AMEE Medical Education Guide No. 9. Assessment of clinical competence using the Objective Structured Long Examination Record (OSLER)

FERGUS GLEESON
James Connolly Memorial Hospital & The Royal College of Surgeons in Ireland, Dublin, Ireland

SUMMARY *Much criticism has been directed at the assessment of clinical competence and at the long case in particular in recent years. In the traditional long case candidates spend one hour with a patient from whom they take a history and whom they examine. An examiner is not present. The student is then examined by a pair of examiners over a 20–30 minute period. This has been to the extent that the problems associated with the long case in terms of objectivity, validity and reliability are such that some critics have suggested that it should be abandoned altogether. Others would take the view that before we dispense with this method we should attempt to remodel and improve it. Furthermore, tradition and practically would suggest that the long case will be with us for some time to come. The justifiable criticism of the long case is directed on a number of fronts, a major one being that the history-taking process is not observed by the examiners. Bearing these criticisms in mind, the Objective Structured Long Examination Record (OSLER) has been developed. The OSLER is a 10-item analytical record of the traditional long case which attempts as far as is possible within the limits of practicality to improve the objectivity, validity and reliability of existing practices. All candidates are assessed over 20–30 minutes by the examiners on the same 10 items, thus improving reliability and items are included that are representative of what would be regarded as having an acceptable degree of construct or face validity with regard to the long case. Attention is paid to communication skills and the history-taking process in particular. In attempting to standardize the long case and minimize the 'luck of the draw' aspect, examiners are requested to formally document the difficulty of the case. The figure of 10 with regard to the number of items assessed is not coincidental and is a deliberate act to include a minimum of the essential in terms of what should be assessed. This allows examiners to concentrate on the candidate's performance with a structured guide that is not so intrusive as to interrupt the examiner's concentration. The four items on history include pace and clarity of presentation, communication skills process, systematic approach and establishment of the case facts. Three items on physical examination include systematic approach, examination technique and establishment of the correct physical findings. During these activities the candidate's affective behaviour is also assessed. The remaining three items include construction of appropriate investigations in a logical sequence, appropriate management and final clinical acumen. The latter item draws on the previous nine to assess candidates' ability to identify and solve problems. The initial assessment is essentially criterion referenced through a P+, P, P− system which is followed by the selection of an appropriate mark, each of which has its own written descriptive profile. The perfect method for long case clinical assessment has yet to be established. Indeed perfection may be no more than a pious hope bearing in mind that any method will always be a compromise between objectivity, validity and reliability on one hand and practicality on the other. While the search for the perfect long case method continues, the OSLER is suggested as a practical approach to what is universally recognized as an ongoing assessment challenge.*

Introduction

Assessment is treated with great reverence in the vast majority of medical schools. Lowry (1993), however, has recently posed the question: Is assessment as powerful as we think, and if it is, are most medical educators using it

Correspondence: Fergus Gleeson, FRCPI, Associate Professor of Medicine, James Connolly Memorial Hospital, Dublin, Ireland. Fax: 8204708.

0142-159X/97/010000-00

effectively? Clinical assessment in many medical schools, in spite of frequent criticism, has continued to be a combination of what are commonly termed long- and short-case examinations. This combination is likely to persist if one accepts that clinical assessment, to be truly valid, must be patient centred. Over the past 20 years, the short case examination has received much attention with the introduction of the Objective Structured Clinical Examination (OSCE) (Harden & Gleeson, 1979). With most attention being paid to these improvements, the long case has largely been ignored. While improvements such as the OSCE have focused attention on the individual components of clinical competence, it is widely agreed that there is still need for a method to assess students on the patient as a whole. The traditional long-case examination has been our method of fulfilling this role. There has been much justified criticism of the long case in which different examiners examine different candidates on different patients. This has very rightly been referred to as the 'luck of the draw' (Stokes, 1974). What, therefore, educationally constitutes a good assessment method? It should be objective, valid and reliable. Why does the existing long case fail to meet such criteria to an acceptable degree? Such assessments are frequently heavily subjective in that there is little prior agreement between pairs of examiners or indeed by institutions as a whole as to what constitutes a valid assessment. In other words, there is a lack of or no agreement as to what has to be measured during the course of the examination. What constitutes a valid assessment? Such an assessment measures what it is supposed to measure, i.e. clinical skills. These skills include the ability to obtain information by way of history, physical examination and investigations, to use this information to solve patients' problems and finally to utilize the solution to problems by way of management. In most existing long-case assessments history taking as such is not validly assessed. While the product may be assessed, the far more important history-taking process is not observed and therefore not validly assessed. This is a highly significant omission when one considers the relative value of the history in terms of overall diagnostic problem solving (Hampton *et al.*, 1975; Miall, 1992).

