

The structure of medical competence and results of an OSCE

J.C.G. Jacobs^{1*}, E. Denessen², C.T. Postma³

Department of ¹Medical Education and ³General Internal Medicine, University Medical Centre Nijmegen, the Netherlands, ²Department of Education, University of Nijmegen, Nijmegen, the Netherlands, *corresponding author, e-mail: A.Jacobs@educ.umcn.nl

ABSTRACT

Background: Medical competence is a central concept in medical education. Educational efforts in medical training are directed at the achievement of a maximal medical competence. The concept of the structure of medical competence (multidimensional or one-dimensional with strongly interrelated competences) therefore affects the educational developments and assessment procedures.

Purpose: To examine the applicability of a one or more dimensional character of medical competence in student assessments, by analysing the results of 356 students in the history taking station of an objective structured clinical examination (OSCE), in relation to other assessment procedures.

Methods: The performances of 356 students in a history taking station of an OSCE were analysed. Analyses of the checklist scores were aimed at the dimensionality of history taking skills. External criteria were used to test the validity of the scores on the checklist.

Results: The analyses of the scores on the history taking checklist indicated at least five dimensions of history taking skills: the frequency of patient-centred skills, the quality of performance of patient-centred skills, complaint-oriented skills, general social skills, and the provision of procedural information.

Conclusion: Medical competence, as a subject of assessment, can be seen as a multifaceted construct. This study shows that history taking alone might be composed of five different dimensions, suggesting that medical competence

in respect of assessment might be viewed as a multifaceted construct which in that sense has implications for the assessment of medical competence.

INTRODUCTION

Medical competence is a central concept in medical education. Most of the educational efforts in medical training seek to accomplish a growing medical competence of the students, eventually to such a level that they can take up medical practice independently. But although it is one of the central elements in medical education there is no agreement as to what its structure should be. Some¹⁻⁴ see it as multidimensional, encompassing distinct competences, others⁵⁻⁸ consider it to be a one-dimensional construct containing strongly interrelated competences that can hardly be separated. At present, opinion tends to treat medical competence as a complex of knowledge, skills, emotions, values and habits, most of which are seen as strongly interrelated.⁹ An example of medical competence viewed as a multidimensional assessment entity is that of Metz,¹ introduced in 1984. This model of medical competence was constructed on the basis of four separate skills: perceptive, intellectual, motor, and social skills. In this model, *perceptive skills* are defined as the abilities to discern and interpret by perceptive means, various elements indicative of diseases. Central in *intellectual skills* is cognitive functions, more precisely not just the theoretical knowledge itself, but its application, for example knowing the right questions when taking a history. *Motor skills* indicate the ability to conduct

the appropriate manual procedures in medical examinations. *Social skills* refer to communication and interactions both with patients and with other healthcare professionals. One of the important advantages of this four-dimensional model is that it very adequately enables the construction of assessment procedures because the observations and scores can be based on these four different skills. In this way the judgement of the achievements of the students and the formative feedback could be more detailed.

In 1985, Norman³ undertook a methodological review of the models of competence that were then available. He concluded that at that point no single model could adequately define the prerequisite knowledge, skills and attitudes required for a competent physician. Therefore, he introduced the categorisation of clinical competence in clinical and technical skills, knowledge and understanding, interpersonal attributes and capabilities in problem-solving and clinical judgement. So, taking these together, medical competence can be thought of as a multifaceted construct whereby the various contributing elements are interdependent and overlapping and should be assessed as such.^{1,2}

At present there are indications that medical competence could be assessed by using global ratings, as these would be as reliable and valid as more comprehensive checklists.¹⁰ Although this might be true in general for giving an overall impression of a certain clinical skill, the question remains whether such a global rating is precise enough to detect shortcomings in the learners sufficiently to warrant precise feedback, corrections or educational changes. So, for purposes of assessment, it is of great importance to obtain more insight into the question whether medical competence should be seen as being a one-dimensional entity and tested as such, or as a more dimensional model, built up of multiple recognisable elements and justified in testing developing medical competence of students.

The aim of this study was to examine if a one or more dimensional character of medical competence could be uncovered in the assessment of medical competence of undergraduate medical students. If analysis of the results were to point towards one dimension, this would support the validity of the one-dimensional construct in medical competence testing. But if more dimensions can be discerned in history taking skills, a 'multifaceted' model of medical competence testing would seem more appropriate and global ratings would not seem detailed enough to cover the assessment of competence.

