus, of potential test items.* These consisted of Hindi versions of the items on the MMSE as well as the other MoVIES cognitive tests.

EXPERT CONSENSUS. A 'Delphi panel' of investigators and consultants (including bilingual clinicians) from Pittsburgh and New Delhi developed, explicitly for the Ballabgarh population, a series of measures which seemed to include the same general cognitive domains as, and were as similar as possible to, MMSE items as employed in the MoVIES project. We took into account data from a MoVIES non-demented subsample (described below) and our collective clinical and research experience. We considered each subtest and item in the MMSE

separately and prepared one, or typically more, putatively analogous Hindi items with due regard to their appropriateness in the rural North Indian culture. Each item and subtest was examined for cultural relevance, translatability and adaptability, without deviating from the conceptual basis of the test and the cognitive domain(s) being tapped by the test. We considered whether each item was appropriate and workable as far as possible in an analogous form; and if not, whether a different test or different format should be suggested. Our major objective at this stage was to determine which subtests could simply be translated; which would require only minor modification for cultural appropriateness and local acceptability; which would require major modification; and which subtests appeared unusable and would require substitution. We also selected for pre- and pilot testing more items than we would eventually include, anticipating that some would be eliminated during the process of instrument development.

MOVIES NON-DEMENTED (ND) COMPARISON GROUP. An age- and education-stratified subsample of 90 subjects, non-demented at baseline, was drawn from the MoVIES cohort. This cohort was by definition unavoidably older than the Ballabgarh population, its lower bound having been 65 years at study entry; further, the youngest MoVIES subgroup was the most highly educated. Thus, we selected the least educated ( $\leq$  9th grade) subgroup as our reference sample for developing a version for Ballabgarh elderly. The distribution of their scores on each MMSE subtest, and item by age, gender and education, was examined, item by item, to determine the *relative difficulty* of each item.

Our goal was to identify alternative versions of (and if necessary alternative tests for) the MMSE subtests for use in Ballabgarh, such that they would be of *comparable difficulty* for the Indian population, for example that the test scores would have comparable distributions in both populations, or that comparable proportions would obtain perfect scores on each tests, or the rank order of difficulty of subjects would be similar in the two populations. To accomplish this goal, we might have only to set a lower cutpoint in the final survey, or score more liberally, or actually set a simpler task, depending on the test. This objective of comparable difficulty was not, as it could seem, a self-fulfilling prophecy of setting screening cutpoints such that the same proportions would be classified as cognitively impaired or of ensuring similar prevalence rates in the two populations. Pretesting was conducted on *non-*

*demented volunteers*, so as to obtain normative data which would be compared to data from *non-demented members* of the MoVIES cohort, but not to establish cutpoints or criteria for cognitive impairment in the general elderly population.

A Hindi translation of the selected and modified tests at this stage was prepared in New Delhi. A Hindi-speaking group in Pittsburgh then back-translated the Hindi version into English, allowing clarification of remaining ambiguities. The resulting battery was administered to a small group of elderly rural volunteers and the experience used to fine-tune the items and instructions further.

Procedures for obtaining informed consent were approved by both the Human Volunteers Protection Committee of CARI and the Institutional Review Board for Psychosocial Research of the University of Pittsburgh.

**2. Pretesting.** Pretesting of these items was carried out on successive samples of 30 volunteer subjects each (designated as 30-1, 30-2, etc), aged 60 and older, in Ballabgarh, after obtaining informed consent. The objective was to examine the level of difficulty, acceptability, comprehensibility and relevance of the potential items and the distribution of scores on each subtest and item.

In 30-1, the goal for the tests was to remain as close as possible to that used in the MoVIES study, and to avoid as far as possible either ceiling or floor effects among the volunteers. In subsequent phases, the goal was to modify items that experience showed to be inappropriate. The distribution of data from these subjects was examined and compared to MoVIES *ND* data, allowing us to alter the level of difficulty such that the two normal populations performed at comparable levels. We were prepared to modify the tests as required and repeat this step as often as necessary; we were in fact able to develop a potential battery after two further iterations of this stage (30-2 and 30-3) were carried out.

**FIELD STAFF TRAINING AND OPERATIONS MANUAL.** Epidemiological investigations in developing countries often rely on field staff who themselves are from the study population and have minimal education but have been trained to carry out study procedures. It was therefore also critical to our study design that the tests be simple enough to be administered reliably by field workers with only partial high school education, unlike the college graduates who conduct cognitive testing in the MoVIES study. We required these personnel to deliver instructions and clarification to subjects according

to a script and to record responses verbatim; actual scoring would be carried out later by experienced clinicians. Therefore, a detailed operations manual with explicit instructions for field staff was prepared and modified on an ongoing basis both as the test itself was modified and also as unexpected responses were obtained from subjects. Particular attention was paid to listing allowable prompts and probes to be used if the subject did not respond or gave a non-specific or irrelevant response. The limits of permissible repetition, clarification and non-specific encouragement were discussed and specified and field workers required to record any 'help' that was given. Analyses of these 'help' data allowed us to identify tests which frequently required help and therefore might need reframing.