It is for this reason that the OSCE has achieved much of its deserved success as observed history taking plays a significant part in such assessments. However, while the OSCE displays a relatively high content validity, it has a relatively low face or what is termed construct validity. In other words, while it assesses the parts very well it does not assess the whole candidate/patient interaction on one and the same patient which, after all, is what occurs in the practice of medicine. In view of what has been stated in terms of objectivity and validity, there is thus a high probability that there will be inconsistencies or a lack of reliability in the marking of the long case if there is not a clear agenda to be followed. While examiners may at times be following a similar agenda, the items may receive significantly different emphasis by the individuals of pairs of examiners so that marking inconsistencies are a recognized problem (Fleming *et al.*, 1976).

While the importance of patients' histories is universally recognized, the emphasis placed on physical examination in the long case needs to be critically reappraised for a number of reasons. One of these is the lack of gross physical signs in the majority of patients examined in practice. Furthermore, many very difficult clinical problems do not have any physical signs and are therefore frequently not used in such assessments (Weatherall, 1991). Social and psychological factors play a significant part in the day-to-day problems encountered in both hospital and community practice.

Increasingly therefore, recognition of the value of communication skills is being highlighted (Irwin *et al.*, 1989; Doherty *et al.*, 1990; McManus *et al.*, 1993). The communication skill necessary to acquire information on difficult clinical problems is very real and consequently places a responsibility on institutions and their examiners to establish that such skills have been developed by way of assessment. This need therefore emphasizes again the relative importance of the history. Another key factor is the degree of case difficulty over a wide range of long cases, which is very variable and must be given due recognition by examiners. A further danger in the course of long-case examinations is that during the assessment, unless there is a structure to be followed, the emphasis may shift from the clinical to that of a viva or oral assessment. This is not an infrequent occurrence and in such circumstances the validity of the clinical examination is obviously seriously compromised. There are therefore genuine concerns about the existing traditional long case which frequently result in the making of global pass/fail decisions in a non-structured fashion. Such decisions at times result in questionable outcomes in terms of justice to the candidate and to the public at large, which is the ultimate reason for all assessments. Such a scenario could be likened to referees adjudicating in different games, using different rules and in the end miraculously producing an overall winner. Can one imagine a similar scene in any other field of human endeavour?

In the traditional long case students spend one hour with a patient from whom they take a history and whom they examine. An examiner is not present. The student is then examined by the examiner over a 20–30 minute period.

Long-case assessments at both undergraduate and postgraduate level are in most instances carried out over a short period of time, e.g. one week. During this time frame, large numbers of candidates have to be assessed. There is therefore a need for an improved long-case format to assess such large numbers that is practical to implement but at the same time recognizes the essential criteria of objectivity, validity and reliability in so far as this is possible under the time constraints already referred to. In spite of the obvious problems that exist with regard to the long case, there is a natural reluctance to change established practices unless very real benefits are possible. For any educational innovation to succeed, there are certain criteria to be fulfilled (Collingwood, 1979). The innovation must have a relative advantage over existing practice. The complexity of the innovation must not be such as to evoke an immediate and negative attitude. It must have trialability in that it can be introduced and removed in the event of failure without producing a major convulsion in the system. Finally, it must be seen to have observability in that the more visible the effect of the innovation, the more likely will be its acceptance.

The current unstructured global marking of the long case has major potential for unreliable assessment. An essential requirement therefore is the need to structure a

number of items for examiners to deliberate on. This results in turn in the introduction of the concept of the checklist. For a comprehensive long-case examination, the potential length of such a list would be so great as to be impractical to implement. Such an instrument would end up being both an invalid and unreliable instrument in that the examiner would spend more time concentrating on the checklist rather than on the actual measurement of the candidate's performance. More realistic approaches such as the observed long case proposed by Newble (1991) and a similar approach by Price & Byrne (1994), for assessment skills in psychiatry are both very expensive in terms of examiner time. Both approaches require examiners to be present for the whole history-taking process carried out by the candidates. This extra time element would make such assessments impractical for the vast majority of institutions, particularly in those situations in which large numbers of candidates have to be assessed in a relatively short time frame. The method adopted, therefore, must be comprehensive enough to allow for valid judgements by examiners, be practical to use and at the same time be perceived as fair by the candidates. Such perception demands that, as with all assessment instruments, it must be seen to be objective. It must also be structured so that all candidates are assessed using the same criteria leading to greater consistency or reliability.

