Programmatic Assessment: A Paradigm Shift in Medical Education

*

Roger Ellis OBE[†]

Elaine Hogard[‡]

[†]Emeritus Professor, Ulster University; University of Chester UK
[‡]Director of Assessment and Programme Evaluation, Northern Ontario School of Medicine, Canada

Abstract

This article describes Programmatic Assessment, a significant innovation in Medical Education, which involves such a change in thinking and action as to merit the label paradigm shift. Programmatic Assessment is a radical approach to assessment throughout the medical education programme devised to address endemic problems in assessment and its deleterious their deleterious effects on the curriculum and student learning. While the focus is on Medical Education the ideas,; issues and approaches are relevant to professional education more broadly.

Keywords: Programmatic Assessment, Medical Education, Professional education, problems in assessment.

* (Insert Article type via "File|Properties|Custom Properties"). URL: http://ojs.aishe.org/index.php/aishe-j/article/view/[insert ojs 295

This article describes Programmatic Assessment, a significant innovation in Medical Education, which involves such a change in thinking and action as to merit the label paradigm shift. Programmatic Assessment is a radical approach to assessment throughout the medical education programme devised to address endemic problems in assessment and its deleteriousthei r deleterious effects on the curriculum and student learning. While the focus is on Medical Education the ideas,; issues and approaches are relevant to professional education more broadly.

Programmatic Assessment is so described for several reasons. First it represents an attempt to consider assessment as constituting a programme as important as the curriculum itself, and thus meriting being planned and reviewed in a similar way. Second it underlines that while assessment is necessary as a basis for progress and award decisions, it may also be considered as a learning programme for students in itself. Third it stresses that the data derived from assessment charts the progress and development of students throughout their programme. Fourth it reflects the belief that no one assessment point should determine progress or award but that such decisions should be based on an aggregation of points through the programme.

The relationship between the curriculum and assessment is a complex one but there is little doubt that the impact of assessment on desirable learning is often very negative. (Cilliers CF et al 2010,2012). Aiming to pass examinations can lead to poor learning styles in students. It can also encourage a ‘grade culture’, where achieving the highest grades is the main objective. The feedback on performance provided to students by grades is simplistic and lacks any qualitative assessment of content and performance. Such feedback does not align with curricular goals and culminates in a non-meaningful aggregation of assessment information with no clear longitudinal element.

Assessment in medical education tends to be concentrated on a limited range of methods with a pervasive use of Multiple Choice Tests. Given the complex objectives of programmes no single method can suffice. Each individual method has significant limitations and any one method represents a compromise on reliability, validity and educational impact. The apparent objectivity of multiple choice tests and other other such methods can minimise the role of

expert judgement even though such judgements have been demonstrated to be the most valid and reliable form of assessment. (GMC 2009).

The paradox is that assessment is designed to measure the extent to which students have achieved the learning objectives of the programme. But assessment itself can then have a negative effect on the learning which the programme is intended to cultivate. The programme requires assessment but then assessment drives the programme. (Van der Vleuten C et al 2012). Devising and operating an assessment system is a major use of staff time and yet its impact on desirable learning is minimal or even negative.

To address these problems a radical new approach to assessment known as Programmatic Assessment has been developed, first in the Netherlands at the University of Maastricht and subsequently at medical schools throughout the world. Van Vleuten and his colleagues (Van Vleuten et al 2012) proposed a model for what they described as progressive assessment in action, which simultaneously optimised assessment for learning and improved assessment for decision making about student progress. The model was based on a set of assessment principles that were derived from empirical research. The approach specifies cycles of training, assessment and learner support activities that are complemented by intermediate and final moments of evaluation on aggregated assessment data points.

A key principle of the approach is that individual data points are maximised for learning and feedback value whereas high stake decisions are based on an aggregation of many data points. Thus each assessment point is optimised for learning using meaningful feedback but the key decisions about progress on the programme or the award of credit are never taken on single assessment points but only on an aggregation of points. Expert judgement plays an important role in the assessment programme. Fundamental is the notion of sampling and bias reduction to deal with the inevitable subjectivity of this kind of judgement. Bias reduction is further sought in procedural assessment strategies derived from criteria for qualitative research.

This approach can be considered a paradigm shift in several ways. First it approaches assessment in a similar way as the design and review of the curriculum where elements are planned, arranged and coordinated and where every element is systematically evaluated and

reformed. Second it bases high stake decisions on a longitudinal aggregation of assessment points. Third it employs a wide range of methods for individual data points including qualitative assessments involving expert judgements, verified through established qualitative methods. Fourth it optimises feedback from each assessment point to promote student learning. Individually and together these elements are in sharp contradistinction from established practice where assessment is rarely considered as a system; high stake pass fail decisions are taken on single key assessment points; a limited number of assessment methods are used; expert judgements are minimised; and feedback is limited to numerical or grade scores.