A final (30-4) pretesting phase was carried out as a training exercise for the field workers, who administered and recorded the tests under the direct supervision of the project neuropsychologist (SS). Interrater reliability of field workers, examined qualitatively with regard to test administration and data recording, was high; discrepancies were discussed and resolved by further amendments to the operations manual. Actual scoring of subjects' recorded responses, which requires some degree of clinical judgement, continues to be performed by the neuropsychologist. In 30-4, we attempted to measure interrater reliability of test scoring using a two-level design intended to maximize 'unreliability', so that areas of potential discrepancy could be identified and addressed. Each subject was tested by a field worker who recorded the subject's responses on a test form. The neuropsychologist observed the testing and independently completed a second form. The field worker's form was then scored by the neuropsychologist while the neuropsychologist's form was scored by the medical officer (RP). Agreement was then examined between the neuropsychologist's and medical officer's scores. In 24 (80%) out of 30 subjects, there was 100% agreement on total HMSE score (with the days backwards subtest of attention described below) and in the remaining six subjects there was a discrepancy of one point. These discrepancies were resolved by discussion and further guidelines were added to the operations manual.

**TEST MODIFICATION.** An item-by-item description follows.

**Orientation to time:** In the English MMSE, a point each is given for correctly named day, date, month, year and season, for a total of five points. In an early exploration of feasibility, we found that

many elderly Ballabgarh residents did not appear to know or keep track of years in either the Roman or the Hindu calendar. Rather than delete this question altogether and reduce the overall weighting of temporal orientation within the total HMSE score, we attempted to find a substitute question tapping large-scale temporal orientation. Alternatives such as naming the most recent major religious festival were tested but found inappropriate; we could obtain no consensus on what constituted a major festival and minor festivals were celebrated very often, leading almost every response to be scored as correct. For 30-3, we selected a closed-ended question: 'Is it morning, afternoon, or evening?', recognizing this taps a more limited scale of temporal orientation than the 'year'. Naming the 'season' also caused some difficulty because the equivalent Hindi word can also be interpreted as meaning 'weather'; when this occurs, we provide further clarification, specifying 'season of the year'. Thus, the five final items are: time of day, day of week, date, month and season.

*Orientation to place:* In variations on the English MMSE, five questions are asked, which could include 'name of this place/building', floor (storey), street address, city, county, state and country. It was difficult to identify five appropriate spatial orientation questions for the Ballabgarh sample. Buildings in the villages are all single-storeyed and do not have street numbers; streets do not have names. Enquiring into the official designations of state, district, area, village and block, we found that several of the elderly did not appear attuned to geographic location other than the name of the village and the majority could not name the state. We have finally selected 'district', 'post office' (postal district), 'village', 'block (or area or neighbourhood)' and either 'which place is this' or 'whose house is this', depending on whether the testing was being conducted in a home or in the health centre or other location in the village.

*Registration:* In the English MMSE, the names of three objects are given. We have used the Hindi words for 'mango, chair, coin', equivalent in familiarity to 'apple, table, penny' used in the MoVIES version. Tables are not ubiquitous pieces of furniture in the rural setting. The word '*paisa*', meaning coin, is also the term for the lowest denomination of Indian currency, thus equivalent to 'penny'. We found that although subjects could register these words initially, they appeared confused when asked to recall them later. We introduced a cueing device to counter this problem: the instruction now begins

'I have brought with me from Delhi the following three things: a mango, a chair, and a coin ...'.

*Attention:* The MMSE has two alternative attention subtests: (i) backwards spelling of WORLD and (ii) serial subtractions of seven starting at 100. In MoVIES, we administer both subtests and found (Ganguli *et al.*, 1990) only a weak association between scores on the two subtests, suggesting that they are not equivalent and should not be used interchangeably. The two tasks share the requirement that the subject be able to keep his attention on the cognitive problem for a period of time, maintain response set and have sufficient working memory capacity to hold and manipulate several pieces of information in mind while solving a multistep problem. They differ in that the former involves serial reordering of a presumably well-learned sequence, while the latter involves a presumably novel serial calculation. For Ballabgarh, we devised alternative forms of each subtest which attempted to retain their essential features including these similarities and differences.