A valid method for the assessment of long-case clinical competence must include essential principles. These are the recognition of and observation of the history-taking process. While such is being observed the examiner has the opportunity of assessing the communication skill of the candidate. Physical examination skill is essential, as is the ability to construct a series of investigations. All of the foregoing allow the examiner to deliberate on the candidate's ability to identify and solve problems. Finally there is a requirement to assess the candidate's ability to manage the problem, which again involves skills of communication as well as overall management. During all these activities, the examiners will also have the opportunity of assessing the affective behaviour or attitude of the candidate towards the patient. The assessment instrument must be practical in terms of its length and usage by the examiner whose primary function is to concentrate on the candidate's performance. The Objective Structured Long Examination Record (OSLER) has been developed in an attempt to fulfil the stated foregoing criteria and principles. Examiners spend 20–30 minutes with the student who has already examined and taken a history from the patient.

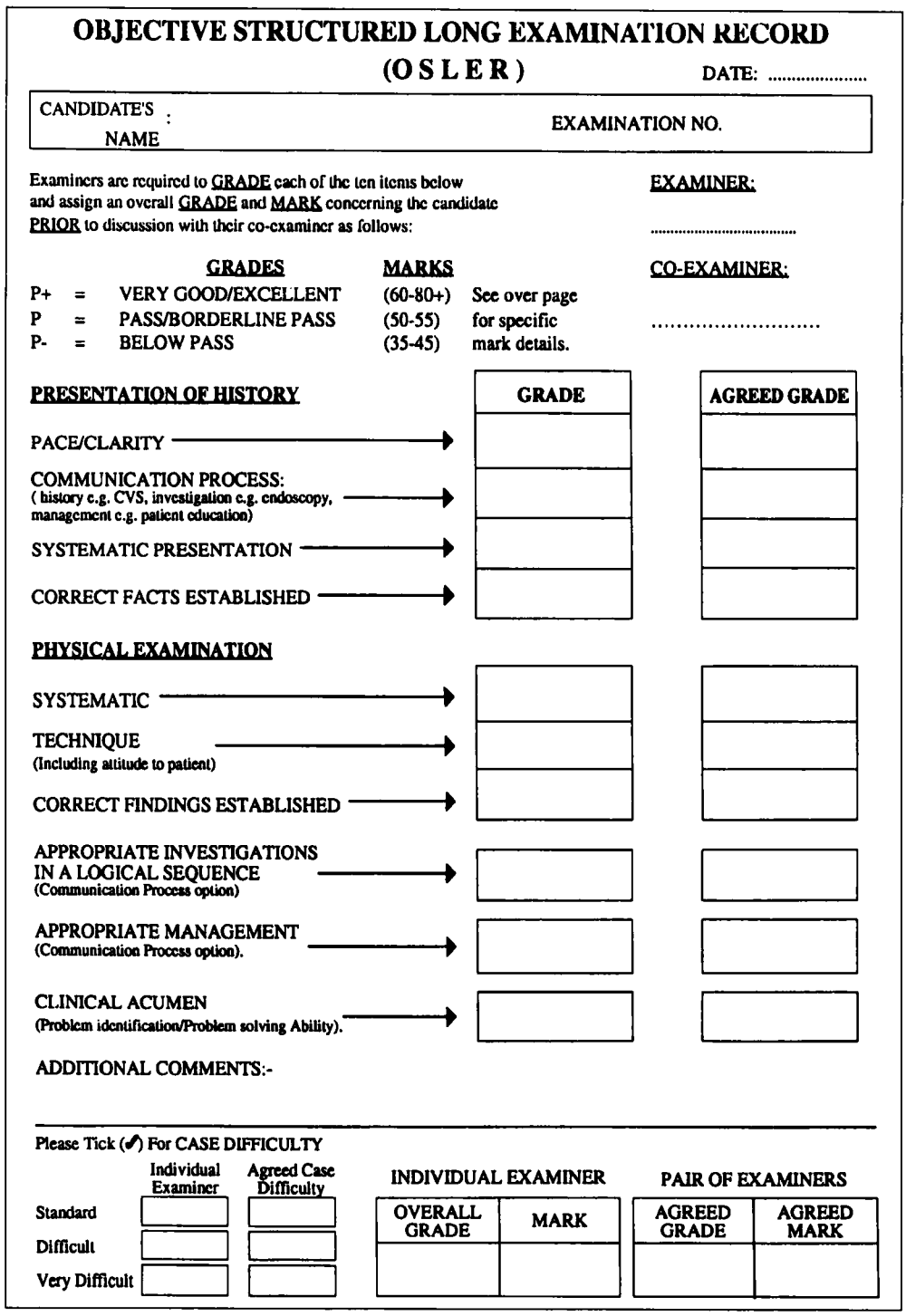
Method

Presentation of history

The OSLER consists of 10 items (Figure 1) which include four on history, three on physical examination and the remaining three cover investigation, management and clinical acumen. The figure of 10 is not coincidental and is a deliberate act to include as much as is essential but as little as possible. This is to allow the examiner to concentrate on the candidate's performance with a guide that is not so intrusive as to interrupt the examiner's concentration. The four items assessed on the history are pace and clarity of presentation, communication skills process, systematic approach and establishment of the case facts. Pace/Clarity essentially assesses communication between the candidate and the examiner. Pace of presentation measures rate of speech with appropriate pauses. Too rapid and it is unintelligible, too slow and it is inefficient in terms of time economics. Clarity is obviously allied to pace but at the same time recognizes the need and ease with which the examiner observes the unfolding story that is the history. Greater emphasis is now being placed on communication skills in medical schools (GMC, 1993) and the inclusion of the first item recognizes this fact. Graduates of medical schools are employed worldwide and if the candidate cannot effectively communicate with the examiner, he/she cannot be validly assessed. More importantly, if the candidate cannot make him/herself understood by the examiner, what chance has the patient? It is essential therefore that the examiner has an opportunity of observing the communication