Programmatic assessment is thus an integral approach to the design of an assessment programme with the intent to optimise its learning function, its decision-making function and its curriculum quality-assurance function. Individual methods of assessment, purposefully chosen for their alignment with the curriculum outcomes and their information value for the learner, the teacher and the organisation, are seen as individual data points. The information value of these individual data points is maximised by giving feedback to the learner. There is a decoupling of assessment moment and decision moment. Intermediate and high-stakes decisions are based on multiple data points after a meaningful aggregation of information and supported by rigorous organisational procedures to ensure their dependability. Self-regulation of learning, through analysis of the assessment information and the attainment of the ensuing learning goals, is supported through a mentoring system. Each student is allocated a counsellor who works closely with the student to interpret feedback and plan learning.

While the approach was developed to address problems in medical education, it is clear that its principles and procedures could be applied in other areas of professional training including nursing, the therapies, social work and education.

Implementation of a Programmatic Assessment system requires a number of steps which are now outlined. These have been distilled from the various descriptions and prescriptions produced by the pioneers in the Netherlands and Australia. Each of these steps is desirable in itself and might have a beneficial effect on assessment. However, their totality characterises a Programmatic Assessment system.

A master plan should be produced for assessment analogous to that normally produced for a modern curriculum. This plan should relate to an overarching structure for the curriculum. In medicine this is usually in the form of a competency framework, most recently realised as a description of the professional role as Entrustable Professional Activities. The plan s is important since in programmatic assessment pass/fail decisions are not taken at the level of each individual assessment moment, but only after a coherent interpretation can be made across many assessment moments. An individual assessment can be considered as a single data point. The traditional dichotomy between formative and summative assessment is redefined as a continuum of stakes, ranging from low- to high-stakes decisions. The stakes of the decision and the richness of the information emanating from the data points are related, ensuring proportionality of the decisions: high-stake decisions require many data points. In order to meaningfully aggregate information across these data points an overarching structure is needed, such as a competency framework. Information from various data points can be combined to inform the progress on domains or roles in the framework. For example, information on communication from an objective structured Clinical examination (OSCE) may be aggregated with information on communication from several mini-clinical evaluation exercise (Mini-CEX) and a multisource- feedback tool.

Examination regulations should be devised that promote feedback rather than just pass/fail, progress and award. Individual data points are optimised for providing information and feedback to the learner about the quality of their learning and not for pass/fail decisions. Pass/fail decisions should not be made on the basis of individual data points – as is often the case in traditional regulations. Examination regulations traditionally connect credits to individual assessments; this should be prevented in programmatic assessment. Research has shown that feedback is ignored in assessment regimes with a summative orientation (Harrison et al. 2013). Because aligning g credits to individual assessments raises their stake, learners will primarily orientate themselves to passing the test instead of on feedback reception and follow-up (Bok et al. 2013). Credit points should be linked only to high stake decisions, based on many data points.

A robust system must be devisedis required for collecting and storing information. In programmatic assessment, detailed information about the learner is essential and considerable information is gathered over time. Being able to handle this information flexibly is vital. One way of collecting information is through the use of (electronic) portfolios. Here,

portfolios have a dossier function allowing periodic analyses of the student's competence development and learning goals. The (e-)portfolio should therefore serve three functions: (1) provide a repository of formal and informal assessment feedback and other learning results (i.e. assessment feedback, activity reports, learning outcome products, and reflective reports), (2) facilitate the administrative and logistical aspects of the assessment process (i.e. direct online loading of assessment and feedback forms via multiple platforms, regulation of who has access to which information and by connecting information pieces to the overarching framework), and (3) enable a quick overview of aggregated information (such as overall feedback reports across sources of information). User friendliness is vital. The (e-)portfolio should be easily accessible to whichever stakeholder has access to it. Many e-portfolios are commercially available, but care should has beto be taken to ensure that the structure and functionalities of these portfolios are sufficiently aligned with the requirements of the assessment programme.

The idea of low stake and high stake assessment is fundamental to Programmatic Assessment. High stake assessment decisions such as award of credits; progress on the programme; and award of the final qualification, should, ideally, always be based on the aggregation of a number of low stake assessments. Single low stake assessments should would provide feedback for learning. This feedback should becould be facilitated by the reduction in the stake level of the assessment. Information richness in the feedback is the cornerstone of programmatic assessment. Without rich assessment information programmatic assessment will fail.