- (a) **WORLD BACKWARDS.** Since most of the sample is illiterate, spelling (either forwards or backwards) is not an option. We ask subjects to name the days of the week backwards, starting from Sunday. Initially, responses were scored for the first 5 days named, not including Sunday. When we found subjects performed better if given an example, the instruction was changed to include the statement: 'For example, before Sunday comes Saturday; what comes before that?'. Credit is not given for naming Saturday or Sunday.
- (b) **SERIAL SEVENS.** The concept of abstract mental arithmetic was incomprehensible and the particular task of subtracting of sevens from 100 too difficult for our subjects. We gave the following subtraction task in the form of a story: '*A man has 20 rupees for bus fare. Every day, he spends 3 rupees on his bus fare. After spending the first day's bus fare, he will be left with 17 rupees. How much money will be left after the next day's bus fare ... and the next day's bus fare ...*'. The first five consecutive responses are scored.

*Recall:* The subject is asked to recall the three objects (mango, chair and coin) named earlier. Initially, there were difficulties because subjects did not understand which three objects were being referred to. However, when we introduced the cue

'What are the three objects I told you I brought with me from Delhi?', the task became much clearer and produced more correct responses.

*Naming:* As in the English version, the subject is shown a wristwatch and asked to name it. As a second object we use a pen rather than a pencil because the word used for pencil could mean either pencil or pen. Also, (ball-point) pens are the more familiar objects, being used more than pencils because they do not require sharpening. Our illiterate subjects typically used neither, and several of them gave incorrect responses to 'pen'. However, the majority of these responses were the Indian-English word 'holder', a corruption of the archaic 'pen-holder' (consisting of a stem with a slot for a removable nib which was dipped into ink for writing). Although this is technically an incorrect term for a ballpoint pen, it appears to be colloquially appropriate and not an example of dysnomia in this population; thus, we have now classified 'holder' as an acceptable alternative response.

*Repetition:* In English, the standard phrase for repetition is 'no ifs, ands or buts'. This familiar expression is also a test of fluency (Kertesz, 1982), being composed entirely of functor words such as prepositions and conjunctions being used as nouns, which patients with non-fluent aphasia find particularly difficult to repeat. In other translations of the MMSE, for example in Chinese (Salmon *et al.*, 1989) and Spanish (Perez *et al.*, 1991), we have seen substitution by phrases translated as '44 stone lions' and 'three dogs in a wheatfield', which are described as lexically equivalent but do not appear to us to serve the same cognitive testing function. After extensive exploration, we identified a meaningful phrase which translates as 'no, neither this nor that' and which consisted of five monosyllabic words as does the original and appeared to test the same aspects of fluency as 'no ifs, ands or buts' does in English. However, subjects appeared to be confused by the first word 'no', apparently perceiving it as part of the instructions. We then dropped the 'no' and settled on 'neither this nor that' as the phrase for repetition. One final problem was that, after performing this task, some subjects proceeded to repeat everything the examiner said. This prompted us to include a further instruction 'Now I am going to ask you a different kind of question' before proceeding to the next subtest.

*Visual command* (read and follow command): In the English version, the subject is shown a written command 'Close your eyes' and asked to do as it says. For our mostly illiterate subjects, the examiner says 'Look at me and do exactly what I do' and then closes his own eyes for 3 seconds (follow example),

while the coexaminer observes and records the subject's response.

*Three-step task:* As in the English version, the subject is asked to pick up (or take) a given piece of paper with his right hand, fold it in half with both hands and give the paper back to the examiner, with a point being given for each step remembered and correctly executed.

*Sentence:* This was an unexpectedly difficult task. In the English version, the subject is asked to write a complete sentence and is given a full point if the sentence has a subject and a predicate, regardless of spelling, grammar and syntax. We attempted to ask our illiterate subjects to 'say a sentence', as was also the instruction given in the Shanghai study (Zhang *et al.*, 1990). However, the standard Hindi word for 'sentence' appeared to be of ambiguous meaning to the Ballabgarh population and we could not find an alternative in the Haryanvi dialect. We then tried asking the subject 'Tell me something' and accepting any answer given in the form of a complete sentence. A response of 'What shall I tell you?' thus qualified as a correct response but also led to an awkward and somewhat pointless exchange between interviewer and subject. Some subjects responded with overlearned material such as stories or poems. We are now asking subjects 'Tell me something about your house' and awarding a full point to any complete sentence offered in response. While this item is now quite different from the original, we feel that it at least taps the ability to understand the task of generating a complete thought.

*Copying a figure:* In the original MMSE, the subject is asked to copy a figure consisting of two intersecting pentagons. A point is awarded only if there are 10 angles and two of them intersect. We were initially concerned that any task involving paper and pencil would be unfamiliar and intimidating in this population, but have since been impressed with subjects' willingness to attempt the copying task. However, we did find the intersecting pentagons to be too difficult for these subjects to copy and therefore substituted a simpler figure (a diamond within a square).

We have also simplified scoring; we have anchored scores to subjects' actual drawings in 30-3 categorized by neuropsychologists at both sites as acceptable or unacceptable. This subtest is an example of why clinical judgement is required for scoring, as opposed to administration and recording, of responses.