METHODS

Participants

The results of 356 students (160 male (45%) and 196 female (55%) students) at the history taking station of an objective

structured clinical examination (OSCE) were analysed. The students participated in the OSCE of the practical clinical training module 1, just before the start of their clerkships, in the fifth year of their study. (In the Netherlands the medical undergraduate curriculum usually contains four years of mainly preclinical education, followed by two years of mainly clerkships).

Description of the OSCE

The OSCE consists of twelve stations with five minutes for each station. A distinction was made between process- and product-centred stations: trained expert-observers observed three process-centred stations (one of which a history taking station) with checklists, while the results of nine product stations (for example interpretation of ECG) were rated on the final product. The OSCE was based on the four skills (perceptive, intellectual, motor and social) of Metz's model of medical competence.

Procedure

In five minutes, students completed part of a history taking of one of the standardised patients. Four standardised patients carefully trained for their roles participated in this study. The validity of simulated patients has been demonstrated before.¹¹ Two expert-observers were trained in the use of the history taking checklist. Their training started with an explanation of the content of the items, followed by a hands-on training with video material. The training aimed to accomplish a 90% agreement in scoring behaviour that was obtained in two half-day sessions.¹² Every four weeks one of the two observers participated in the OSCE, mostly real time by a one-way screen, sometimes from a videotape.

The checklist

The history taking checklist consisted of 24 items: 8 were directed at social skills and 16 addressed intellectual skills. These intellectual skills, especially items 15 to 19, focussed on the medical content in a general way. They are directed at the achievements of the students to gather medical information irrespective of case content. Students did not know the content of the list, but the checklist items reflected the goals of the training activities. The construction of the checklist was based on the MAAS-R (Revised Maastricht history taking and advice checklist)¹³ that has proved to be a valid instrument in assessing the essential elements in history taking. To guarantee further the content validity of the checklists, the construction was supervised by a steering group of medical experts from different disciplines. The checklist used in this study is directed at the basic communication skills and at the more general elements of medical data gathering, not directly at the medical content of the case histories. In *table 1* the items of the checklist history taking are shown. The items 1, 2, 3, 13, 20, 21, 22 and 24 referred to social skills; the remaining items were aimed

Table 1
Checklist 'history taking' with frequency distribution of items

	YES	NO	OPEN		
1. Introduction	352	3	1		
	GOOD	MODERATE	BAD		
2. Explanation of position (clerkship, training)	134	192	30		
	YES	NO			
3. Proposition of plan	103	253			
	GOOD	MODERATE	BAD	NOT SHOWN	OPEN
4. Questions about reasons for encounter	334	22	0	0	
5. Exploration reason for encounter	225	104	20	6	1
6. Questions about expectation of consultation	17	14	2	312	11
7. Questions about presumptions, ideas about complaints	68	18	4	261	5
8. Exploration of impact complaint on daily life	25	20	6	297	8
9. Information about self-help and results	68	35	2	245	6
10. Reaction of social environment on complaint	20	4	0	324	8
11. Patient's activities out of house (resumption of work)	7	3	0	337	9
12. Recapitulation of history so far	144	94	25	92	1
13. Use of common, understandable language	305	46	3	1	1
14. Checking information in recapitulation	143	81	27	102	3
15. Questions about medical history	78	184	29	55	10
16. Getting a clear view of complaint	140	195	20	1	0
17. Questions about period of complaint (origin, development)	198	135	17	3	3
18. Questions about course and duration of present complaint	168	170	16	2	0
19. Questions about noted correlation with other symptoms	70	241	42	3	0
20. Results in relation to complaints and expectations	12	38	7	298	1
21. Checking if reasons for encounter have been discussed	3	8	3	342	0
22. Creation of space for patient to express himself/herself	246	97	12	n.a.	1
23. Proper winding up consultation/making appointments	50	100	32	173	1
24. Showing empathy	188	151	15	n.a.	2

n.a. = not applicable.

at intellectual skills. Response categories on the respective items were 'yes-no' or 'good – moderate – poor – not shown'. To validate the checklist, results of the detailed checklist were compared with the global appreciation of history taking skills by each group's own tutor. The tutor, a medical specialist, supervises a group of 12 students very closely for four weeks during training sessions and coaches them in clinical skills including history taking, together with a psychologist. At the end of this period the tutor gives each student: 1) a global grade for history taking skills, based on the overall performance during the entire four weeks, 2) a grade for medical knowledge and 3) one overall grade.

