International Psychogeriatrics, Vol. 9, Suppl. 1, 1997, pp. 87-94
© 1997 International Psychogeriatric Association

Mental Status and Neuropsychological Assessment

A Guide to the Standardized Mini-Mental State Examination

D. WILLIAM MOLLOY AND TIMOTHY I. M. STANDISH

ABSTRACT. The Mini-Mental State Examination (MMSE) is a widely used screening test for cognitive impairment in older adults. Because the guidelines for its application are brief, the administration and scoring of the test can vary between different individuals. This can diminish its reliability. Furthermore, some of the items must be changed to accommodate different settings, such as the clinic, home, or hospital. Because there are no time limits, it is not clear how long one should wait for a reply to a question. It is also not clear how one deals with answers that are "near misses." The goal of the Standardized Mini-Mental State Examination (SMMSE) was to impose strict guidelines for administration and scoring to improve the reliability of the instrument. The reliability of the MMSE was compared with the reliability of the SMMSE in 48 older adults who had the tests administered by university students on three different occasions to assess the interrater and intrarater reliability of the tests. The SMMSE had significantly better interrater and intrarater reliability compared with the MMSE: The interrater variance was reduced by 76% and the intrarater variance was reduced by 86%. It took less time to administer the SMMSE compared with the MMSE (average 10.5 minutes and 13.4 minutes, respectively). The intraclass correlation for the MMSE was .69, and .9 for the SMMSE. Administering and scoring the SMMSE on a task-by-task basis are discussed.

Since Dr. Marshall Folstein first published the Mini-Mental State Examination (MMSE) in 1975 (Folstein et al., 1975), it has become widely used as a screening test for cognitive impairment. The test, originally developed as a short screening instrument for clinicians, is now used as a screening tool to evaluate patients' suitability for inclusion/exclusion in clinical trials, and as an outcome measure in clinical trials. The test has gained wide acceptance because it covers a variety of cognitive domains, such as orientation to time and place, short- and long-term memory, registration, recall, constructional ability, language, and the ability to understand and follow commands.

This short test, usually completed in approximately 10 minutes, can be administered

From the McMaster University, Hamilton, Ontario, Canada (D. W. Molloy, MB, MRCP[I], FRCP[C], and T. I. M. Standish, MA).

Offprints. Requests for offprints should be directed to Dr. D. William Molloy, 440 Orkney Road, RR#1 Troy, Ontario, LOR 2BO, Canada.

effectively, after a short training period, by physicians and nurses. However, the original MMSE had few instructions for administration and scoring; consequently, these were left to the discretion of each rater. As a result, different raters developed their own styles and techniques of administering and scoring the instrument, which differed widely in some cases, compromising the reliability and widespread application of the instrument. For example, one colleague gave half points to patients who named a day, month, or season incorrectly, arguing that this was better than someone who could not name any day, month, or season at all. Others gave hints, told patients if they gave the wrong answer on the first attempt, and encouraged them to have a second guess. This variability in administration and scoring decreased the reliability of the test. Given that this instrument was being used to screen patients for experimental treatment trials and as an outcome measure to assess the effectiveness of new treatments, it was important to ensure that the test should have satisfactory interrater and intrarater reliability.

Use of the MMSE in clinical trials posed problems of variability, not only with the same rater, but also when different raters scored the same patient in a test-retest situation. Without clear explicit guidelines, multiple raters inevitably administered and scored the test in their own way unless clear instructions were provided.

Clinicians intuitively recognized their own intrarater (same rater) variability and attempted to minimize it. Many different groups developed their own unique standardization procedures to settle disputes about how to score or administer the MMSE. Over time, if they could agree, they developed a consensus on a system of administration and scoring. However, in clinical trials with multiple sites, the problem of intersite variability, a result of interrater variability, confounded interpretation of results. This was apparent when raters from different sites had different administration and scoring procedures compared with other sites. Attempts to deal with this interrater variability wasted significant amounts of time when multidisciplinary clinical teams and different clinicians at investigators' meetings attempted to try to develop a consensus on how the MMSE should be administered and scored. The wheel was constantly being reinvented.

The Standardized Mini-Mental State Examination (SMMSE) was developed and tested to provide clear explicit administration and scoring instructions to reduce variability and increase the reliability of the MMSE (Molloy et al., 1991a).

STANDARDIZATION OF THE MMSE

In our clinic, different individuals from a multidisciplinary team and research assistants working on different clinical trials came to a consensus on scoring and administration guidelines for the MMSE after many hours of discussion. We answered the questions that different raters raised, such as: How is a question to be asked? What is an acceptable answer? How long should one wait for an answer? The SMMSE attempted to provide clear, unequivocal answers to these questions. Standardized instructions and administration techniques were developed and refined for each item.