skills of the candidate with the patient through the second item, communication process. This is achieved by requesting the candidate to take a history for three minutes concentrating on one system, e.g. cardiovascular, or segment of the history, e.g. social history. By observing this process and listening to the remainder of the history, the examiner can form an opinion as to the candidate's ability to communicate with the patient. Alternatively the communication skill of the candidate can be assessed during the assessment of the investigation or management sections. This could be achieved by the candidate describing to the patient a particular investigation, e.g. colonoscopy. Alternatively a candidate could be asked to explain to the patient, as part of the management, the usage and dangers of anticoagulants. By listening to the remainder of the history, the candidate's ability to systematically go through the story in a logical sequence can be assessed. Finally it is essential that the candidate demonstrates his/her ability to accurately establish the correct facts of the case.

Physical examination

Three items are a minimum of the essential for inclusion in relation to physical examination. Here again the process as well as the product is being observed and assessed. A systematic approach will reveal something of the candidate's ability to logically approach the subject to obtain the necessary information to problem solve. However, the key to successful physical examination lies in a well-developed technique. This item deals with the candidate's psychomotor skills and, like all such skills, frequent practice is the essential requisite. An experienced astute examiner will be in a position to decide on the merits of a candidate in this section. Not alone are the pure psychomotor skills being observed but also the candidate's confidence and attitude towards the patient. Influences other than technique can affect the performance of psychomotor skills on any particular occasion; however, the candidate with a truly professional approach which includes attention to detail can overcome such influences. The most obvious of these is the relative difficulty of the case the candidate is assigned. The 'luck of the draw' is a well-accepted factor and the experienced examiner will recognize this. All examiners therefore need to be consistently conscious of this factor and an



OBJECTIVE STRUCTURED LONG EXAMINATION RECORD

(O S L E R) DATE: .....

CANDIDATE'S NAME : EXAMINATION NO.

Examiners are required to GRADE each of the ten items below and assign an overall GRADE and MARK concerning the candidate PRIOR to discussion with their co-examiner as follows:

EXAMINER:

.....

CO-EXAMINER:

.....

		GRADES	MARKS	
P+	=	VERY GOOD/EXCELLENT	(60-80+)	See over page
P	=	PASS/BORDERLINE PASS	(50-55)	for specific
P-	=	BELOW PASS	(35-45)	mark details.

