The use of programmatic assessment in the clinical workplace: A Maastricht case report

ERIK W. DRIESSEN[1], JAN VAN TARTWIJK[2], MARJAN GOVAERTS[1], PIM TEUNISSEN[1] & CEES P. M. VAN DER VLEUTEN[1]

[1]Maastricht University, The Netherlands, [2]Utrecht University, The Netherlands

Abstract

The differences of learning experiences in the workplace put challenges on the assessment: the assessment programme should be aligned with the general competency framework of the curriculum and also fit to the differences in learning contexts of the workplace. We used van der Vleuten's programmatic assessment model to develop a workplace-based assessment programme for final year clerkships. We aimed to design a programme that stimulates learning, supports assessment decision, is feasible and non-bureaucratic. The first experiences with the programme show that students think that the programme has high learning value and the assessment is sufficiently robust. Many of the commonly reported weaknesses of work-based assessment (not a good fit with the educational context, too complex, too bureaucratic and too much work) were not mentioned by the students.

Introduction

From an educational viewpoint, learning in the clinical workplace is far less structured than during the preclinical phase of the medical curriculum, therefore specific challenges need to be overcome in designing an effective programme for workplace-based assessment. Even though the goals and competencies to be achieved at the end of clinical training can be determined in advance, the fact that learning opportunities depend on patient mix and presence of clinical supervisors makes them difficult to plan and causes them to vary considerably between students (Billet 2006). The main challenge for those designing a workplace-based assessment programme is to structure assessments in such a way that, despite differences between learning environments, all students are stimulated to work towards the required competency levels and have sufficient opportunities to demonstrate that they have attained them. Assessment programmes should be aligned with the general competency framework of the curriculum and also fit to the different learning contexts of the workplaces.

In the literature, there are frequent reports of dissatisfaction with workplace-based assessment among students, residents and doctors, who criticise it for not fitting into the clinical context and for being too complex, too bureaucratic and too much work (Overeem et al. 2007; Hrisos et al. 2008; Davis et al. 2009; Pereira & Dean 2009; Sabey & Harris 2011). As GP trainees commented in a recent study: 'the assessment thus becomes a set of hoops to jump through, rather than a robust system for feedback and learning' (Sabey & Harris 2011). In response to such criticisms, several authors have advocated for a different design of work-based assessment. In their AMEE guide to work-based assessment, Norcini and Burch (2007)

Practice points

- A programmatic assessment approach during the clerkships helps to provide students with feedback.
- A simple structure of the assessment programme and its separate parts results in a programme that is feasible for students and teachers.
- A 'lean' assessment programme contributes to the feasibility.
- A learner led assessment approach in which the learner can choose areas for observation and feedback contributes to the learning value of the programme.
- A final assessment that is based on the principles of qualitative research is perceived by students as honest and robust for taking high-stake decisions.

conclude that 'long term use (of work-based assessment) may require further modification and simplification of existing methods so as to make them more user-friendly in busy clinical settings where patient care is the first priority and trainee assessment of less importance'. The General Medical Council in the UK advocates 'learner-led' assessment programmes, with trainees creating learning opportunities and determining areas for observation and feedback (GMC 2010). Further, several authors have recommended an 'approach that relies on qualitative information and thus on professional judgment to overcome shortcomings of the current assessment program' (van der Vleuten et al. 2010; Sabey & Harris 2011).

Based on the literature and van der Vleuten's programmatic assessment model described in this issue, we developed a workplace-based assessment programme for the final

Correspondence: E. Driessen, Department of Educational Development and Research, Faculty of Health, Medicine and Life Sciences, P.O. Box 616, 6200 MD Maastricht, The Netherlands. Tel: +31 43 3885774; email: e.driessen@maastrichtuniversity.nl

