



Assessing Competences in Medical Students with a Newly Designed 360-Degree Examination of a Simulated First Day of Residency: A Feasibility Study

Sigrid Harendza^{1*}, Pascal O Berberat² and Martina Kadmon³

¹Faculty of Medicine, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

²TUM School of Medicine, Technical University of Munich, Munich, Germany

³Faculty of Medicine, University of Augsburg, Deanery, Augsburg, Germany

*Corresponding author: Sigrid Harendza, Faculty of Medicine, University Medical Center Hamburg-Eppendorf, III. Medizinische Klinik, Martinistr. 52, D-20246 Hamburg, Germany, Tel: +49-40-7410 53908; Fax: +49 40 7410 40218; E-mail: harendza@uke.de

Rec date: Aug 25, 2017; Acc date: Aug 28, 2017; Pub date: Aug 31, 2017

Copyright: © 2017 Harendza S, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution and reproduction in any medium, provided the original author and source are credited.

Abstract

Background: Many medical schools worldwide are in the process of curricular change towards competence-based undergraduate medical education to ease the transition to residency. Aspects of patient management, interprofessional interaction and the concept of entrustable professional activities are in the focus of medical educators. Such curricular changes require an assessment with valid and reliable measurements of competences feasible for large numbers of medical students.

Methods: Seventy medical students from three German medical schools (Hamburg, Oldenburg, and Munich) participated in our newly designed 360-degree examination based on selected competences relevant for first year residents. A consulting hour with five simulated patients is followed by a patient management phase of 2.5 hours with interprofessional disturbances, followed by a 30-minutes handover of the patients to a resident. Different competences are assessed by the supervisors, simulated patients, nurses, residents, and by the participants themselves. All participants and research assistants evaluated the assessment regarding aspects of organization and content.

Results: All participants, assessors, and research assistants were satisfied with the process, technical equipment, and organization of the assessment. All assessors and research assistants stated that their respective training prepared them well for their role in the assessment. The participating students felt satisfied with the content of the assessment. They considered the patient cases to be very realistic. While there was no difference between students in their final year and students below semester 11 with respect to feeling confident during history taking and handover, final year students felt significantly more confident ($p=0.02$) during the patient management phase.

Conclusion: It is feasible to implement a competence-based 360-degree assessment for a large group of students. Further analysis of the data will provide evidence whether students from medical schools with different undergraduate curricula perform differently with respect to competences relevant for first year residents.

Keywords: Competences; Competence-based assessment; Diagnostic procedures; Empathy; Entrustable professional activities; Interprofessional communication; Simulated patients; Patient handover

Introduction

The goal of undergraduate medical education is students' acquisition of knowledge, skills and attitudes to be prepared for the complex tasks they will encounter in their daily routine as physicians [1]. Currently, undergraduate and postgraduate medical education develop toward outcome and milestone-oriented education, which forms the basis of the concept of competency-based medical education (CBME) and enables supervisors to monitor the progress of their students [2]. The concept of CanMEDS, developed as a competence-based framework with professional roles for postgraduate medical training [3] has served meanwhile also as a prototype for the development of competence-based undergraduate medical curricula

[4]. The main intention of competence-based medical training is to enable students to apply their knowledge and skills to solve clinical problems [5]. Some medical competences, e.g. oral case presentation, a competence, which is needed for every ward round and patient handover, can be used as teaching and assessment tools in competence-based medical education [6]. However, most types of competence-based assessment in undergraduate medical education such as mini-clinical exercises (mini CEX [7]) or Direct Observation of Procedures (DOPS [8]) do not provide enough standardization due to the changing clinical context. Therefore, a claim for better standardization and constructive alignment of competence-based education and assessment has been made [9,10].

In 2013, Wijnen-Meijer et al. developed a competence-based assessment called UHTRUST (Utrecht Hamburg Trainee Responsibility for Unfamiliar Situations Test) simulating the first day of a beginning resident and provided an argument-based approach to its validation [11]. This assessment was based on facets of competence,

which were gathered, assessed, and ranked with respect to their importance for a beginning resident by physicians and physician educators from the Netherlands and Germany [12,13]. Students from a vertically integrated curriculum in the Netherlands reached significantly higher scores than students from Hamburg for the competence “active professional development” comprising ‘reflection’ and ‘asking for feedback’ and for the entrustable professional activity “solving a management problem” [14]. Furthermore, the Dutch students ordered significantly less laboratory and radiology tests compared to the German students with no difference in the diagnostic accuracy for the five simulated patients [15]. We repeated the previous ranking study performed among physician educators in the Netherlands and Germany with 25 essential competences that enable trust in graduates [13,16]. In our study including 202 physicians from three German medical schools with different undergraduate curricula, we found 90% congruency with respect to the top 10 competences with the Dutch cohort [17].