These grades were used as external criteria in the study. Another external criterion is the mean score of the student in the nine product stations of the OSCE, these stations (for example interpretation of ECG) were scored on their final product and not observed.

Analysis

First, frequency distributions of the items were inspected. Next, factor analyses were conducted to assess the dimensionality of the test scores. Furthermore, correlations of the scores on the history taking checklist with external criteria were calculated to assess the external validity of the scores on the history taking checklist.

RESULTS

Table 1 shows the frequency distributions of the scores on the items of the observation list for history taking. Frequency distributions revealed that for several items the frequencies of the scoring category 'not shown' were quite high (table 1). For the items 6 to 11, 20, and 21, the percentage of students who did not pay attention to these items was above 50%. Factor analysis of the scores did not show an interpretable solution (percentages explained variance for the

first three factors were 15.04, 10.08 and 7.33% respectively). The reliability of the scores on all 21 items (Cronbach's alpha) was 0.66 with a mean inter-item correlation of 0.10. For the two distinct groups of items, social skills (8 items) and intellectual skills (16 items), the reliability of the scores (Cronbach's alpha) was 0.50 (mean inter-item correlation was 0.13) and 0.58 (mean inter-item correlation was 0.09). The poor scalability of the items may be due to the high frequencies of the category 'not shown'. Because of this, we found it useful to further explore the meaning of the scoring category 'not shown'. The question that we addressed was: is a 'not shown' activity worse than a badly performed activity? Therefore the frequency of performed activities and the relation between the number of performed activities and the level of performance was explored.

Intellectual skills

Because of the content of items 6 to 11, and due to the fact that the observations showed a large number of students not performing these items, we initially focussed on these items, each of which referred to a patient-centred intellectual skill. To further explore the scores on these items, the number of items the students actually performed was counted (range: 0 to 6). The frequency distribution of the number of demonstrated patient-centred items is shown in table 2. Table 2 shows that a total number of 165 students did not show any of the patient-centred intellectual items, and that a total number of 191 students demonstrated one or more of those items. None of the students demonstrated all six items. To address the issue of the relationship between the number of items performed and the level of performance, the mean item score for students who actually demonstrated any item was computed. The mean score of the quality of performance of the 191 students who performed at least one of the items was 2.63 (sd = 0.48).

For the 191 students who demonstrated one or more patient-centred items, the correlation between the number of demonstrated items and the quality of performance was -0.08 (df = 189, p=0.30). This means that there is no rela-

tionship between the number of demonstrated items and the quality of performance. This result implies that the scoring categories 'not shown, poor, moderate, and good' can not be perceived as a one-dimensional interval, or even an ordinal, scale. The number of performed activities does not indicate the level of performed activities. Therefore, the two variables were used separately in further analyses.

For the remaining intellectual items (complaint-oriented items) factor analysis was applied to assess the dimensionality of the item scores. Students who did not perform one of these activities were removed from the analysis. Table 3 shows the pattern coefficients and communalities of the five intellectual items. A one-dimensional solution explained 33.9% of item variance. Factor analysis revealed one dimension in these complaint-oriented items. This dimension can be labelled as the 'complaint orientation of history taking' (Cronbach's alpha = 0.71).

Social skills

For the presumed remaining social skills, factor analysis was applied to assess the dimensionality of the item scores. Students who did not perform one of these activities were removed from the analysis. Table 4 shows the pattern coefficients and the communalities of five social skills. A two-dimensional solution explained 48.1% of item variance. Factor analysis showed that two dimensions of skills could be discerned. The first dimension could be labelled 'general social skills' (Cronbach's alpha = 0.61), the second dimension 'providing procedural information to the patient' (abbreviated as: 'procedural information'; Cronbach's alpha = 0.73). Items hardly differentiating between students (items 1, 4, 20, 21) and items which did not, or not enough, relate with other items (items 5, 12, 14, 23) were kept outside the analysis. For each scale, scores were calculated by computing mean scores for items referring to one of the distinguished dimensions. Descriptive statistics of the scales are summarised in table 5.