	Below expectation	As expected	Above expectation
Clinical performance (for instance as judged by mini-CEX)	Slow in taking a history and performing a physical examination. Considers irrelevant aspects	Adequate speed in taking a history and performing a physical examination. Relevant aspects are considered	Conducts an adequate and efficient history and physical examination
	Slow in making a diagnosis. Misses important conclusions	Adequate speed in making a diagnosis. Diagnosis contains important conclusions	Arrives at an accurate diagnosis within adequate time
	Frequently unable to formulate management plan and needs considerable guidance	Formulates an adequate management plan for simple clinical presentations	Formulates an adequate management plan for simple clinical presentations
		Needs some guidance	Needs little guidance
		Achieves these goals in the second half of the internship	Has achieved these goals at the start of the internship
Professionalism (for instance as judged by 360-degree feedback)	Does not keep appointments	Keeps appointments	Keeps appointments
	Occasionally fails to ask for supervision when this is necessary. Reacts defensively to feedback	Asks for supervision when this is necessary	Asks for supervision when this is necessary
	Is unable to cope with stress	Needs help in reflecting and considering alternatives and responds adequately to feedback	Is able to reflect critically; responds adequately to feedback and is prepared to acknowledge errors
	Does not pay attention to his/her personal appearance	Occasionally needs help in coping with stress	Is able to cope with stress adequately
	Frequently shows inappropriate behaviour or behaves disrespectfully	Appropriate personal appearance; behaves respectfully	Looks well cared for and behaves respectfully

Table 1. Examples of rubrics, defining the competency level to be attained at the end of the final-year clerkship and detailing when performance is below or above this level (Source: Maastricht University).

clerkship of the Maastricht undergraduate curriculum. We aimed to design a programme that stimulates learning, supports trustworthy assessment and steers clear of the reported shortcomings of many of the current work-based assessment programmes (van der Vleuten et al. 2012). We conducted a case study using questionnaires and focus groups to investigate students' experiences with the new programme.

The Maastricht assessment programme for the final, 18-weeks, clinical clerkship

Based on a programmatic assessment model and on the available evidence about workplace-based assessment, an assessment programme was developed for the final year clerkship at Maastricht University, the Netherlands. The clerkship, which was first implemented in 2006, is scheduled in year 6, and differs from the clerkships in years 4 and 5 in the following respects:

- The clerkship lasts 18 weeks at one department, considerably longer than the other, 4–8 weeks, clerkships.
- Students apply for a clerkship in a discipline and location of their preference.
- For the duration of the clerkship, students have a personal mentor from the hospital department, who is responsible for supervision and assessment.
- Mentor and student share patient care responsibilities and students have more responsibility for patient care than during earlier clerkships.

In describing the assessment programme, we consecutively discuss the building blocks of the programmatic assessment

model: learning activities, assessment activities, supporting activities, intermediate evaluation and final assessment.

Learning activities

The structure of the programme is based on the CanMEDS competency framework. For each competency, a rubric is designed, defining the competency level to be attained at the end of the clerkship and detailing when performance is below or above this level (Table 1).

In order to attune learning activities to different clerkship settings as well as to students' individual learning needs, students prepare a learning plan in consultation with their mentor before the clerkship starts. It is important that the learning plan is tailored to students' competencies attained during previous clerkships and the clerkship's specific learning opportunities.

Assessment activities

The *assessment activities* are designed to ensure that students receive feedback in different competency areas, with emphasis on qualitative feedback. All assessment activities contribute to the final assessment.

Since assessment activities are designed to fit all disciplines and locations that can be selected by students for the final clerkship, the assessment instruments are generic and suited to different disciplines (i.e. family medicine, surgery or psychiatry) and locations (i.e. Maastricht, Ghana or Italy). In designing the programme, care has been taken that assessment activities fit within the possibilities and limitations of the clerkship, are easy to use and time efficient, and that teachers and students experience minimal problems in conducting

them in the busy clinical workplace. We strived to make the assessment activities in line with recommendations from the literature to make work-based assessment efficient, simple and 'learner led' (Norcini & Burch 2007; GMC 2010). We chose for short and simple assessment instruments, with limited numeric items and enough space for narrative feedback. We made use of rubrics to provide the users with information on what is expected of the students.

The programme sets minimum requirements for five *single assessment activities*:

- five mini-Clinical Evaluation Exercises (mini-CEX);
- two multisource feedback (MSF) procedures;
- two Critical appraisals of a topic (CAT);
- two progress tests; and
- one OSCE.