Our aim was to extend the competence-based assessment developed by Wijnen-Meijer et al. [11] based on the top 10 competences ranked by German physicians to a 360-degree examination simulating the first day of residency. We wanted to explore whether it was feasible to implement such a type of competence-based examination for a greater

cohort of undergraduate medical students from three different medical schools.

Material and Methods

Design of the 360-degree competence-based assessment

This 360-degree assessment for undergraduate medical students close to their graduation was based on the top 10 competencies identified by Fürstenberg et al. [17] and comprised an authentic simulation for medical students in the role of a beginning resident on a busy first day in hospital developed earlier [11]. The assessment includes five phases, which can be completed by 30 participants per day, six groups of five students each (Figure 1). After a 20-minutes briefing of five participating students per round (phase 1), each participant has a 5-minute meeting with his or her supervisor. This is followed by a consultation hour (phase 2), which is videotaped, where every participant meets the same five simulated patients in a different order for ten minutes each. Afterwards, during a third phase of 2.5 hours (phase 3), the candidates can order medical tests from a virtual laboratory and radiology department using a laptop (both are also available by phone) and gather further information from the internet.

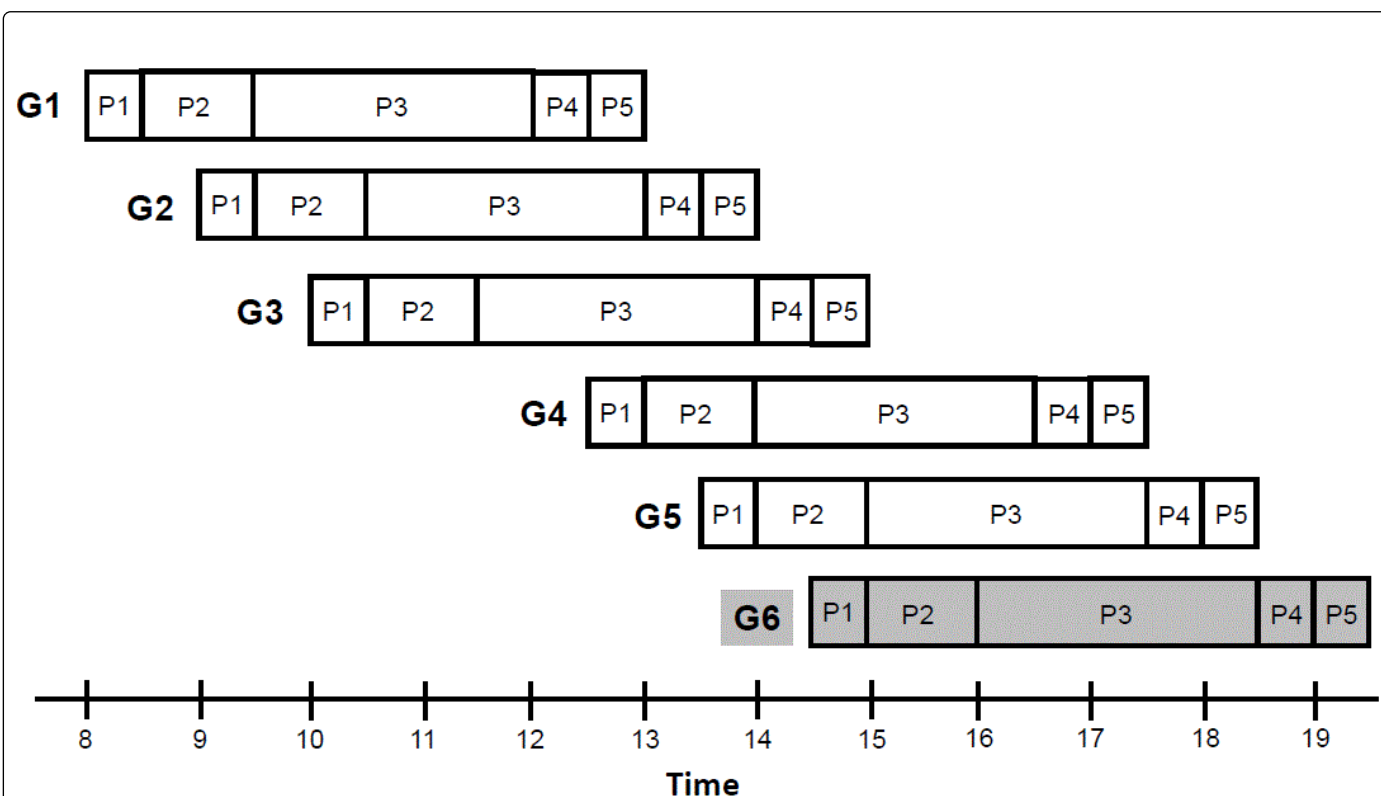


Figure 1: Setting of the competence-based 360-degree assessment per day for 30 participants-G: group, P1 (phase 1): briefing, P2 (phase 2): consultation hour, P3 (phase 3): patient management phase with disturbances, P4 (phase 4): handover, P5 (phase 5): debriefing; G6 is depicted with grey color because our study only included five groups.