To examine the interconnectedness of scores on these separate scales, correlations between the scale scores were calculated (table 6). Except for the correlation between 'complaint orientation' and 'general social skills', correlations between the scales were quite low.

Criterion-related validity of history taking skills

For validation of the identified dimensions in history taking skills, scale scores were correlated with other performance data: 1) the tutor's global appreciation of history taking skills, 2) knowledge estimate by tutor, 3) the overall grade by tutor and 4) the mean score on the product stations of the OSCE (table 7). Correlations between these scores were quite low, which means that the scores on history taking skills were weakly related to these external criteria.

Table 2
Frequency distribution of the total number of demonstrated patient-centred intellectual items (items 6 through 11)

VALUE	FREQUENCY	%
0	165	46.4
1	108	30.3
2	51	14.3
3	27	7.6
4	3	0.8
5	2	0.6
6	0	0

Table 3

Factor analysis of complaint-oriented items with communalities (h^2) and percentage explained item variance ($n=286$)

	F ₁	H ²
15. Questions about medical history	0.54	0.29
16. Getting a clear view of complaint	0.59	0.35
17. Questions about period of complaint (origin, development)	0.68	0.46
18. Questions about course and duration of present complaint	0.67	0.45
19. Questions about noted correlation with other symptoms	0.38	0.14
		33.9% explained item variance

F₁ = complaint orientation of history taking.

Table 4

Factor analysis of social items with communalities (h^2) and percentage explained item variance ($n=351$)

	F ₁	F ₂	H ²
2. Explanation of position (clerkship, training)	0.14	0.81	0.78
3. Proposition of plan	-0.02	0.77	0.80
13. Use of common, understandable language	0.41	0.06	0.46
22. Leaving room for patient to express himself/herself	0.60	0.06	0.56
24. Showing empathy	0.78	-0.02	0.57
			48.1% explained item variance

F₁ = general social skills, F₂ = providing procedural information to the patient.

Table 5

Descriptive statistics for five scales of history taking skills

SCALE	N	M	(SD)
Intellectual skills			
Patient-centredness/frequency (items 6-11)	356	0.88	(1.03)
Patient-centredness/quality (items 6-11)	191	2.63	(0.48)
Complaint orientation (items 15-19)	355	2.31	(0.39)
Social skills			
General social skills (items 13, 22, 24)	356	2.64	(0.38)
Procedural information (items 2, 3)	356	1.94	(0.69)

Theoretically, the scores on all scales, except for the scores on patient-centredness/frequency, can range from 1 to 3.

Table 6

Correlations between the five scales of history taking skills

	PATIENT-CENTREDNESS/ FREQUENCY	PATIENT-CENTREDNESS/ QUALITY	COMPLAINT ORIENTATION	GENERAL SOCIAL SKILLS	PROCEDURAL INFORMATION
Intellectual skills					
Patient-centredness/frequency	---				
Patient-centredness/quality	-0.08	---			
Complaint orientation	0.24*	0.12	---		
Social skills					
General social skills	0.17*	0.13	0.40*	---	
Procedural information	0.12*	-0.07	0.12*	0.12*	---

* $p < 0.05$.

Table 7

Correlations between the five scales of history taking skills and four external criteria (n=593, except patient-centredness/quality (n=191))

SCALES OF HISTORY TAKING SKILLS	TUTOR'S GLOBAL APPRECIATION OF HISTORY TAKING SKILLS	KNOWLEDGE ESTIMATE BY TUTOR	OVERALL GRADE (BY TUTOR)	MEAN SCORE ON PRODUCT STATIONS OF THE OSCE
Patient-centredness/frequency	0.02	0.02	0.00	-0.12
Patient-centredness/quality	0.16	0.10	0.14	0.10
Complaint orientation	0.17*	0.26*	0.13	-0.10
General social skills	0.29*	0.06	0.17	0.04
Procedural information	0.06	0.05	0.03	-0.14*

* $p < 0.05$.