In an effort to ensure broad coverage of the layers of Miller's pyramid, we included two progress tests ('knows' and 'knows how' level) (Schuwirth et al. 2010) and one standardised assessments of competence (OSCEs) ('shows how' level). At the start of the clerkship, the portfolio contains only the student's learning plan, to which, over the course of the clerkship, various items are added: assessments, *intermediate assessment*, reports of progress meetings, students' reflections on their learning process and any other materials students may wish to add. The assessment activities listed above are the minimum requirements, but students and mentors are free to add to them. It is the task of the student to collect workplace assessments from different assessors in different situations in order to achieve a sufficiently varied sample across assessors and cases. Since the rubrics and the single assessments are not specifically designed for a particular discipline or location, the same instruments can be used for all clerkship settings, from psychiatry to surgery.

Supporting activities

The *supporting activities* are designed to help students direct their learning. In using feedback, received on various occasions, to reflect on their performance and progress, students are supported by their mentor and by the portfolio. The mentor is one of the clinical teachers of the placements' hospital department. Student and mentor are expected to schedule a total of three *progress meetings*, in weeks 4, 8 and 12, to monitor the student's progress. The discussion is guided by the information in the portfolio and by a brief self-analysis of the student's development in relation to the learning plan and the CanMEDS competencies which the student prepares for the meeting. The mentor gives feedback and guides the student's self-reflection. Based on the outcomes of the meeting, mentor and student can adjust the learning goals and plan activities for the upcoming period.

Intermediate evaluations

Since it was deemed not feasible for students to collect a sufficient number of assessments to justify several intermediate evaluations, the assessment programme deviates from the model for programmatic assessment in that it contains only

one *intermediate evaluation*, scheduled 4 weeks into the clerkship. Since intermediate evaluations of a cohort of 340 students were considered too much of a burden for one committee of examiners, the mentors conduct the intermediate evaluation, and decide, based on the information in the portfolio and on their judgement of the student, whether a student is allowed to continue with the clerkship.

Final evaluations

An assessment committee of the medical faculty conducts the *final assessment*. Since this is one of the most important assessments of the final year, several measures have been taken to guarantee the quality of this – high-stakes – assessment (Driessen et al. 2005a; van der Vleuten et al. 2010). In accordance with the recommendation to focus on qualitative information (Sabey & Harris 2011), measures taken to ensure trustworthiness are largely derived from qualitative research. As the first safeguard, the assessment committee uses an assessment form containing rubrics defining the required performance standards for each competency. As the second safeguard, students and mentors submit their recommendations with regard to the assessment to the committee at the end of the clerkship. To this end, special student and mentor forms, similar to the assessment form used by the committee, have been developed, which students and mentors use to indicate their judgements of the quality of the portfolio and the student's performance. In this way, the assessment committee can base its assessment of a student on the portfolio, supplemented by information from both mentor and student. Our earlier experiences with a similar procedure for the first-year portfolio in Maastricht were quite encouraging (Driessen et al. 2005a), with both students and mentors stating that the procedure had no negative impact on mentoring and student self-reflection (Driessen et al. 2003, 2005b). Other safeguards are the built-in feedback cycles of intermediate assessment and progress meetings, an appeal procedure for students, training and benchmarking of examiners and an assessment committee of limited size (six members) (van Tartwijk & Driessen 2009).

Evaluation

Between 2007 and 2009, 670 students participated in the assessment programme. All these students filled in a web-based questionnaire to assess students' perceptions regarding:

- possibilities to engage in learning activities;
- feedback from the assessment instruments;
- the robustness of the final assessment; and
- the quality of the supporting activities.

Questionnaire items were rated on a five-point Likert scale (1 = totally disagree, 5 = totally agree). Table 2 shows the results for each year of the programme.

The questionnaire paints a fairly positive picture of students' perceptions of the educational value and the robustness of the assessment programme during the 3 years of the study. In addition to the questionnaire, we used focus groups to gain more insight into students' perceptions of the educational value and the feasibility of the assessment

Table 2. Students' ratings of statements about the workplace-based assessment programme for final year clerkships, Maastricht University.