Setting Up

Before the test was administered, raters were given seven general rules. For example, raters were advised to make sure subjects have their hearing and/or visual aids. A simple test of hearing and comprehension is performed routinely to establish whether or not the subject understands and hears and to ensure the right person is being tested, by asking, "What is your name?" Next, subjects are warned that they will be asked some questions and permission is sought to ask these questions, "Would it be all right to ask you some questions about your memory?"

Each question may be asked three times if the subject does not seem to understand or has not attempted to answer. If subjects attempt to answer and are incorrect, they score zero and the rater moves on, giving no clues verbally or physically. Certain props are required, such as a pencil (sharpened at one end with a rubber on the other), and a large piece of paper with "Close Your Eyes" written on one side and two five-sided figures intersecting to make a four-sided figure on the other side. Raters also need a clock to measure time in seconds.

Administration

Each administration of the SMMSE begins with "I am going to ask you some questions and give you some problems to solve. Please try to answer as best you can." Each item has its own specific instruction, based on clinical experience and on what most people understood and found unequivocal. For example, in the orientation section of the MMSE, the rater is instructed to ask questions about day, date, season, etc., but is not told exactly what to say. Different instructions to the patient could possibly elicit different responses. One rater may ask, "What is the date?" or "What is today's date?" whereas another rater may ask, "What's the date?" and after the response follow with abbreviated instructions such as "The day?", "The month?", etc.

When the instructions are left to individual raters, they use their own questions, which differ from those of other raters and, therefore, act as a potential source of variability. The SMMSE provides exact verbatim instructions for every item. For example, the SMMSE sets questions such as "What year is this?" and "What month of the year is this?" The rater is instructed to say the questions exactly as they appear in the SMMSE.

Standardized Scoring

Some issues arise in interpreting correct answers and responses when scoring the MMSE. Some tasks are relatively straightforward and easily scored. For example, subjects are asked to read the statement "Close your eyes" and to perform the task. The scoring is simple: The subject either closes his eyes or he does not.

In contrast, in scoring the spelling of "World" backwards, the combinations and permutations of omissions and reversals can lead to problems. A problem arises when the subject cannot spell "World" correctly forwards or spells "Word" instead. If the subject cannot spell "World" forwards, he is not given the task to spell it backwards and is scored zero.

In scoring the orientation tasks, such as naming the present month, problems

arise when testing is being conducted within 1 day or 2 of the change of the month. Raters may feel that it is normal to mix up a single date. If the subject gives the wrong month on the final day of 1 month or the first day of the next month, some raters may give the point. However, another more exacting rater may insist on the exact date and withhold the point.

In the SMMSE, in scoring the date item, subjects are awarded the point if they give the previous or next day's date, e.g., on the 7th, the 6th, or 8th is acceptable. Subjects are also awarded a point if, on the first day of a new month or the last day of the previous month, they say either of these months. In addition, during the last week of an old season or first week of a new season, either season is accepted.

Problems also arise when scoring the five items of the SMMSE that deal with orientation to place: country, province/state/county, city/town, building, and floor. The places are asked in order, from the largest geo-political unit to the smallest. Because the relative importance of geography and political divisions differs from place to place, it should be decided in advance of testing what will be accepted as appropriate answers in each context (excepting country). A general guideline should be based on how most people would describe their location. In some cases, the county may be a more important unit than the state, therefore the name of the county would be asked. Towns incorporated into larger municipalities for administrative purposes may retain a meaningful identity for some, so it may be more useful to ask for the town, even if it has little political status.

The name of the building where the testing is being conducted is the next question. For example, we work in a hospital called the Henderson General Hospital, which was part of the Hamilton Civic Hospitals. It was known locally as the "Nora Francis" because the hospital was originally named after Nora Francis Henderson. In the last year, the hospital has merged with Chedoke-McMaster Hospitals and is now called Hamilton Health Sciences Corporation. Within our hospital, there is a cancer center and a maternity wing. Because this item is so confusing, a list of acceptable responses has been created.

In the final question referring to location, the patient is requested to name the floor of the building where the test is taking place. This question loses some relevance in cases where there is only one floor. However, the question should be posed anyway. Decide if "main" or "first" or "ground" are to be accepted equally. This can be particularly confusing where there are multiple entrances at different levels. In multiple floor situations, with many different "ground" floors, a list of acceptable responses for each building may be necessary.

Questions and answers have been standardized for different settings in the community, such as the home, e.g., "What street is this?" (as an alternative to the question referring to the hospital/building) or, "What room are we in?" (as an alternative to the question pertaining to the floor).

Variability also occurs from simple errors in calculation or recording responses. The SMMSE provides a score sheet that is clear and easy to use. Space is also provided to record the subject's actual answers as well as the score for

each item, which allows scoring to be verified after the test is completed.