Based on the results, participants can develop management plans for the five patients. During this phase, the participants face other tasks, i.e. usual disturbances occurring on a ward like interprofessional interactions with nursing staff and telephone calls from simulated colleagues. Participants can call their supervisor at any time during

this phase. Halfway through this phase, participants have a brief meeting with their supervisor to discuss questions and progress.

Phase 4 (30 minutes) includes a handover of the five patients to a real resident who will take over the next shift. During this phase, which

is also videotaped, the supervisor is present to assess the participant but does not interfere in the discussion of the five patients between participant and resident. The assessment ends with a 30-minutes audio recorded debriefing of each group of five participants (phase 5) to evaluate their experiences during the simulation.

Content

The patient cases used in this assessment were: 1) a 42-year-old woman with palpitations (atrial fibrillation), 2) a 53-year-old man with fatigue and bloody cough (granulomatous polyangiitis), 3) a 58-year-old woman with severe abdominal pain (perforated sigma diverticulitis), 4) a 54-year-old man with loin pain (covered perforated infrarenal aortic aneurysm), 5) a 36-year-old woman with rheumatoid arthritis and fever (herpes zoster). Specific laboratory or radiology reports were provided to the participants at request in a slightly shortened but realistic time (e.g. if an ECG was ordered, it would be provided within 15 minutes, a CT scan would take one hour and a urinary culture could be ordered if participants felt the necessity to do so, but no result would be provided).

The disturbances during the patient management phase were as follows. The nurse calls and tells the participant that the 36-year old woman has a very low blood pressure approximately 30 minutes after the patient management phase started. Fifty-five minutes into the patient management phase a simulated resident calls the participant and explains that she cannot cover the afternoon shift today because her grandmother died. The nurse approaches the participant personally about 75 minutes into the patient management phase and reports that the 53-year old man coughed up some more blood and a drop in oxygen saturation.

One-hundred minutes after the patient management phase began, the nurse calls again and tells the participant that a patient from the ward has a potassium of 5.8 mmol/L. Approximately 115 minutes into the patient management phase, the nurse calls again and reports that the 36-year-old woman developed a rash on her left shoulder. In the middle of the patient management phase, the supervisor surprisingly visits the participant and talks to him/her for about five minutes.

Setting and equipment

The assessment took place during three days on two floors of the seminar building (Campus Lehre) of the University Hospital Hamburg-Eppendorf. The following rooms are needed: five consultations rooms, one recreation room for the actors, one room for the supervisors, one room for the nurses, one room for the residents, one room for simulated residents and other research assistants, three patient management rooms for five students each, one room for the laboratory and radiology team, five handover rooms, one room for telephone calls, one room for breaks and catering.

Each supervisor, participant (15 simultaneously), nurse, resident, the laboratory department, the radiology department, and three simulated residents have to be equipped with a cellular phone. Ten video cameras (one for each consultation room and one for each handover room) and one audio recorder for the debriefing are needed.

During the patient management phase every participant is equipped with a laptop including WLAN (15 simultaneously) and six further laptops with printers were necessary for the laboratory and radiology results (results were ordered electronically and delivered to the participants in an analog way).

Instruments used in the 360-degree assessment

In total, 13 different instruments were used for the assessment (Table 1) and one instrument for its evaluation.

Self-Assessment	
NEOFFI	Assessment of five personality domains
IRI	Assessment of individual differences in empathy
MAQ	Assessment of medical attitudes
STRAIPER	Assessment of strain perception
STRAISY	Assessment of strain symptoms
VAS	Assessment of the quality of history-taking per simulated patient
GKCSAF	Assessment of communication skills
PEF	Assessment of clinical reasoning steps after patient encounters
SDQ	Social desirability questionnaire
Simulated Patients	
CARE	Assessment of consultation and relational empathy
GKCSAF	Assessment of communication skills
Supervisors	
FOC	Assessment of facets of competence
Logbook	Documentation of all telephone or personal encounters with participant
EPA	Assessment of entrustable professional activities
Nurses	
FOC	Assessment of facets of competence
Residents	
FOC	Assessment of facets of competence