	2007[a] (N = 160)	2008 (N = 235)	2009 (N = 275)
Learning activities			
During the clerkship, I could pay sufficient attention to all the CanMEDS roles	4.2/0.7[b]	4.2/0.8	4.2/07
Assessment activities			
I received informative feedback from the mini-CEX	4.2/0.8	4.1/0.7	4.2/0.7
I received informative feedback from the MSF	4.1/0.9	4.2/0.7	4.2/0.8
I received informative feedback from the CAT assessment	3.9/0.8	3.9/0.8	4.0/0.8
Direct observation of my performance during patient encounters took place regularly (once a week)	3.5/1.2	3.7/1.0	3.8/1.0
I received adequate feedback on observed patient encounters	4.1/0.8	4.1/0.8	4.1/0.7
Supporting activities			
There was enough time spent on guidance	4.2/0.8	4.2/0.8	4.1/0.9
I was satisfied with the quality of the guidance	4.3/0.8	4.2/0.9	4.2/0.8
During the progress interviews, I received sufficient feedback on my performance and progress	4.3/0.8	4.2/0.8	4.1/0.8
Final assessment			
The mentor provided sufficient arguments for his/her recommendation for the final assessment	4.4/0.7	4.4/0.6	4.3/0.7
The final assessment was conducted with care	4.4/0.7	4.4/0.7	4.4/0.7
The final assessment accurately reflects my performance during the clerkship	4.3/0.8	4.4/0.7	4.3/0.7
All CanMEDS roles were considered for the final assessment	4.0/0.9	4.2/0.8	4.1/0.8

Notes: [a]Year and number of students and [b]mean/standard deviation.

programme and of the robustness of the final assessment. In order to tap into a variety of perspectives, we conducted two focus groups with 19 students, sampled purposively to represent different disciplines, locations (academic and non-academic hospitals), attitudes to portfolio learning recorded earlier and gender. An experienced moderator guided the discussion, while a research assistant took notes and asked questions to clarify points when necessary. We report the results of the focus groups for each of the building blocks of the assessment programme.

Learning activities

During the first weeks of the clerkship, the learning goals received only little attention during clinical work:

> Before the clerkship you think about what learning goals are important, but when you are busy working you don't think of them anymore, it was only in the last weeks of the clerkship did I start to do more interviews, for example, because I knew this was in my learning plan (p. 14).

The clerkship was considered very informative and useful, particularly as a result of the combination of more independence, longer duration, and support and guidance from the assessment programme.

> You have more opportunities for the learning process to return to it, to direct it, you have time to work on certain things...I really think it is another principle, especially because of the assessment (p. 2).

Assessment activities

Single assessments were thought to yield good feedback, but students had to actively seek feedback by specifically asking for explanations of the assessment and for narrative and critical feedback.

> You have to look at the form and notice why did you write that and ask for explanations, because they are not going to go through the list with you (p. 5).

Students were highly appreciative of the fact that the single assessments reflected their improvement in domains where they were having difficulties.

> that is very difficult. At first you go scarlet and you begin to stammer,...it is just going really very badly. And then you get real feedback, and the MiniCEX form forces them to do so. For me it really works (p. 9).

An advantage of working in one department for a longer period was that the assessment activities enabled students to follow their own development over time and improve in areas where it was needed. Students sometimes received only positive evaluations, for example in patient contact assessments in MSF, but they did not consider this very useful.

Supporting activities

Although students were satisfied with the supporting activities, they indicated that they sometimes had to take action to actually obtain support, because some mentors were initially not familiar with the assessment instruments and the portfolio, and did not know what they were supposed to do. However, once the mentors knew the ropes, the progress meetings had sufficient educational value. The meetings were primarily guided by the portfolio. During the first meeting, students came to realise that the mentor acted primarily as a coach, and after that they felt sufficiently secure to be critical of their own functioning and to write this down in their self-reflection. As a result, a growing feeling of trust developed between mentor and student.

> My mentor asked me in the first progress meeting, why didn't you tell me this, then we could have

solved it. Afterwards I told him, because by then there was trust which hadn't been there at first (p. 5).

Discussing the portfolio was considered especially useful when there was a discrepancy between the student's self-reflection and the results of individual assessments.