3. *Pilot phase.* After the pretest data were examined and appropriate modifications made to the test items, an age-stratified (55–64, 65–74, 75+) random sample of 100 subjects was drawn from the Ballabgarh census database for a pilot study. Field workers administered the cognitive battery including the HMSE, as finalized at the end of the pretest phases, to the subjects after obtaining informed consent and basic demographic information.

## RESULTS

### *Pilot phase*

These subjects were an age-stratified random sample of Ballabgarh elderly, who were 53% male and had a mean age of 70.7 (SD = 9.7) years; 26% had some education (10 subjects with 1–5 years, 11 subjects with 6–8 years and five subjects with 9–10 years of education); 74% had no formal education and were illiterate. Two women had 1–5 years of education; the rest of the women were uneducated.

For comparison, we used two subgroups of the

MoVIES ND sample. Those with  $\leq 9$  years of education numbered 26, had a mean age of 75.0 (SD = 6.1) years and were 46.2% male. Those with 10–12 years of education without graduating high school numbered 19, had a mean age of 71.5 (SD = 6.1) years and were 52.6% male.

Table 1 shows data on each MMSE (English or Hindi equivalent) item in the MoVIES ND sample (subgroups with  $\leq 9$ th grade and 10th–12th grades of education) and in the Ballabgarh random sample of 100 individuals participating in the pilot phase (subgroups with zero education and with 1–10 years of education). In addition to means and standard deviations of scores, we also show the percentage of each sample at ceiling (obtaining perfect scores) on each subtest and on the total as estimates of 'relative difficulty', along with the rank order of subtests with respect to relative difficulty in each phase (rank 01 denoting the easiest subtest, on which the largest proportion performed at ceiling).

Several sets of comparisons can be made, bearing in mind that the range for each subtest/item (as opposed to the total score) is small and does not

Table 1. MMSE/HMSE scores in Monongahela Valley and Ballabgarh samples

English/Hindi subtest/item	MoVIES non-demented sample				Ballabgarh pilot random sample			
	10th–12th grades (N = 19)		$\leq 9$ th grade (N = 26)		1st–10th grades (N = 26*)		Zero education (N = 74*)	
	Mean (SD)	% at ceiling (rank)	Mean (SD)	% at ceiling (rank)	Mean (SD)	% at ceiling (rank)	Mean (SD)	% at ceiling (rank)
Orientation to time	4.7 (0.5)	68.4 (08)	4.8 (0.4)	76.9 (07)	4.3 (0.7)	34.6 (13)	3.6 (1.0)	16.2 (12)
Orientation to place	5.0 (0.0)	100.0 (01)	5.0 (0.2)	96.2 (03)	5.0 (0.2)	96.2 (02)	4.4 (0.8)	47.3 (08)
3-object registration	3.0 (0.0)	100.0 (01)	2.9 (0.4)	92.3 (05)	2.8 (0.5)	88.5 (07)	2.2 (0.8)	41.9 (09)
Serial sevens/bus fare	3.4 (1.4)	21.1 (11)	3.1 (1.8)	26.9 (13)	3.8 (1.8)	50.0 (12)	1.6 (1.6)	5.4 (13)
WORLD/days backwards	4.1 (1.1)	52.6 (10)	4.5 (1.1)	73.1 (09)	4.3 (1.3)	76.9 (11)	2.3 (2.3)	37.8 (10)
3-object recall	1.7 (1.0)	21.1 (11)	1.9 (1.0)	30.8 (12)	2.8 (0.5)	80.8 (09)	2.5 (0.8)	60.8 (06)
Name watch	1.0 (0.0)	100.0 (01)	1.0 (0.0)	100.0 (01)	1.0 (0.0)	100.0 (01)	1.0 (0.0)	100.0 (01)
Name pencil/pen†	1.0 (0.0)	100.0 (01)	1.0 (0.0)	100.0 (01)	0.9 (0.3)	92.0 (06)	0.8 (0.4)	79.2 (03)
Repeat phrase	0.6 (0.5)	57.9 (09)	0.6 (0.5)	57.7 (10)	1.0 (0.2)	96.2 (02)	0.6 (0.5)	62.2 (05)
Read command/follow example	1.0 (0.0)	100.0 (01)	0.9 (0.4)	84.6 (06)	1.0 (0.2)	96.0 (05)	0.8 (0.4)	79.2 (03)
3-step task	2.7 (0.5)	78.9 (07)	2.4 (0.8)	53.8 (11)	2.8 (0.5)	80.8 (09)	2.4 (0.7)	54.1 (07)
Write sentence/'Tell me something'	1.0 (0.0)	100.0 (01)	1.0 (0.2)	96.2 (03)	1.0 (0.2)	96.2 (02)	1.0 (0.2)	97.3 (02)
Copy drawing	0.8 (0.4)	78.9 (07)	0.8 (0.4)	76.9 (07)	0.8 (0.4)	84.0 (08)	0.3 (0.5)	33.3 (11)
<b>Total (with WORLD/days)</b>	<b>26.6 (2.5)</b>	<b>5.3</b>	<b>26.6 (3.0)</b>	<b>11.5</b>	<b>27.6 (2.3)</b>	<b>23.1</b>	<b>21.8 (4.6)</b>	<b>1.4</b>
<b>Total (with sevens/bus fare)</b>	<b>26.0 (2.3)</b>	<b>5.3</b>	<b>25.2 (3.6)</b>	<b>7.7</b>	<b>27.0 (2.8)</b>	<b>19.2</b>	<b>21.0 (3.9)</b>	<b>0.0</b>