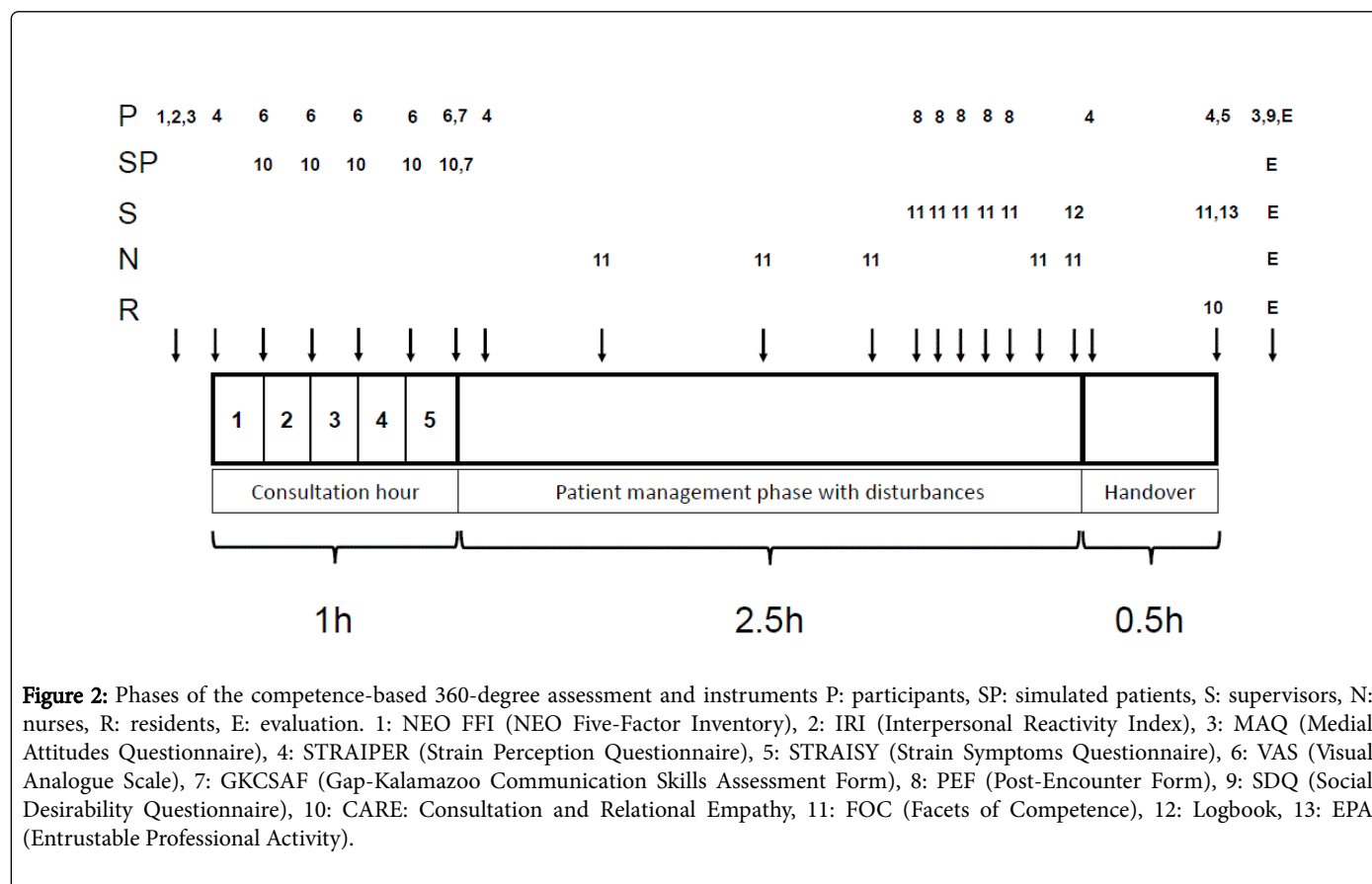
NEO FFI: NEO Five-Factor Inventory; IRI: Interpersonal Reactivity Index; MAQ: Medical Attitude Questionnaire; STRAIPER: Strain Perception Questionnaire; STRAISY: Strain Symptoms Questionnaire; VAS: Visual Analogue Scale; GKCSAF: Gap-Kalamazoo Communication Skills Assessment Form; PEF: Post-Encounter Form; SDQ: Social Desirability Questionnaire; CARE: Consultation and Relational Empathy; FOC: Facets of Competence; EPA: Entrustable Professional Activity

Table 1: Instruments used for the 360-degree competence-based assessment.

The 360-degree assessment included instruments for self-assessment and instruments used by the simulated patients, by the supervisors, by the nurses, and by the residents. Figure 2 displays, which instruments were used in which phase of the assessment. Whenever possible, validated German versions of the questionnaires were used. For questionnaires, which are not validated in German, Cronbach's alpha will be calculated in the further analysis. Before the briefing round, students completed the NEO FFI (NEO Five-Factor Inventory) which measures five basic personality factors [18]. Four different aspects of empathy were measured using the IRI (Interpersonal Reactivity Index, modified according to Davis [19]: social functioning, self-esteem,

emotionality, and sensitivity to others. Medical attitudes were assessed with the MAQ, a Medical Attitude Questionnaire developed according to the professionalism assessment scale (PAS) by Klemenc-Ketis and Vrecko [20]. Every time, when participants entered a new stage of the assessment, they filled out a Strain Perception Questionnaire (STRAIPER) modified according to Müller and Basler [21]: at the beginning of the consultation hour, at the beginning of the patient management phase, at the beginning of the handover and after the handover. After each consultation, the participants filled out a Visual