DISCUSSION

The students' results in the history taking station of the OSCE at the start of the clerkships were analysed in order to gain more insight into the structure of medical competence as a subject of assessment. First, the structure of the complete checklist was examined with a factor analysis. This revealed a non-interpretable solution, because of the fact that for several items of the checklist for this OSCE station, the frequencies in the scoring category 'not shown' were quite high. Apparently, students tended to overlook the patient-centred elements (items 6-11 and 20-21). This might have been partly caused by the pressure of time, but the instruction to the students was to address only the specific history of the present complaint. Furthermore, the students were trained to pay attention to the issues of the items 6 to 11 and 20 and 21 in this part of the history taking process. And most students felt, as expressed during the assessment, that they completed this part of the history taking with the standardised patient.

The correlation between the number of performed activities and the level of performance of these activities was close to zero. This implied that the scoring categories 'not shown, poor, moderate and good' could not be perceived as a one-dimensional interval or even an ordinal scale. Therefore the checklist was analysed in parts.

The first analysis included the patient-centred intellectual items. For a better understanding, two new uncorrelated variables were introduced: the 'frequency of patient-centred intellectual items' and the 'quality of performance of patient-centred items'. Factor analysis of the remaining intellectual items showed that these could be represented by one factor labelled 'complaint orientation of history taking'. Subsequently, factor analysis of the items concerning social skills revealed two factors renamed: 'general social skills' and 'providing procedural information to the patient'.

The correlations between the scale scores of these five scales, covering an important part of the history taking

skills of the students, were low. This suggests that it is very unlikely that any one of these scales might be used to represent history taking, and even more unlikely that one such scale could represent medical competence. Apparently, the content of the history taking checklist represents very different domains, which underlines a 'multifaceted' model of medical competence. Nevertheless, the study offered no support for Metz's model¹ of four dimensions; the structure of history taking skills seems to be more complex.

In the interpretation of the results of this study the validity of the checklist of the history taking skills is an important issue. The checklist was based on a valid instrument in assessing the essential elements in history taking,¹³ further scrutinised and adapted by experienced clinicians to strengthen its validity in the given test procedure. The checklist history taking is directed at the basic communication skills including the skills to collect medical information, independent of case-specific content. To ensure scoring is as uniform as possible, one of the two specially trained expert-observers observed the history taking stations.¹²

To examine the external validation we compared the results on the five scales of the checklist with the global appreciation of history taking skills by the tutors. These correlations were low. First, this low correlation might be due to a low content validity of the checklist. However, the checklist was put together with great care. Furthermore, the assessment procedures were aimed at skills that are specifically taught in our medical curriculum: social skills and intellectual skills in history taking are given special attention. Second, the quality of the observations by the tutors might be less thorough than expected. It is known that personal observations of a teacher in close and intensive contact with students introduce subjective elements (e.g. halo effects) in tutor judgements.^{14,15} That was one of the very reasons for introducing objective structured clinical examinations in 1979.¹⁶ Another explanation for the low correlation might be the difference in time; a four-week period is compared

with a five-minute station in the OSCE. In conclusion, the observation that the results in the history taking station of the objective assessment were not in accordance with those of the tutors does not necessarily indicate shortcomings in the OSCE.

The OSCE described here was scheduled at the end of a training period and had to assess the general clinical competence of the students at that particular moment. Of course no final judgement of their competence in history taking may be inferred by this assessment because generalisations about a person's competence can not be based on one or two patient encounters.¹⁷ Moreover, earlier studies have demonstrated that clinical performance is, apart from content specificity, very variable and unpredictable.¹¹ By comparing the results of this large group of students and the relation of these results with other outcomes, we sought to get an impression of the structure of the history taking skills of these students. It was not the purpose of this study to find differences between individuals.

The ideal final assessment of a student's medical competence will be an appraisal of his daily work in clinical practice. The final examination at our medical school consists of a four-week internship. In this internship the responsibilities of the undergraduate student resemble those of a resident, but the goals of this internship are different from a residency.