\*Because some subjects did not complete all subtests/items, means and SDs on some items are based on 1–3 subjects fewer than the entire sample.

†Incorrect responses to naming 'pen' in Ballabgarh included the colloquial 'holder', which has since been reclassified as an acceptable alternative response.

allow large differences to be manifested; that the age and education ranges of the Monongahela Valley and Ballabgarh cohorts are quite different; and that the sample sizes are small. Thus, we have focused on qualitative comparisons.

Comparing all four groups, the proportions obtaining perfect scores vary more across samples than do the mean scores (at least partly owing to the small range for each item). When rank-ordered by relative difficulty, the subtraction task (serial sevens or bus fare) was among the hardest for all groups, while naming, particularly of the watch, was the easiest item. The orientation tasks (for both time and place) as well as the copying task continue to pose more difficulties in Ballabgarh than in the Monongahela Valley. Even though the Monongahela Valley subjects had less difficulty initially registering the names of the three objects, the Ballabgarh elderly appear to have less difficulty recalling them subsequently. In all groups, the WORLD/days backwards attention subtest was easier than the subtraction (serial sevens or bus fare) subtest; therefore, the mean total HMSE score calculated with the former alternative was higher than the total calculated with the latter.

In general, mean scores and proportion at ceiling were higher in the Monongahela Valley than in Ballabgarh, and, in Ballabgarh, higher in the educated group. Mean totals were highest in the MoVIES non-demented sample (without much difference between the two educational groups), intermediate in the Ballabgarh educated subsample ( $N = 26$ ) and lowest in the Ballabgarh uneducated group ( $N = 74$ ).

In Ballabgarh, the educated group obtained uniformly higher mean scores on all tests with less 'relative difficulty'. While 100% of all subjects were able to name the watch, 92% of the educated (compared to 79.2% of the uneducated) subjects were also able to name the pen; 96.2% of the educated (in contrast to 47.3% of the uneducated) members obtained perfect scores on orientation to place; however, only 16.2% of the uneducated and 34.6% of the educated obtained perfect scores on orientation to time.

## DISCUSSION

Results of pretesting and pilot testing of the Hindi Mental State Exam have both quantitative and qualitative implications. The MMSE is among the most widely used of general mental status tests; it is

important that attention is paid to the appropriateness of each of its subtests and items for the population being studied, as well as to the comparability of different versions. Our data suggest that, carefully modified, the Hindi version is useful for assessing the cognitive functioning of our Ballabgarh elderly population, and that the same subtests or their equivalents are the easiest and hardest for the Ballabgarh and Monongahela Valley populations.

Some subtests required substantial modification which seemed clearly related to education. Examples include the attention subtest requiring abstract mental arithmetic skills for which this population has no other use and all subtests which require reading, writing and drawing. Since uneducated subjects can name a wristwatch with ease, it is likely that their relative difficulty with naming a pen reflects not aphasia but relative unfamiliarity with writing implements; this assumption is borne out by the finding that the educated subjects had little difficulty naming the pen. The notion of being asked to 'say a sentence', in no apparent context, is probably also less baffling to those who have been called upon to perform such tasks in the course of their schooling. Although we have previously shown that scores on all tests (including the MMSE) in MoVIES were directly related to level of education (Ganguli *et al.*, 1991), the differences are not apparent between the two small MoVIES subgroups (with less than 9th grade and 10th–12th grades) examined here. However, the differences in mean scores and relative difficulty between those with zero education and those with some education in Ballabgarh are fairly marked. A growing body of epidemiological literature suggests that lower levels of education are associated with higher prevalence of dementia and that lack of education may be a risk factor for Alzheimer's disease (Katzman, 1993; Mortimer and Graves, 1993). If there is a genuine difference in risk, one would expect to find higher prevalence rates of Alzheimer's diseases in societies with lower educational levels, perhaps in pandemic proportions in subgroups with zero education. While such a hypothesis renders the study of largely uneducated societies highly desirable, it simultaneously magnifies the methodological challenge of developing appropriate cognitive tests for uneducated and illiterate populations. In the Ballabgarh cohort, comparisons among the educated and uneducated members will be facilitated by the use of cognitive tests which can be performed by even illiterate individuals and will help to distinguish between the



effects of brain damage and those of lack of education.