Analogue Scale (VAS) on their self-assessed quality of the history-taking while the simulated patients assessed the participants' consultation and relational empathy after every encounter with the CARE questionnaire [22]. After the last conversation with a simulated patient, every participant and every simulated patient completed a Gap-Kalamazoo Communication Skills Assessment Form (GAKCSA) [23] for this final encounter, which assesses communication skills in general.



During the patient management phase, participants completed a Post-Encounter Form (modified according to Durning et al. [24]) for each simulated patient. Meanwhile the supervisors filled out a Form for Facets of Competence (FOC, modified according to [11]) for their assessment of the participants' management of every patient and documented every telephone encounter with a participant in a logbook. The nurses completed an FOC form for each single disturbance per participant. At the end of the patient management phase, the nurses additionally filled out one FOC form per participant, assessing all facets of competence independently of any disturbances.

After the handover, residents and supervisors completed a general FOC form per participant and the supervisors also filled out a form to assess to what degree they would hand over responsibility for certain entrustable professional activities (EPA) [25]. The EPAs were based on the facets of competences defined as relevant for first year residents. They were partly adapted from the UHTRUST-project and partly extended to reflect all facets of competences defined in the present study. Before the debriefing round, participants filled out a strain symptoms questionnaire (STRAISY), the modified MAQ with respect to their behavior during the assessment and a Social Desirability

Questionnaire [26] to control for socially desirable answers in general as well as an evaluation form with questions regarding the general organization and content of the 360-degree assessment. The debriefing was organized like a focus group with semi-structured questions for discussion. At the end of the assessment, the simulated patients, the supervisors, the nurses, and the residents also completed evaluation forms with questions regarding the feasibility of the 360-degree assessment.

Participants and staff

Students in semester 10 (before the written part of the national licensing examination) and in their final year (semester 11 and 12) from three universities with different undergraduate curricula were invited to participate in this study. On completion of the study, they were rewarded for participation with a book voucher of 25 Euro. Originally, 90 slots were available but only 75 students signed up for this study. Due to short term illnesses, 70 students completed the assessment (54.3% were female). Their mean age was 25.9 ± 2.2 years. In total, 34 students were in their final year and 36 students had not

started their final year yet. Furthermore, seven supervisors were involved in the assessment (six male, one female), five simulated patients (two male, three female), three nurses (one male, two female), five residents (two male, three female), and fourteen research assistants (four male, ten female).

Analysis of the evaluation data

Evaluation data included aspects of organization of the assessment as well as content of the assessment. Evaluation items were assessed on a 5-point Likert scale (1: does not apply, 2. somewhat applies, 3: partly applies, 4: rather applies, 5: fully applies). Comparisons between final

year students and students before their final year were calculated with a one-sided t-test. Significance levels were set to $p < 0.05$.

Results

All 70 participants completed the assessment in timely manner and no technical or other problems that could not be solved occurred during the process. Three hundred fifty history-taking videos (10 minutes each) and 70 handover videos (30 minutes each) were taken as well as 15 debriefing audio recordings (30 minutes each) which will be analyzed. All questionnaires and assessment sheets were filled out in

Item	Participants (N=70)	Simulated Patients (N=5)	Supervisors (N=7)	Nurses (N=3)	Residents (N=4)	Research Assistants (N=14)
The spatial circumstances (rooms etc.) were adequate.	4.29 ± 1.08	4.60 ± 0.89	4.57 ± 1.13	4.33 ± 0.57	5.00 ± 0.00	4.85 ± 0.37
The technical equipment (computer etc.) was adequate.	4.33 ± 1.05	n/a	5.00 ± 0.00	5.00 ± 0.00	n/a	4.62 ± 1.39
The duration of the assessment was too long.	1.74 ± 0.94	1.40 ± 0.54	3.29 ± 1.25	2.33 ± 2.31	2.75 ± 0.95	2.46 ± 1.39
The changing of rooms was well organized.	4.33 ± 1.09	5.00 ± 0.00	4.57 ± 0.53	5.00 ± 0.00	n/a	n/a
The training equipped me adequately for the task I had to perform during the assessment.	n/a	5.00 ± 0.00	4.00 ± 0.81	4.67 ± 0.57	4.00 ± 1.41	4.25 ± 1.05
My tasks during the assessment were clearly defined.	n/a	5.00 ± 0.00	4.86 ± 0.37	4.67 ± 0.57	4.00 ± 0.81	4.62 ± 0.65

Table 2: Evaluation of the organizational aspects of the 360-degree assessment.