	GRADE	AGREED GRADE
PRESENTATION OF HISTORY		
PACE/CLARITY →		
COMMUNICATION PROCESS: (history e.g. CVS, investigation e.g. endoscopy, management e.g. patient education) →		
SYSTEMATIC PRESENTATION →		
CORRECT FACTS ESTABLISHED →		
PHYSICAL EXAMINATION		
SYSTEMATIC →		
TECHNIQUE (Including attitude to patient) →		
CORRECT FINDINGS ESTABLISHED →		
APPROPRIATE INVESTIGATIONS IN A LOGICAL SEQUENCE (Communication Process option) →		
APPROPRIATE MANAGEMENT (Communication Process option). →		
CLINICAL ACUMEN (Problem identification/Problem solving Ability). →		

ADDITIONAL COMMENTS:-

Please Tick (✓) For CASE DIFFICULTY

	Individual Examiner	Agreed Case Difficulty
Standard		
Difficult		
Very Difficult		

INDIVIDUAL EXAMINER		PAIR OF EXAMINERS	
OVERALL GRADE	MARK	AGREED GRADE	AGREED MARK

Figure 1. The OSLER

assessment of the case difficulty is included in the OSLER to aid this process. It should of course also be borne in mind that, in later practice, the 'luck of the draw' will apply on a daily basis and the candidate should have the flexibility to demonstrate that he/she can handle any given situation under the prevailing circumstances. Whatever difficulties are encountered the candidate has to correctly identify the clinical signs to proceed satisfactorily to manage the patient's problem.

Investigation, management and clinical acumen

For the item on investigation, the examiner is requested to assess the candidate's ability to construct appropriate investigations for the case in question in a logical sequence. Frequently, appropriate investigations might be suggested but the sequence would be inappropriate either in terms of invasiveness of the patient or in terms of costs. In addition the examiner also has an opportunity to assess the candidate's ability to logically sequence his/her thought processes in a limited time. This is an additional skill which is essential for later efficient practice. Management is the next skill to be assessed. Here the candidate can range from either killing to curing the patient. The examiner has a duty not to release a candidate on an unsuspecting public who is not properly prepared. This concept can be rela-

tively blurred in a situation where the candidate performs well in the earlier items but in this critical area can be found wanting with potentially disastrous consequences.

Clinical acumen is the overall ability of the candidate to identify the patient's problems and to put the diverse parts of the case together to produce a whole product in terms of problem identification and the ability to solve such problems in overall management terms. Increasingly, the importance of identifying problem-solving ability is being recognized (Barrows & Feltonich, 1987; *Lancet*, 1989; Cassirer, 1992). The inclusion of this item therefore is an essential criterion of clinical competence as the examiners have to attempt to extrapolate from this situation the candidate's ability to perform consistently over a range of such situations or cases. This crucial decision by the examiner has suspect potential if it is made in a global fashion as frequently occurs without the support of the clearly identified previously described nine items. There is evidence to demonstrate that the ability to solve problems will vary from case to case (Elstein *et al.*, 1978). This in turn makes it all the more important to recognize and include this item for valid judgements by examiners. To further assist the examiner in this respect and also to minimize the 'luck of the draw' element for the candidate as far as this is feasible, the difficulty of the case is noted.

Case difficulty

As long cases vary in their degree of difficulty, it is necessary for examiners to establish the relative difficulty of the case under consideration. Not to do so would seriously compromise the validity and reliability of the overall assessment. The case difficulty has been arbitrarily divided into 'standard cases', which would represent a single problem, 'Difficult' cases, which would include up to three problems and 'very difficult' cases, with greater than three problems. However, it will be appreciated that a single problem could amount to a very difficult case. Examiners therefore have to grade difficulty in the context of the case in question and it will be obvious therefore that this decision has to be made prior to commencing the assessment itself.

Grading and marking

It has long been recognized that awarding marks in the long case is unreliable (Wilson *et al.*, 1969) and short training periods for examiners have yielded little improvement (Ludbrook & Marshall, 1971). This is not too surprising as there has been little examiner training on methods that in turn frequently lack objectivity and validity. Prior to the awarding of a mark in the OSLER, a grading system has been adopted. Performance therefore is graded as P + (very good/excellent), P (pass/bare pass) and P − (below pass) for each of the 10 items followed by an overall grade for the complete performance. This is how the vast majority of examiners instinctively make initial assessment decisions. This could be described as an extended criterion-referenced method in that candidates are measured against the criterion for the standard of the clinical assessment in question, i.e. undergraduate or postgraduate. Having decided on an appropriate grade for each individual item and then an overall grade, examiners using the OSLER are then in a position to select an appropriate mark from a designated list of possible marks, each of which is backed up with a stated written mark Profile (Figure 2). Individual examiners, having decided on their overall grade and mark for the candidate, are then in a position to confer with their co-examiner during which time they agree a grade for each of the 10 individual items, an overall grade and finally an agreed mark. This combination of grading and marking amounts to 138 formal decisions being made for any one individual candidate when both examiners are taken into account.