In our sample, items on which there are greater apparent differences in relative difficulty, despite modification, appear to reflect genuine cultural and linguistic differences between the populations. Examples include the orientation items. Orientation to time, in the conventional sense, does not appear to play a major role in the daily functioning of Ballabgarh elderly. The year seems irrelevant to these subjects; with respect to season, individuals are likely to describe the observable weather rather than specify a period of the year. The season, while important for this agricultural community, is represented by a dialect word also meaning 'weather'; the seasons themselves are described as hot, rainy and cool, which also describe the weather on any given day. A further problem for testing is that there is no strict demarcation of seasons in India, for example designated days on which each season officially begins. Further, there are few environmental cues to reinforce knowledge of 'time' since routines and activities do not vary greatly according to day, date or month; the illiterate elderly do not read newspapers or write cheques, which would encourage them to keep track of dates in the same way as their urban or educated counterparts. Our difficulty in finding a meaningful fifth item for orientation to place (particularly among the uneducated) is another example. This generation of elderly has not been exposed in a major way to regions or events outside their own village, and have little reason to pay attention to larger geographic subdivisions. The illiterate do not receive or address mail and thus do not pay attention to postal regions and post offices. Clearly, it is helpful for such disparities to be recognized before employing, as a dementia screen, a scale which has been standardized in a different kind of population.

Variables on which comparisons are made must have similar meanings and be susceptible to similar interpretations in the groups being compared. Tests which are identical on the surface do not always satisfy this condition. The challenge is to identify the important underlying dimensions on which the populations are to be compared and find ways of assessing them that are appropriate to the groups being compared, rather than concentrating on superficial similarity of method. We have attempted to devise, for use in Ballabgarh, a Hindi version of the MMSE analogous to the English version in that it taps similar cognitive domains using tests of similar relative difficulty which involve, where possible,

similar task demands. In doing so, our initial step was an *a priori* attempt to distill the cognitive functions tapped by each test and, particularly, to identify the cognitive impairment or weakness that typically limited subjects' ability to perform the task in question. The next step was to check how well we had matched the tasks by observing Indian subjects perform the task, listening to their feedback and examining their scores. The importance of this 'qualitative' assessment cannot be underestimated. When preliminary observations and data suggested that the Indian subjects were achieving lower scores because they had more difficulty with naming the state, we did not conclude that they were more likely than their Monongahela Valley counterparts to be disoriented to place, in any useful sense of the term. Rather, we interpreted this finding as probable evidence that their culture and lifestyle coded 'place' differently and tried to select an orientation task they appeared to perform with the same ease that the MoVIES subjects identified their state.

In cases where we were satisfied that task demands were as similar as possible but scores remained low, we attempted to reduce the difficulty of the task without further changing its character (eg by simplifying the figure for copying). By 'similar scores' in this context we do not mean that scores need necessarily be numerically equal on all tasks since, as White (1989) points out, scores can be calibrated separately for different cultural settings just as they are frequently adjusted for different educational levels or age groups. However, we did attempt to secure a sufficiently similar distribution that the range of measurement was similar in the two groups and, in particular, that floor effects were not *more* marked in the Indian sample, as this would affect our ability to define population-based cutoff scores at the lower end of the range, as was done in the MoVIES project (Ganguli *et al.*, 1993a). Thus, as in MoVIES (Ganguli *et al.*, 1990), we found that the two attention subtests were not of equal difficulty for Ballabgarh elderly and that the subtraction task was the harder of the two. Across the entire scale, in both the Monongahela Valley and Ballabgarh, the easiest and hardest subtests were the same. The finding that recall of three objects was apparently easier in the Ballabgarh sample may be related to the additional cueing ('*I have brought with me from Delhi*'); possibly, Monongahela Valley scores could also be improved by such cueing without diminishing our ability to test recent memory.

In matching our test components in this way, we are implicitly assuming that the underlying

structure of cognition is similar in the two samples and that differences in their test performance reflect the different degrees to which basic cognitive skills are developed through education, culture and the demands and customs of the two environments in which our subjects live. The illiteracy of the majority of Ballabgarh elderly is a particularly difficult case. We have generally modified test items that require reading or writing to allow for oral presentation and response. In doing so, we assume that literacy is a skill developed through education and not a marker for the adequacy of the basic cognitive domain. We recognize that, in taking this approach, we will not be able to detect a cognitive impairment that manifests itself solely as an acquired dyslexia, but we believe that such situations would be rare. In most cases, we judge that the primary purpose of tasks that involve reading skills (eg spelling WORLD backwards) is to assess some other cognitive domain (in this case, attention) rather than literacy. This remains the purpose of our modified tasks. In situations in which we have had to modify MOVIES tests (eg by using the spoken 'follow example' rather than written presentation of the command 'Close your eyes'), we believe we have gained more, by including a task that is still analogous in important respects, than we have lost by modifying the method of presentation. The alternative would have been to delete the subtest from the instrument, which would have unduly hampered our ability to assess the cognitive function in question.