Item	Students in final year (N=34)	Students below semester 11 (N=36)	p-values
The patient cases were interesting.	4.65 ± 0.48	4.53 ± 0.65	0.38
The patient cases seemed realistic to me.	4.50 ± 1.02	4.50 ± 0.61	1.00
I knew the solution to the patient cases intently.	3.32 ± 0.72	3.19 ± 0.88	0.51
There was enough time to manage the patients.	2.62 ± 1.18	2.56 ± 1.16	0.82
I felt confident during history taking.	3.94 ± 0.81	4.00 ± 0.89	0.77
I felt confident during the patient management phase.	3.32 ± 0.87	2.81 ± 0.95	0.02
I felt confident during the handover.	3.50 ± 0.96	3.19 ± 1.21	0.24
The interaction with the nurses felt realistic to me.	4.00 ± 0.81	3.89 ± 0.91	0.59
The disturbances occurring during the patient management phase were designed realistically.	4.12 ± 0.84	4.08 ± 0.96	0.87

Table 3: Students' evaluation of the content of the 360-degree assessment.

paper and pencil and collected for evaluation.

With respect to the organizational aspects of the assessment, all participants and contributors were quite satisfied with the spatial circumstances, the technical equipment, and the organization of the room changes (Table 2). The duration of the assessment appeared adequate. Simulated patients, supervisors, nurses, residents, and

research assistants felt well prepared by their respective trainings for the specific tasks they had to perform during the assessment. The participating students seemed to be very satisfied with the content of the assessment (Table 3). Students in their final year as well as students below semester 11 regarded the patient cases to be interesting and realistic (4.50 ± 1.02 and 4.50 ± 0.60, respectively).

Nurses also considered the patient cases to be very realistic (4.67 ± 0.57) while to supervisors (4.14 ± 0.90) and residents (3.00 ± 0.81) they seemed a bit less realistic. Participants felt there was not enough time to manage the patient cases. While there was no difference between the two groups of students (final year students versus students below semester 11 and before the written part of the national licensing examination) with respect to feeling confident during history taking and handover, final year students felt significantly more confident ($p=0.02$) during the patient management phase than students below semester 11. Both groups of participating students stated that the interprofessional interaction with the nurses was realistic and that the disturbances occurring during the patient management phase were designed realistically. Students felt in general that it was worth the time and effort to have participated in this voluntary assessment (4.37 ± 0.93).

Discussion

Our newly developed competence-based 360-degree assessment worked well with 70 participating students and felt realistic according to the evaluation by participants, supervisors, nurses, and residents. This assessment organized like a triple jump [27] with history taking, patient management phase, and handover, is very closely related to the daily routine of a first year resident. Even though it lacks the assessment of clinical skills like physical examination, which can be examined by mini-CEX [7] and DOPS [8], it provides the possibility to assess competences like handling difficult patient encounters, clinical reasoning, and patient management. Furthermore, medial attitudes and entrustable professional activities can be rated with this assessment.

Interestingly, the only significant difference in the evaluation of this assessment between students in their final year of undergraduate training and younger students, who have not passed the written part of the national licensing examination yet, was that younger students felt more uncertainty during the patient management phase with interprofessional disturbances than final year students. This finding is underscored by the results of a ranking study of 25 competences relevant for medical graduates performed with undergraduate medical students, where final year students ranked the competence "Coping with uncertainty" to be significantly more relevant for a first year resident than first year students did [28]. Even though undergraduate medical students are nowadays introduced earlier to clinical tasks - even at medical schools with traditional curricula comprising pre-clinical and clinical courses - being assessed with multiple choice exams during undergraduate training might foster the believe of certainty in medical decisions in students [29]. Facing patient management, even when the patients are only simulated, leads to a large number of unnecessary laboratory and radiology tests ordered by German medical students [15] in the attempt to prevent mistakes and to reduce the uncomfortable feeling of uncertainty. Inducing the feeling of uncertainty in less advanced students during the patient management phase of our assessment might be another indicator how realistic the perception of the assessment was. It could also point to the fact that undergraduate medical students rarely experience patient workups before the final year in a setting with interprofessional interaction and other disturbances that occur in the daily routine of hospital based physicians. This phase with at least perceived independent accountability in patient management and decision-making may be especially challenging for the participating students.

All assessors of this competence-based 360-degree assessment, the simulated patients, supervisors, nurses, and residents felt well prepared by the assessor training and clearly knew their tasks. Even though it has been shown for mini-CEX that rater training had no effect on the reliability of the exam scores [30] it is an unconditional requirement to reach solid reliability for the assessment of competences [11]. Another study also showed that rater training had a positive effect on the quality of inferences made by raters in competence modeling [31]. Additionally, a qualitative study pointed out that participation of clinicians in a performance dimension rater training and in a frame of reference rater training equipped participants with assessment skills, which were congruent with principles of criterion-referenced assessment and entrustment, and basic principles of competency-based education [32].