Discussion

The OSLER has now been used for 10 years, during which time important data has emerged. The detailed information that is available following such OSLER assessments has highlighted serious defects in basic clinical skills. This has been noted in both undergraduate and more particularly in postgraduate studies (Gleeson, 1992). The identification of such defects was not too surprising as such findings have been noted in other studies (Maguire & Rutler, 1976; Wiener & Nathanson, 1976; Wray & Friedland, 1983; Sox *et al.*, 1985; Chan Yan *et al.*, 1988). Of even more significance has been the documented immediate marked improvement between two OSLER assessments on 230 postgraduate students within 48 hours (Gleeson, 1995). The time interval was such that only the feedback knowledge of such defects could have influenced the improvement. This finding is all the more important as feedback is regarded as a key step in the development of such skills (Ende, 1983). In a recent *Lancet* commentary the following was stated: "OSLER seems to be a powerful tool for providing feedback and therefore has great potential to increase clinical competence" (Van Der Vleuten, 1996).

What are the advantages of the OSLER in the context of educational assessment criteria? Objectivity is enhanced by prior agreement on what is to be assessed. In any long case there are three variables which are the candidate, the examiners and the patient. Ideally the only variable should be the candidate. Strenuous efforts are being made to standardize patients, particularly through simulation, in North America. For the foreseeable future, however, such standardization will not be practicable or, indeed, for many desirable. In the meantime we must strive to standardize our examiners by assisting them to be as reliable or consistent as possible in their assessments. Recognition of this is already obvious by having two examiners assessing each candidate. The end result, however, is not as perfect as one would anticipate on many occasions. It is not acceptable or good practice for a pair of examiners to confer on the merits of a candidate prior to awarding an individual grade or mark. Many examining authorities increasingly recognize this problem but in some instances have been slow to insist on its implementation. The OSLER, with its increased number of items and fixed structure, will assist individuals of a pair of examiners in their decision making and thus make it easier for examining authorities to insist on the implementation of individual marking prior to examiners conferring. Examiners also require to be conscious that they are assessing broad clinical skills in addition to detailed case-specific skills. There is evidence that such an

The pass mark is 50. Marks should be given in 5s (e.g. 80, 75, 70, 65, 60 etc) in accordance with the following guidelines. Intermediate marks, e.g. 53, 67 should not be used.

EXTENDED CRITERION REFERENCED GRADING SCHEME	EXTENDED MARKING SCHEME
P+	80 Outstandingly clear and factually correct presentation of the patient's history, demonstration of physical signs and organisation of the case management. Clearly a candidate displaying outstanding communication skills and clinical acumen. First class honours. 75 Excellent overall case presentation, communication skills, examination technique and demonstration of the correct facts and physical signs of the case. The candidate may even display outstanding attributes in some but not all measurable criteria. First class honours. 70 Excellent in most respects of overall case presentation, communication skills, examination technqiue and demonstration of the correct facts and physical signs of the case; Also excellent communicator and demonstrates the ability to investigate and appropriately manage the patient with a very well developed clinical acumen. First class honours. 65 Very good overall presentation covering all major aspects; few omissions, good priorities. Very clearly an above average candidate in terms of communication and clinical acumen. Second class honours, division 1. 60 Very good in most respects of presentation and communication but not in all aspects. However, a good solid performance in most areas assessed with a well developed clinical acumen. Second class honours, division 2.
P	55 Good sound overall presentation and communication of the case without displaying any attributes out of the ordinary. The candidate displays an overall adequate standard of examination technique. The patient's problems are identified and a reasonable management outline suggested. 50 Adequate presentation of the case and communication ability. Nothing to suggest more than just reaching an acceptable standard in physical examination and identification of the patient's problems and their management. Clinical acumen just reaching an acceptable standard. Safe borderline candidate who just reaches a pass standard.
P-	45 Poor performance in terms of case presentation, communication with the patient and demonstration of physical signs. Inadequate attempt at a clear identification of the patient's problems. The candidate may display some adequate attributes but does not reach an acceptable pass standard overall. * THE MARK 40 IS NOT USED IN CLINICALS 35 Veto mark. The candidate's performance in terms of case presentation, clinical and communication skills is so poor that the standard required is not even remotely approached. Quite clearly this candidate requires a further period of training.