> often there is a discrepancy between my opinion and someone else's, that is also very revealing. Why did I think my performance was satisfactory…I was criticised for this: why did you judge your performance as satisfactory? Why satisfactory and not good? That is something you can discuss (p. 13).

Even though the clerkship was very time intensive – 55 h a week on average – students did not think the workload for the assessment programme was excessive. As one student put it: *it (the assessment programme) was brief and to the point (p. 17).* It was partly up to the student to determine how elaborate the portfolio was going to be:

> you can make an enormous difference: how often do you ask for assessment? What more do you add? It is how you manage it (p. 6).

It was important to the students that the programme was relevant. Because the portfolio was central to learning and assessment, students had no problem spending time on it.

> The first and second year of the programme, well then I am actually only studying, there's not so much to write about… But the more you get involved with patient care, then my learning goals really changed, because then you learn how to work with doctors, how do you treat patients? What if you are afraid to do something? (p. 13)
>
> I felt that I was writing the same things over and over again every year, until this year…this was the first time I realised hmm, suddenly I'm writing completely different things. This is because it's my final year clerkship. And I think to me this makes a huge difference, for the first time it is useful (p. 1).

Final assessment

Even though the portfolio offered a good representation of students' work and performance, some students thought the assessment was too indirect: they would prefer to be assessed by their mentors, not by some faculty committee. They felt that their mentors had been able to observe their work and were therefore better suited to judge their performance.

Discussion

The main reasons for using van der Vleuten's programmatic assessment model for the final clerkship were to enhance the clerkship's learning value while providing robust assessment and to prevent some of the frequently reported problems with work-based assessment. The first experiences with the programme show that students think the programme has high learning value and the assessment is sufficiently robust. Many of the commonly reported weaknesses of work-based

Table 3.	Success factors for programmatic assessment in the clinical workplace.

- Give single assessments maximum learning value
- Aggregate the single assessments in a portfolio
- Use a holistic assessment procedure that relies on qualitative information
- Organise mentoring on the work floor
- Make the assessment programme 'learner led'
- Strive for a simple structure for the whole programme and its separate parts
- Strive for a lean assessment programme
- Take care of the implementation of the programme

assessment (not a good fit with the educational context, too complex, too bureaucratic and too much work) were not mentioned by the students.

What can we learn from our experiences with the assessment programme? What were in our eyes the key factors for the success? First of all the programmatic approach of the assessment appeared to be effective in providing students with feedback. With the help of the mentors they could use the feedback to direct their learning. The single assessments were mainly perceived as feedback instruments and not as selection instruments. The aggregation of the single assessments in a portfolio helped the students, mentors and examiners to monitor the performance of the student during the clerkship. A second lesson is that the simple structure of the whole programme and its separate parts (assessment activities, supportive activities) results in a programme that is feasible for students and teachers. Third, that a 'lean' assessment programme contributed to this feasibility. Instead of many other assessment programmes, we did not aspire to cover all the separate competences and possible clinical situations by separate assessments instruments. We only prescribed a limited set of required assessments. Fourth, we gave students and mentors – in line with the GMC guidelines – the freedom in choosing areas for observation and feedback (GMC 2010). We think that this 'learner led' approach contributed to the perceived learning value of the assessment and the congruence between the assessment programme and the different learning environments of the students. Overall, we think that these last three factors (simple, lean and student led) made that the students did not experience the assessment as bureaucratic.

The fifth key factor is the holistic assessment procedure. We had already experience with this way of assessing students in our first-year portfolio (Driessen et al. 2005a). In the eyes of the students, this procedure that is based on the principles of qualitative research is also robust for taking high-stake decisions in work-based assessment. Finally, from literature, we know that in the end, implementation is the decisive factor for the success of any assessment in work-based assessment (Norcini & Burch 2007). Especially teachers and students have to be bought in and familiarised with the assessment instruments. Our students were already quite well known with most assessment methods that we used, the only new instrument for them was the MSF. The students played an important role in familiarising the clinical teachers with the different instruments. The students were also the ones who 'teached' the teachers to

use the instruments in such a way that they provide useful feedback for them. Probably, the students were so active and assertive, because they directly gained from the feedback they received during the placement.