The patient cases of our assessment, which were based on the top 10 competences of our ranking study [17] were considered to be very realistic by our participants and assessors. If a medical school wishes to establish such a competence-based 360-degree assessment, it will be necessary to establish a blueprint for the generation of more patient cases similar to a blueprint that has been developed for a competency-based examination of clinical skills with standardized patients [33]. Furthermore, medical students should be given the opportunity to assess their clinical competences longitudinally during their undergraduate studies to prepare for such an assessment [34]. Additionally, interprofessional learning should be integrated further in undergraduate medical education [35] and assessment [36] to accustom medical students to the interprofessional exchange, which they will experience in their daily work as residents and which they encountered as a disturbance in their patient workup in our assessment.

A weakness of our study is that we did not reach the planned number of 90 participating students since the assessment took place during the holiday time and participation was voluntary. Therefore, we only performed the assessment with five groups of students per day instead of six. However, the whole setup was planned for six groups and worked well leading us to the conclusion that it is feasible to implement the assessment in the suggested way.

Conclusion

The implementation of a competence-based 360-degree assessment is feasible for large groups of students. If such an assessment is used in a mandatory fashion, students should be given the opportunity to experience the real situation of residents' work including responsibility, interprofessional interactions, and uncertainty. Further analysis of the data gathered in our assessment will provide evidence whether students from different undergraduate curricula perform differently with respect to patient management in complex medical situations resembling the working reality of a first year resident.

Acknowledgements

We thank Sarah Prediger, Sophie Fürstenberg, Kristina Schick, Jana Deppermann, Dr. Fabian Fincke, Dr. Martin Gartmeier, and Daniela Vogel for the excellent organization and support throughout the duration of this project. Many thanks to Christian Bruhn, Ulrike Johannson, Thomas Klees, Elena Meißner, and Claudia Wiedemer for being the 'best simulated patients we ever had during our whole undergraduate medical training' (participating student's comment). We thank all participating students, nurses, residents, and supervisors

as well as all assistants who provided technical support. This study is part of the ÄKHOM-project, funded by the German Ministry of Education and Research (BMBF), reference number: 01PK1501A/B/C. The study was performed in accordance with the Declaration of Helsinki. The Ethics Committee of the Chamber of Physicians, Hamburg, confirmed the innocuousness of this study with consented, anonymized, and voluntary participation (PV3649).