Examiners should not be hesitant in awarding high or low marks when justified.

Figure 2. The OSLER marking profile.

approach is more reliable (Van Thiel *et al.*, 1991). Is the traditional long case valid, i.e. does it measure what it is supposed to measure in assessing how the student handles the patient as a whole? Clearly there are problems when it comes to measuring history taking and this has already been referred to. By highlighting the construct and content validity, i.e. increasing the number and construct of the items on history taking to be measured, these problems could be expected to be improved by the OSLER. In addition to content validity, overall construct or face validity will be improved by ensuring that all 10 items are formally assessed in a structured manner. It can and does happen that, in existing assessment methods, some item(s) receive undue attention to the exclusion of others. By having a fixed number of items to be measured, examiners will not have to generalize from what they have assessed to what they should have assessed, as is frequently the situation.

Most assessment innovations run into problems of practicality in terms of organizational logistics. The OSLER is singularly unaffected in this respect. Indeed it could be described as organizational friendly as the organization is identical to existing practices. The OSLER could also be described as examiner friendly in that it assists the examiner as an 'aide-memoire' in reminding him/her to consistently cover the same general areas for all candidates to be assessed. The provision of a checklist of items for the

long case was suggested over 20 years ago (Fleming *et al.*, 1974) as a reasonable approach. This in turn makes it candidate friendly in that the assessment will be regarded as more fair by the candidates. There are also a number of other advantages associated with usage of the OSLER that should make it potentially more acceptable. Since it is essentially in line with the traditional long case, it fulfils the innovational criteria already referred to, thus making it more acceptable to more conservative forces. The same number of examiners are required and the examiner time is identical. Structured examinations are frequently criticized by examiners, who feel that their 'independence' is in some way interfered with. In strict educational and institutional objective terms such a stand would be unacceptable: however, the fact remains that such a view is strongly held by a significant number. The OSLER has attractions for such situations in that it allows the examiner to continue to operate as before. The examiner continues to exercise his/her independence, particularly through the item on clinical acumen. However, it will be obvious that the grade in this area will have to correlate with the grades recorded in the other nine items. The feedback potential already referred to is obvious in terms of identification of clinical skills defects.

If one accepts that for clinical assessment to be truly valid it must be patient centred then it would seem reasonable to conclude that the long case is going to be maintained to a greater or lesser extent in the short to medium term at least. Instead of bemoaning this fact, an effort should be made to maximize its potential while at the same time minimizing its faults until such time as a method emerges which will allow full observation of the candidate during the long case. The OSLER as described is both examiner (user) friendly and candidate friendly and could be implemented with relatively little effort. The small extra effort required to implement it would be offset by the more detailed data obtained on candidates' performances rather than the more frequent global data currently available. The perfect method for the assessment of clinical competence has to date not been developed and for reasons of practicality will not be available in the foreseeable future. Until such time, some improvement in the long case is necessary. The OSLER is suggested as that improvement.

Note on contributor

PROFESSOR GLEESON is Associate Professor of Medicine at the Royal College of Surgeons in Ireland and Consultant Physician/Gastroenterologist at James Connolly Memorial Hospital, Dublin. He has had a long standing interest in the assessment of clinical competence and was closely involved in the development of the Objective Structured Clinical Examination (OSCE).

References

BARROWS, H.S. & FELTOUICH, D.J. (1987) The clinical reasoning process, *Medical Education*, 21, pp. 86–91.

CASSIRER, J.P. (1992) Clinical problem-solving—a new feature in the journal, *New England Journal of Medicine*, 326, pp. 60–61.

CHAN YAN, C., GILLIES, J., REUDY,. J.H., MONTANER, J.S.G. & MARSHALL, S.A. (1988) Clinical skills of medical residents: a review of physical examination, *Canadian Medical Association Journal*, 139, pp. 629–632.

COLLINGWOOD, V. (1979) Planning of innovation in higher education, *Programmed Learning and Educational Technology* 16, pp. 8–15.

DOHERTY, E., O'BOYLE, C.A., SHANNON, W., MCGEE, H. & BURY, G.(1990) Communication skills training in undergraduate medicine, *Irish Medical Journal*, 83, pp. 54–56.

ELSTEIN, A.S., SHULMAN, L.S. & SPRAFKA, S.A. (1978) *Medical Problem Solving—An Analysis of Clinical Reasoning* (Cambridge, MA and London, UK, Harvard University Press).