In Table 3, we summarise what in our view were the success factors for a programmatic approach for workplace-based assessment for final-year clerkships.

Our experiences show that the model for programmatic assessment can be used for facilitating learning and assessment in the clinical workplace. We hope this case description will be helpful for those who further want to improve the workplace-based assessment.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the article.

Notes on contributors

ERIK DRIESSEN, is an associate professor, Department of Educational Development and Research, Faculty of Health, Medicine and Life Sciences, Maastricht University, The Netherlands.

JAN VAN TARTWIJK is a professor of Education, Faculty of Social and Behavioural Sciences, Utrecht University, The Netherlands.

MARJAN GOVAERTS is an assistant professor, Department of Educational Development and Research, Faculty of Health, Medicine and Life Sciences, Maastricht University, The Netherlands.

PIM TEUNISSEN is an assistant professor, Department of Educational Development and Research, Faculty of Health, Medicine and Life Sciences, Maastricht University, The Netherlands.

CEES VAN DER VLEUTEN is a professor of Education, Chair of the Department of Educational Development and Research and Scientific Director of the School of Health Professions Education, Faculty of Health, Medicine and Life Sciences, Maastricht University, The Netherlands, Honorary Professor at King Saud University (Riyadh, Saudi Arabia), Copenhagen University (Copenhagen, Denmark) and Radboud University (Nijmegen, The Netherlands).

References

Billet S. 2006. Constituting the workplace curriuclum. J Curric Stud 38:31–48.

Davis MH, Ponnamperuma GG, Ker JS. 2009. Student perceptions of a portfolio assessment process. Med Educ 43:89–98.

Driessen E, van der Vleuten C, Schuwirth L, van Tartwijk J, Vermunt J. 2005a. The use of qualitative research criteria for portfolio assessment as an alternative to reliability evaluation: A case study. Med Educ 39:214–220.

Driessen E, van Tartwijk J, Vermunt JD, van der Vleuten CP. 2003. Use of portfolios in early undergraduate medical training. Med Teach 25:18–23.

Driessen EW, van Tartwijk J, Overeem K, Vermunt JD, van der Vleuten CP. 2005b. Conditions for successful reflective use of portfolios in undergraduate medical education. Med Educ 39:1230–1235.

GMC. 2010. Workplace based assessment: A guide for implementation. London: GMC.

Hrisos S, Illing JC, Burford BC. 2008. Portfolio learning for foundation doctors: Early feedback on its use in the clinical workplace. Med Educ 42:214–223.

Norcini J, Burch V. 2007. Workplace-based assessment as an educational tool: AMEE Guide No. 31. Med Teach 29:855–871.

Overeem K, Faber MJ, Arah OA, Elwyn G, Lombarts KM, Wollersheim HC, Grol RP. 2007. Doctor performance assessment in daily practise: Does it help doctors or not? A systematic review. Med Educ 41:1039–1049.

Pereira EA, Dean BJ. 2009. British surgeons' experiences of mandatory online workplace-based assessment. J R Soc Med 102:287–293, doi: 10.1258/jrsm.2009.080398.

Sabey A, Harris M. 2011. Training in hospitals: What do GP specialist trainees think of workplacebased assessments? Educ Prim Care 22:90–99.

Schuwirth L, Bosman G, Henning RH, Rinkel R, Wenink AC. 2010. Collaboration on progress testing in medical schools in the Netherlands. Med Teach 32:476–479.

van der Vleuten CP, Schuwirth LW, Scheele F, Driessen EW, Hodges B. 2010. The assessment of professional competence: Building blocks for theory development. Best Pract Res Clin Obstet Gynaecol 24:703–719.

van Tartwijk J, Driessen EW. 2009. Portfolios for assessment and learning: AMEE Guide no. 45. Med Teach 31:790–801.

van der Vleuten CPM, Schuwirth LWT, Driessen E, Dijkstra J, Tigelaar D, Baartman LKJ, van Tartwijk J. 2012. A model for programmatic assessment fit for purpose. Med Teach 34:205–214.