References

1. Fraser SW, Greenhalgh T (2001) Coping with complexity: Educating for capability. *BMJ* 323: 799-803.
2. Frank JR, Mungroo R, Ahmad Y, Wang M, De Rossi S, et al. (2010) Toward a definition of competency-based education in medicine: A systematic review of published definitions. *Med Teach* 32: 631-637.
3. Frank JR, Danoff D (2007) The CanMEDS initiative: Implementing an outcome-based framework of physician competencies. *Med Teach* 29: 642-647.
4. Fischer MR, Bauer D, Mohn K (2015) Finally finished! national competence based catalogues of learning objectives for undergraduate medical education (NKLM) and dental education (NKLZ) ready for trial. *GMS Z Med Ausbild* 32: Doc35.
5. Ten Cate O, Snell L, Carraccio C (2010) Medical competence: The interplay between individual ability and the health care environment. *Med Teach* 32: 669-675.
6. Melvin L, Cavalcanti RB (2016) The oral case presentation: A key tool for assessment and teaching in competency-based medical education. *JAMA* 316: 2187-2188.
7. Norcini JJ, Blank LL, Arnold GK, Kimball HR (1995) The mini-CEX (clinical evaluation exercise): A preliminary investigation. *Ann Intern Med* 123: 795-799.
8. Barton JR, Corbett S, van der Vleuten CP (2012) The validity and reliability of a direct observation of procedural skills assessment tool: Assessing colonoscopic skills of senior endoscopists. *Gastrointest Endosc* 75: 591-597.
9. Govaerts MJ, van der Vleuten CP, Schuwirth LW, Muijtjens AM (2007) Broadening perspectives on clinical performance assessment: Rethinking the nature of in-training assessment. *Adv Health Sci Educ Theory Pract* 12: 239-260.
10. Crossley J, Johnson G, Booth J, Wade W (2011) Good questions, good answers: Construct alignment improves the performance of workplace-based assessment scales. *Med Educ* 45: 560-569.
11. Wijnen-Meijer M, Van der Schaaf M, Booij E, Harendza S, Boscardin C, et al. (2013) An argument-based approach to the validation of UHTRUST: Can we measure how recent graduates can be trusted with unfamiliar tasks? *Adv Health Sci Educ Theory Pract* 18: 1009-1027.
12. Wijnen-Meijer M, ten Cate O, van der Schaaf M, Harendza S (2013) Graduates from vertically integrated curricula. *Clin Teach* 10: 155-159.
13. Wijnen-Meijer M, van der Schaaf M, Nillesen K, Harendza S, Ten Cate O (2013) Essential facets of competence that enable trust in medical graduates: A ranking study among physician educators in two countries. *Perspect Med Educ* 2: 290-297.
14. Wijnen-Meijer M, Ten Cate O, van der Schaaf M, Burgers C, Borleffs J, et al. (2015) Vertically integrated medical education and the readiness for practice of graduates. *BMC Med Educ* 15: 229.
15. Harendza S, Alofs L, Huiskes J, Wijnen-Meijer M (2013) Ordering patterns for laboratory and radiology tests by students from different undergraduate medical curricula. *BMC Med Educ* 13: 109.
16. Wijnen-Meijer M, van der Schaaf M, Nillesen K, Harendza S, Ten Cate O (2013) Essential facets of competence that enable trust in graduates: A Delphi study among physician educators in the Netherlands. *J Grad Med Educ* 5: 46-53.
17. Fürstenberg S, Schick K, Deppermann J, Prediger S, Berberat PO (2017) Competencies for first year residents-physicians' views from medical schools with different undergraduate curricula. *BMC Med Educ*.
18. Costa PT, McCrae RR (2008) The revised Neo personality Inventory (NEO-PI-R). *The SAGE Handbook of Personality Theory Assessment* 2: 179-198.
19. Davis MH (1980) A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology* 10: 85.
20. Klemenc-Ketis Z, Vrecko H (2014) Development and validation of a professionalism assessment scale for medical students. *Int J Med Educ* 5: 205-211.
21. Müller B, Basler HD (1993) Kurzfragebogen zur aktuellen Beanspruchung (KAB). Beltz, Weinheim, Germany.
22. Mercer SW, Maxwell M, Heaney D, Watt GC (2004) The consultation and relational empathy (CARE) measure: Development and preliminary validation and reliability of an empathy-based consultation process measure. *Fam Pract* 21: 699-705.
23. Peterson EB, Calhoun AW, Rider EA (2014) The reliability of a modified Kalamazoo Consensus Statement Checklist for assessing the communication skills of multidisciplinary clinicians in the simulated environment. *Patient Educ Couns* 96: 411-418.
24. Durning SJ, Artino A, Boulet J, La Rochelle J, Vleuten CV, et al. (2012) The feasibility, reliability, and validity of a post-encounter form for evaluating clinical reasoning. *Med Teach* 34: 30-37.
25. Ten Cate O, Chen HC, Hoff RG, Peters H, Bok H, et al. (2015) Curriculum development for the workplace using entrustable professional activities (EPAs): AMEE Guide No. 99. *Med Teach* 37: 983-1002.
26. Paulhus DL (1984) Two-component models of socially desirable responding. *J Pers Soc Psychol* 46: 598-609.
27. Smith RM (1993) The triple-jump examination as an assessment tool in the problem-based medical curriculum at the University of Hawai. *Acad Med* 68: 366-372.
28. Fürstenberg S, Harendza S (2017) Differences between medical student and faculty perceptions of the competencies needed for the first year of residency. *BMC Med Educ*.
29. Freiwald T, Salimi M, Khaljani E, Harendza S (2014) Pattern recognition as a concept for multiple-choice questions in a national licensing exam. *BMC Med Educ* 14: 232.
30. Cook DA, Dupras DM, Beckman TJ, Thomas KG, Pankratz VS (2009) Effect of rater training on reliability and accuracy of mini-CEX scores: A randomized, controlled trial. *J Gen Intern Med* 24: 74-79.
31. Lievens F, Sanchez JI (2007) Can training improve the quality of inferences made by raters in competency modeling? A quasi-experiment. *J Appl Psychol* 92: 812-819.
32. Kogan JR, Conforti LN, Bernabeo E, Iobst W, Holmboe E (2015) How faculty members experience workplace-based assessment rater training: A qualitative study. *Med Educ* 49: 692-708.
33. Mookherjee S, Chang A, Boscardin CK, Hauer KE (2013) How to develop a competency-based examination blueprint for longitudinal standardized patient clinical skills assessments. *Med Teach* 35: 883-890.
34. Bodamer C, Feldman M, Kushinka J, Brock E, Dow A et al. (2015) An internal medicine simulated practical examination for assessment of clinical competency in third-year medical students. *Simul Healthc* 10: 345-351.
35. Solomon P, Risdon C (2011) Promoting interprofessional learning with medical students in home care settings. *Med Teach* 33: e236-241.
36. Domac S, Anderson L, O'Reilly M, Smith R (2015) Assessing interprofessional competence using a prospective reflective portfolio. *J Interprof Care* 29: 179-187.