Time Limits

The original MMSE did not have any suggested time limits for subjects to complete a task. However, time is a factor in controlling variability. Most subjects will answer a question fairly quickly whereas others will think for a long time before answering. In the silence, between question and response, raters may not know whether the subject is thinking about the answer or has forgotten the question. Furthermore, the rater may be reluctant to pose the question again in case this interrupts the subject's train of thought. In these circumstances, different raters have their own arbitrary time limit before they either ask the question again or move on to the next question. Patient raters may wait longer for an answer and continue to ask the question until the subject gets it right or gives up.

Some subjects will give several answers, some correct, others incorrect. Which one should be accepted? Some raters will give the point if one of the answers was correct, whereas others will only award the point if the first or last answer was correct and ignore the rest. Variability can also be caused by variations in the total time taken to complete the test. Some raters who have long, drawn-out sessions may make subjects fatigued and this ultimately affects performance.

The SMMSE imposes explicit time limits on all tasks to reduce variability and enhance reliability. In all cases, the rater begins timing from the end of the instruction. Subjects are not aware that the questions are being timed, because this could add a sense of urgency and lead them to rush their answers. The rater times the test inconspicuously and, should a patient take longer than the allowed limit, the rater ends the task simply by saying "Thank you, that's fine" and proceeds to the next question. If a patient is trying hard to complete a task, for example, drawing the five-sided figures, and exceeds the 1-minute time limit, a rater may allow the patient to complete it. The scoring is still based on the work performed during the time limit and the patient would not get the point if it was not completed correctly during this period. Other patients having difficulty with a task are relieved when the task is stopped by the time limit. This latitude is allowed to preserve the patient's dignity and prevent catastrophic reactions.

SMMSE Props

Another potential source of variability in the traditional MMSE was in the use of props for the "Close your eyes" task and drawing of the five-sided figures. It was left to the rater to write the words or copy the figures onto a piece of paper as an example to show the patient. Often these words or figures were hand-written on the spur of the moment and the quality of the example was quite variable. The SMMSE provides the instruction "Close Your Eyes" and the two five-sided figures on opposite sides of a laminated card, each in clear black-on-white format, with no other distractions on the card. "Close your eyes" is in a very large and plain font. The five-sided figures clearly show the four-sided overlapping figure that must be copied.

"Ball, Car, Man" — Registration

In the original test, the three words "Apple," "Table," and "Penny" were suggested to test registration and recall. In some cases, where subjects were tested repeatedly in studies, raters would say "I am going to name three objects and I want you to repeat them back to me," and before they could say the words, the subject would say "Apple, Table, Penny." It became obvious to us that we needed alternate forms of these three words. As a result, we created alternate three-word sets that had the same word frequency, e.g., "Ball, Car, Man" and "Bull, War, Pan."

In this task, the rater slowly names three objects to test the subject's ability to register new information. The subject is asked to repeat them within 20 seconds. One point is scored for every word correctly recalled. The order of recall is not scored. The rater may not repeat the words, so it is important to enunciate clearly and control for distractions when this task is being performed. After the subject has recalled as many as he or she can, but has not registered all three, the rater will help the subject to learn all the words for the delayed recall segment. The rater repeats the words at 1-second intervals and then asks the subject to repeat them until all three are repeated. The rater repeats to a maximum of five times or until they are learned. The subject is advised that he or she will be asked to recall them later, after the task of the spelling of "World" backwards.

"World"

In this task, the patient is asked to spell the word "World." After successfully spelling it, he or she is asked to spell it backwards to test the subject's ability to reverse the letters of the word. Five points are given for correctly reversing the letters, 4 points are given for omission of one letter, 3 points for omission of two letters, 2 points for reversal or omission of three letters, and 1 point is given for reversal of four letters. A list of possible answers with scores for each is provided in the SMMSE.

Serial Sevens

The serial sevens are presented as an alternative task to spelling "World" backwards. However, the two tasks are not equivalent. The advantage of serial sevens is that the scoring is easier, because the subject is scored according to the number of correct subtractions (Molloy et al., 1991a).

Watch and Pencil

In the naming tasks, the subject is asked to name two common objects, a watch and a pencil. In each case the subject is shown the object and is asked to say what it is. "Clock" or "time" are not accepted for the watch, and "pen" is not accepted for the pencil. The watch used should be generic in appearance with a dial (not digital). The pencil should be sharpened to a point and have an eraser on the other end. Ten seconds are allowed to respond to each.

"No ifs..."

In this task, the subject is asked to repeat the phrase "no ifs, ands, or buts" after the rater. The subject has 10 seconds to respond and must say the phrase verbatim to score the point. Raters must be careful to enunciate the words properly

because patients with high-frequency hearing loss often do not hear the sibilants.