To summarise, the structure of history taking skills and medical competence is complex. An adequate theoretical basis could have an important impact on the development of education and assessment programmes both in undergraduate and in postgraduate continuing medical education. If the various scales as identified in our study indeed make up the competence of history taking, this should be taken into account, both in training and assessing this skill, also during the internships. Regarding medical competence, one might conclude that medical competence comprises at least these five elements. To accomplish proper assessment, the various components determining medical competence have to be clear. But, based on the presented results, it is clear that medical competence should be tested in a more detailed way. We recommend a structured assessment supported by a convenient checklist, with or without certain weights reflecting the educational goals of the specific programme. An interesting follow-up research design directed at the assessment of history taking would be to ask the tutors for five global marks on 'patient-centredness/frequency', 'patient-centredness/quality', 'complaint orientation', 'general social skills' and 'providing procedural information to the patient' for each student and compare these marks with the scales of the items on the checklist. Possibly this would support the generation of a short, comprehensive checklist and provide more insight into the relative importance of each scale of history taking.

REFERENCES

1. Metz JCM. Medische competentie. Een onderzoek naar de betrouwbaarheid en de validiteit van het gestructureerd klinisch examen [proefschrift]. (Medical competence. A study on the reliability and the validity of the structured clinical examination [dissertation]) [with English summary]. Nijmegen, 1984.
2. Norman GR. Defining competence: a methodological review. In: Neufeld VR, Norman GR (eds). *Assessing clinical competence*. New York: Springer, 1985:15-35.
3. Verhulst SJ, Colliver JA, Paiva REA, Williams RG. A factor analysis study of performance of first-year residents. *J Med Educ* 1986;61:132-4.
4. Newble DI. Assessment of clinical competence. State of the art. In: Bender W, Hiemstra RJ, Scherpbier AJJA, et al (eds). *Teaching and assessing clinical competence*. Groningen: BoekWerk Publications, 1990:23-7.
5. Miller GE. The assessment of clinical skills/competence/performance. *Acad Med* 1990;65(suppl 9):S63-7.
6. Vleuten CPM van der, Norman GR, Graaf E de. Pitfalls in the pursuit of objectivity: issues of reliability. *Med Educ* 1991;25:110-8.
7. Vleuten CPM van der, Newble DI. How can we test clinical reasoning? *Lancet* 1995;345:1032-4.
8. Vleuten CPM van der, Scherpbier AJJA, Dolmans DHJM, Schuwirth LWT, Verwijnen GM, Wolfhagen HAP. Clerkship assessment assessed. *Med Teach* 2000;22(6):592-600.
9. Epstein RM, Hundert EM. Defining and assessing professional competence. *JAMA* 2002;287(2):226-35.
10. Regehr G, Freeman R, Robb A., Missiha N, Heisey R. OSCE Performance evaluations made by standardized patients: comparing checklist and global rating scores. *Acad Med* 1999;74(suppl 10):S135-7.
11. Norman GR, Tugwell P. A comparison of resident performance on real and simulated patients. *J Med Educ* 1985;19:344-56.
12. Jacobs JCG, Hein PR, Lutterman JA, Metz JCM, Postma CT. Veranderingen in de anamnesevaardigheden van co-assistenten, voor en na het co-assistentenschap Interne Geneeskunde. (Changes in history taking skills of interns, before and after the clerkship Internal Medicine). In: Cate ThJ ten, Dijkers JH, Houtkoop E, et al (eds). *Proceedings Gezond Onderwijs Congres*. Houten/Diegem: Bohn Stafleu van Loghum, 1996:191-7.
13. Thiel J van, Kraan HF, Vleuten CPM van der. Reliability and feasibility of measuring medical interviewing skills: the revised Maastricht history-taking and advice checklist. *Med Educ* 1991;25:224-9.
14. Coulson CC, Kunselman AR, Cain J, Legro RS. The mentor effect in student evaluation. *Obstet Gynecol* 2000;95(4):619-22.
15. Daelmans HEM, Verheul CC, Vleuten CPM van der. Interbeoordelaarsbetrouwbaarheid in het stationsexamen van het algemeen co-assistentenschap. (Inter-observer reliability at the station history taking of the Objective Structured Clinical Examination in the general clerkship.) *Bulletin Medisch Onderwijs* 1999;18(3):102-9.
16. Harden RM, Gleeson FA. ASME Medical Education Booklet No.8: Assessment of clinical competence using an objective structured clinical examination (OSCE). *Med Educ* 1979;13:41-54.
17. Wass V, Vleuten C van der, Shatzer J, Jones R. Assessment of clinical competence. *Lancet* 2001;357(24):945-9.