Cognitive test performance is affected not only by abnormal conditions affecting mental state (eg delirium and dementia) but also by demographic factors such as age, gender and education (O'Connor *et al.*, 1989; Holzer *et al.*, 1984; Escobar *et al.*, 1986), and cultural factors including comprehensibility, acceptability, perceived relevance of test content and familiarity with the language and with testing situations, concepts, procedures and materials. Therefore, no test is completely 'culture-free'; at best, it can be relatively 'culture-fair', ie it can avoid penalizing members of one culture for poor performance on tests designed for and standardized on members of another. The possible confounding of 'culture' by education must also be recognized. If the presence of dementia is to be suspected, and its severity to be defined, by neuropsychological test scores, it is essential to distinguish the effects of dementia on test performance from the preexisting and independent effects of culture, language and education. Tests which, for any of these reasons,

cannot be performed by a large proportion of non-demented individuals within a given culture are probably unsuitable as screens for dementia in that cultural group; and therefore also unsuitable for comparing that group with another group for which the test may be more appropriate.

In this report, we have described the adaptation of the MMSE for use in a largely illiterate Hindi-speaking population and shown the normative values obtained from a random sample of the elderly in this population. Further, ongoing work will address the utility of the Hindi Mental State Exam as a screening instrument for dementia in this population. Ideally, the validity of a cognitive test as a screen for dementia should be confirmed by its ability to discriminate between samples of demented and non-demented individuals and among demented individuals with varying degrees of severity. Additional confirmation can be obtained by longitudinal follow-up of patients with progressive dementias, to determine whether test scores decline over time. Such assessments can be carried out with relative ease in a case-control design if a pool of demented patients is available through existing clinical or research resources, which unfortunately was not the case in our Indian setting. Finally, it should be recognized that the objective, cross-sectional measurement of cognitive functioning can only identify subjects with a higher probability of being demented; it cannot be used to diagnose dementia in the absence of a history of cognitive and functional decline. Such information is ideally obtained from collateral sources, such as a reliable informant and/or medical records. Scales such as IQCODE (Jorm and Jacomb, 1989) designed to quantify cognitive decline data obtained from informants are a promising avenue for further research.

Even after translation and standardization of a 'culture-fair' test, a specific score may vary considerably in clinical meaning (ie in its reflection of a specific level of brain impairment) from group to group, implying that tests must be independently calibrated for each group (White, 1989). While the resulting instrument should be 'fair' to the populations being studied, it will not necessarily be equally fair for all populations; however, the systematic approach to instrument development should be generalizable to investigations of other populations. We believe that our experience, documented in this article, will be useful to others attempting similar research.

## ACKNOWLEDGEMENTS

The authors acknowledge with gratitude the following individuals whose efforts and support made this work possible: in New Delhi and Ballabgarh, Dr Lalit Nath, Dr Arun Mehta, Mr B. S. Nair, Mr R. K. Kaushik, Mr Roshan Lal, Mr Gajraj Singh, Mr Desh Raj and Mr Vijay Ram; in Pittsburgh, Dr Lewis Kuller, Ms Catherine Moran, Ms Deborah Echement, Mr Richard Rode, Mr Karl Weiner and Ms Rita Wolk. We also appreciate the suggestions of Dr Marshall Folstein and of an anonymous reviewer.

The work was supported in part by grants AG09202, AG07562 and AG0312 from the National Institute on Aging, US Department of Health and Human Services.