Folding Paper

This task measures the subject's ability to follow a three-step command. The rater holds up a piece of paper in front of the subject and says "Take this paper in your [nondominant] hand, fold the paper in half once with both hands, and put the paper down on the floor." Thirty seconds are allowed and 1 point is given for each of the three steps properly executed. The subject's nondominant hand should be determined at the beginning of the SMMSE administration. The rater should not allow the subject to take the paper before the entire command is given, but should hold the paper in the midline of his or her body, just out of reach of the subject while the command is being given. Once the three-step command has been finished, the rater pushes the paper forward within reach of the subject.

Write Sentence

In this task, the subject is given a piece of paper and a pencil and is asked to write a complete sentence. Thirty seconds are allowed and 1 point is scored for a sentence that makes sense, i.e., it should have a subject, verb, and object. Spelling mistakes are ignored.

Overlapping Pentagons

Constructional ability is tested by asking the subject to copy a design of two overlapping pentagons. A pencil, eraser, and paper are set before the subject. One minute is allowed for the subject to complete the diagram, and a four-sided figure must be created by the overlap of the two pentagons to score 1 point. The SMMSE presents several examples of diagrams drawn correctly and incorrectly to assist the rater in scoring. One common scoring problem is that the elderly have difficulty drawing straight lines and lines tend to curve rather than form distinct angles. In such cases, the rater may have to determine when a line is merely wiggly and when an angle is intended.

When the final task of the SMMSE is completed, the rater thanks the subject and offers some reassurance such as "You did well." The rater then tallies up the score. The maximum score is 30.

RELIABILITY OF THE SMMSE

The reliability of the SMMSE has been compared with that of the traditional MMSE (Molloy et al., 1991a). In this study, university students with no experience at neuropsychological test administration were recruited as raters.

Half of the students were randomized to administer the MMSE and half were randomized to administer the SMMSE. Separate training sessions were held for each group, consisting of a short talk tailored to their particular version, but of equal duration, and each trainee was allowed a practice session. Forty-eight subjects from a nursing home and a chronic-care hospital unit were recruited to have the tests administered on three separate occasions, at intervals of 1 week.

Each rater tested the same subjects on the first and third weeks while a different rater administered the tests at the second week. This design allowed both intrarater (test–retest) and interrater

reliability to be evaluated. Intrarater variance was reduced significantly, by 86% ($p < .003$), and interrater variance was reduced by 76% when the SMMSE was used. Intraclass correlation for the MMSE was .69, and .90 for the SMMSE. The mean duration of assessments was 13.39 minutes for the MMSE and 10.47 minutes for the SMMSE ($p < .004$).

CONCLUSION

The SMMSE offers useful administration and scoring guidelines for the MMSE. It increases reliability and requires less time to administer. It may be applied as a screening test for clinical trials (Field et al., 1995; Molloy et al., 1991b, 1991d) and, in particular, as an outcome measure in clinical trials (Molloy et al., 1991c). The SMMSE also differentiates accurately between people who can learn and ultimately complete advance directives from those who cannot (Molloy et al., 1996). In addition, it demonstrates equivalent reliability when administered in the clinic or the patient's home (Bédard et al., 1995). French, German, Italian, and Spanish translations are available on request.

REFERENCES

Bédard, M., Molloy, D. W., Standish, T., Guyatt, G. H., D'Souza, J., et al. (1995). Clinical trials in cognitively impaired older adults: Home versus clinic assessments. *Journal of the American Geriatrics Society, 43,* 1127-1130.

Field, S. J., Jackson, H. J., Hassett, A. M., & Pattison, P. (1995). Ability of the Mini-Mental State Examination to discriminate diagnostic entities in a psychogeriatric population. *International Journal of Geriatric Psychiatry, 10,* 47-53.

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

Molloy, D. W., Silberfeld, M., Darzins, P., Guyatt, G. H., Singer, P. A., et al. (1996). Measuring capacity to complete an advance directive. *Journal of the American Geriatrics Society, 44,* 660-664.

Molloy, D. W., Alemayehu, E., & Roberts, R. (1991a). Reliability of a standardised Mini-Mental State Examination compared with the traditional Mini-Mental State Examination. *American Journal of Psychiatry, 148,* 102-105.

Molloy, D. W., Clarnette, R. M., McIlroy, W. E., Guyatt, G. H., Rees, L., et al. (1991b). Clinical significance of primitive reflexes in Alzheimer's disease. *Journal of the American Geriatrics Society, 39,* 1160-1163.

Molloy, D. W., Guyatt, G. H., Wilson, D. B., Duke, R., Rees, L., et al. (1991c). Effect of tetrahydroaminoacridine on cognition, function and behaviour in Alzheimer's disease. *Canadian Medical Association Journal, 144,* 29-34.

Molloy, D. W., McIlroy, W. E., Guyatt, G. H., & Lever, J. A. (1991d). Validity and reliability of the Dysfunctional Behaviour Rating Instrument. *Acta Psychiatrica Scandinavica, 84,* 103-106.