ENDE, J. (1983) Feedback in clinical medical education, *Journal of the American Medical Association*, 250, pp. 777–781.

FLEMING, P.R., MANDERSON, W.G., MATTHEWS, M.B., SANDERSON, P.H. & STOKES, J.F. (1974) Evolution of an examination: MRCP(UK), *British Medical Journal*, 2, pp. 99–106.

FLEMING, P.R., SANDERSON P.H., STOKES, J.F. & WALTON, H.J. (1976) *Examinations in Medicine*. (Edinburgh, Churchill Livingstone).

GENERAL MEDICAL COUNCIL (1993) *Tomorrow's Doctors—Recommendations on Undergraduate Medical Education* (London, General Medical Council).

GLEESON, F. (1992) Defects in postgraduate clinical skills as revealed by the Objective Structured Long Examination Record (OSLER), *Irish Medical Journal*, 85, pp. 11–14.

GLEESON, F. (1995) The effect of immediate feedback on clinical skills using the OSLER, in: *Proceedings of the Sixth Ottawa Conference* (Toronto, University of Toronto Bookstore Custom Publishing).

HAMPTON, J.R., HARRISON, M.J.G., MITCHELL, J.R.A., PRICHARD J.S. & SEYMOUR, C. (1975) Relative contributions of history-taking, physical examination and laboratory investigations to diagnosis and management of medical out-patients, *British Medical Journal*, 2, pp. 486–489.

HARDEN, R.M. & GLEESON, F.A. (1979) ASME Medical Education Booklet No. 8: Assessment of Medical Competence Using an Objective Structured Clinical Examination (OSCE), *Medical Education*, 13, pp. 39–54.

IRWIN, W.G., MCCELLAND, R. & LOVE A.H.G. (1989) Communication skills training for medical students: an integrated approach, *Medical Education*, 23, pp. 387–394.

Lancet (1989) Editorial: Solving problems taking decisions, *Lancet* ii, p. 1306.

LOWRY, S. (1993) *Medical Education* (London, BMJ).

LUDBROOK, J. & MARSHALL, V.R. (1971) Examiner training for clinical examinations, *British Journal of Medical Education*, 5, pp. 152–155.

MAGUIRE, G.P, & RUTTER, D.R. (1976) History taking for medical students: deficiencies in performance, *Lancet* 3, pp. 556–558.

MCMANUS, I.C., VINCENT, C.A., THOM, S. & KIDD, J. (1993) Teaching communication skills to clinical students, *British Medical Journal*, 306, pp. 1322–1327.

MIALL, L.S. & DAVIES, H. (1992) An analysis of paediatric decision-making: how should students be taught?, *Medical Education*, 26, pp. 317–320.

NEWBLE, D.I. (1991) The observed long case in clinical assessment, *Medical Education*, 25, pp. 369–373.

PRICE, J. & BYRNE, G.J.A. (1994) The direct clinical examination: an alternative method for the assessment of clinical psychiatry skills in undergraduate medical students, *Medical Education*, 28, pp. 120–125.

SOX, H.C., MORGAN, W.L., NEUFELD, U.R., SHELDON, G.F. & TONESK, X. (1985) Subgroup, report on clinical skills, *Journal of Medical Education*, 59, pp. 139–147.

STOKES, J. (1972) *The Clinical Examination—Assessment of Clinical Skills*, Medical Education Booklet No. 2. (Dundee, Association for the Study of Medical Education).

VAN DER VLEUTEN (1996) Making the best of the 'long case'; *Lancet*, 347, pp. 704–705.

VAN THIEL, J., KRAAN, H.F. & VAN DER VLEUTEN, C.P.M. (1991) Reliability and feasibility of measuring medical interviewing skills: the revised Maastricht history-taking and advice checklist, *Medical Education*, 25, pp. 224–229.

WEATHERALL, D.J. (1991) Examining undergraduate examinations, *Lancet*, 338, pp. 37–39.

WIENER, S. & NATHANSON, M. (1976) Physical examination: frequently observed errors, *Journal of the American Medical Association*, 236, pp. 852–855.

WILSON, G.M., LEVER, R., HARDEN, R.McG., ROBERTSON, J.L.S. & MACRITCHIE, J. (1969) Examination of clinical examiners, *Lancet* i, pp. 37–40.

WRAY, N.P. & FRIEDLAND J.A. (1983) Detection and correction of house staff error in physical diagnosis, *Journal of the American Medical Association* 249, pp. 1035–1037.