## REFERENCES

- Escobar, J. I., Burnam, A., Karno, M., Forsythe, A., Landsverk, J. and Golding, J. M. (1986) Use of the MMSE in a community population of mixed ethnicity: Cultural and linguistic artifacts. *J. Nerv. Ment. Dis.* **174** (10), 607-614.
- Folstein, M. F., Folstein, S. E. and McHugh, P. R. (1975) Mini-Mental State: A practical method for grading the cognitive state of patients for the clinician. *J. Psychiatr. Res.* **12**, 189-198.
- Ganguli, M., Ratcliff, G., Huff, F. J., Belle, S., Kancel, M. J., Fischer, L. and Kuller, L. H. (1990) 'Serial Sevens' vs. 'World Backwards': A comparison of the two measures of attention from the MMSE. *J. Geriatr. Psychiat. Neurol.* **3**, 203-207.
- Ganguli, M., Ratcliff, G., Huff, F. J., Belle, S., Kancel, M. J., Fischer, L., Seaberg, E. C. and Kuller, L. H. (1991) Effects of age, gender, and education on cognitive tests in a rural elderly community sample: Norms from the Monongahela Valley Independent Elders Survey. *Neuroepidemiology* **10**(1), 42-52.
- Ganguli, M., Belle, S., Ratcliff, G., Seaberg, E., Huff, F. J., von der Porten, K. and Kuller, L. H. (1993a) Sensitivity and specificity for dementia of population-based criteria for cognitive impairment: The MoVIES Project. *J. Gerontol. Med. Sci.* **48** (4), M152-M161.
- Ganguli, M., Seaberg, E., Belle, S., Fischer, L. and Kuller, L. H. (1993b) Cognitive impairment and the use of health services in an elderly rural population: The MoVIES project. *J. Am. Geriatr. Soc.* **41** (10), 1065-1070.
- Hall, K. S., Hendrie, H. C., Brittain, H. M., Norton, J. A., Rodgers, D. D., Prince, C. S., Pillay, N., Blue, A. W., Kaufert, J. N., Nath, A., Shelton, P., Postl, B. D. and Osuntokun, B. O. (1993) The development of a dementia screening interview in two distinct languages. *Int. J. Meth. Psychiatr. Res.* **3**, 1-28.
- Hendrie, H. (1992) Indianapolis-Ibadan dementia project. In Multi-national epidemiological studies of dementia (symposium abstract) (J. D. Curb and A. B. Graves, Eds). *Gerontologist* **32** (Suppl. 2), 219.
- Holzer, C. E., Tischler, G. L., Leaf, P. J. and Myers, J. K. (1984) An epidemiological assessment of cognitive impairment in a community population. In *Research in Community and Mental Health*, Vol. 4. JAI Press, Greenwich, Connecticut.
- Jorm, A. F. and Jacomb, P. A. (1989) The Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE): Socio-demographic correlates, reliability, validity, and some norms. *Psychol. Med.* **19**, 1015-1022.
- Katzman, R. (1993) Education and the prevalence of dementia and Alzheimer's disease. *Neurology* **43**, 13-20.
- Kertesz, A. (1982) *Western Aphasia Battery*. Psychologic Corporation, Harcourt, Brace, Jovanovich, New York.
- Larson, E. (1992) The Ni-Hon-Sea project: An overview. In Multi-national epidemiological studies of dementia (symposium abstract) (J. D. Curb and A. B. Graves, Eds). *Gerontologist* **32** (Suppl. 2), 219.
- Lowenstein, D. A., Arguelles, T., Barker, W. W. and Duara, R. (1993) A comparative analysis of neuropsychological test performance of Spanish-speaking and English-speaking patients with Alzheimer's Disease. *J. Gerontol.* **48**(3), P142-P149.
- Mortimer, J. A. and Graves, A. B. (1993) Education and socioeconomic determinants of dementia and Alzheimer's disease. *Neurology* **43**, Suppl. 4, S39-44.
- O'Connor, D. W., Pollitt, P. A., Treasure, F. P., Brook, C. P. B. and Reiss, B. B. (1989) The influence of education, social class and sex on Mini-Mental State scores. *Psychol. Med.* **19**, 771-776.
- Park, J. H. and Kwon, Y. C. (1990) Modification of the MMSE for use in the elderly in a non-western society. Part I. Development of Korean version of MMSE. *Int. J. Geriatr. Psychiat.* **5**, 381-387.
- Perez, E. A., et al. (1991) The Pan-American Health Organization (PAHO) Study. Presented at National Institute of Aging conference 'The Epidemiology of Alzheimer's Disease: The International Search for Environmental Risk Factors'. National Institutes of Health, Bethesda, MD, USA, July 9-12, 1991.
- Salmon, D. P., Riekkinen, P. J., Katzman, R., Zhang, M., Jin, H. and Yu, E. (1989) Cross-cultural studies of dementia: A comparison of MMSE performance in Finland and China. *Arch. Neurol.* **46**, 769-772.
- Tombaugh, T. N. and McIntyre, N. J. (1992) The Mini-Mental State Examination: A comprehensive review. *J. Am. Geriatr. Soc.* **40**, 922-935.
- White, L. R. (1989) Comparative international studies of dementia: Problems and strategies. Presented at the IVth International Congress of Psychogeriatrics, Tokyo, Sept. 7, 1989.
- Zhang, M., Katzman, R., Salmon, D., Jin, H., Cai, G., Wang, Z., Qu, G., Grant, I., Yu, E. (1990) The prevalence of dementia and Alzheimer's disease in Shanghai, China: Impact of age, gender, and education. *Ann. Neurol.* **27**(4), 428-437.

Copyright of International Journal of Geriatric Psychiatry is the property of John Wiley & Sons Inc. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.