Mostly, conventional feedback from assessments, that is, grades and pass/fail decisions, are poor information carriers (Shute 2008). Meaningful feedback may have many forms. One is to give out the test material after test administration with information on the correct or incorrect responses. In standardised testing, score reports may be used that provide more detail on the performance (Harrison et al. 2013), for example, by giving online information on the blueprint categories of the assessment done, or on the skill domains (i.e. in an OSCE), or longitudinal overview for progress test results (Muijtjens et al. 2010). Sometimes verbal feedback in or after the assessment may be given (Hodder et al. 1989). In unstandardized assessment, quantitative information usually stems from the rating scales being used. This is useful, but it also has its limitations. Feedback for complex skills is enhanced by narrative information (Govaerts et al. 2007). Narrative information may also enrich standardised assessment. For

example, in one implementation of programmatic assessment narrative feedback is given to learners on weekly open-ended questions (Dannefer & Henson 2007). Given the fact that putting a metric on things that are difficult to quantify may actually trivialise what is being assessed. Metrics such as grades often lead to unwanted side effects like grade hunting and grade inflation. Grades may unintentionally ‘‘corrupt’’ the feedback process. Some argue we might should replace scores with words (Govaerts & van der Vleuten 2013), particularly in unstandardized situations where complex skills are being assessed such as in clinical workplaces.

In Programmatic Assessment each student is allocated a mentor. Now that the student has access to rich feedback from each low stake assessment on a longitudinal basis, support for its interpretation and use is essential. Feedback alone may not be sufficient for learners to be heeded well by students. (Hattie & Timperley 2007). Research findings clearly indicate that feedback, reflection, and follow-up on feedback are essential for learning and expertise development (Ericsson 2004; Sargeant et al. 2009). Reflection for the mere sake of reflection is not well received by learners, but reflection as a basis for discussion and action is appreciated (Driessen et al. 2012). Feedback should ideallyshould, ideally, be part of a (reflective) dialogue, stimulating follow-up on feedback. Mentoring is an effective way to create such a dialogue and has been associated with good learning outcomes (Driessen & Overeem 2013). In programmatic assessment mentoring is used to support the feedback process and the feedback use. In a dialogue with an entrusted person, performance may be monitored, reflections shared and validated, remediation activities planned, and follow-up may be negotiated and monitored. This is the role of a mentor. The mentor is a regular staff member, preferably having some knowledge over the curriculum. Mentor and learner meet each other periodically. It is important that the mentor is able to create a safe and trusted relationship. For that purpose, the mentor should be protected from having a judgemental role in the decision-making process (Dannefer & Henson 2007). The mentor’s function is to get the best out of the learner. In conventional assessment programmes, adherence to minimum standards can suffice for promotion and graduation. In programmatic assessment, individual excellence is the goal and the mentor is the key person to promote such excellence.

High stake decisions with their aggregation of multiple data points require trustworthy decision making. The assuring of this requires that procedural measures should be put in place that bring evidence to this trustworthiness. These procedural measures may include (Driessen et

al. 2013):

- An appointment of an assessment panel or committee responsible for decision-making (pass–fail–distinction or promotion decisions) having access to all the information, for example, embedded in the e-portfolio. Size and expertise of the committee will matter for its trustworthiness.
- Prevention of conflicts of interest and ensuring independence of panel members from the learning process of individual learners.
- The use of narrative standards or milestones.
- The training of committee members on the interpretation of standards, for example, by using exceptional or unusual cases from the past for training purposes.
- The organisation of deliberation proportional to the significance of information. Most learners will require very little time; very few will need considerable deliberation. A chair would be responsible for preparing efficient sessions.
- The provision of justification for decisions with high impact, by providing a paper trail on committee deliberations and actions, that is documented very carefully.
- The provision of mentor and learner input. The mentor knows the learner best. To eliminate bias in judgement and to protect the relationship with the learner, the mentor should not be responsible for final pass–fail decisions.
- Provision of an appeals procedures.

This list is not exhaustive, and it is helpful to think of any measure that would stand up in court, such as factors that provide due process in procedures and expertise of the professional judgement. These usually lead to robust decisions that have credibility and can be trusted.

Intermediate decision making assessments should could be organised. Ideally, hHigh-stakes decisions at the end of the course, year, or programme should never be a surprise to the learner. Therefore, provision of intermediate assessments informing he learner and prior feedback on potential future decisions is in fact another procedural measure adding to the credibility of the final decision. Intermediate assessments are based on fewer data points than final decisions. Their stakes are in between low-stake and high-stake decisions. Intermediate assessments are diagnostic (how is the learner doing?), therapeutic (what should might be done to improve further?), and prognostic (what might happen to the learner; if the current development continues to the point of the high-stake decision?). Ideally, an assessment committee provides all inter- mediate evaluations, but having a full committee assessing all students may well be a too resource-intensive process. Less costly compromises are to be considered, such as using subcommittees or only the chair of the committee to produce these

evaluations, or having the full committee only looking at complex student cases and the mentors evaluating all other cases.

Programmatic Assessment encourages and facilitates personal remediation, that is the student helping him/her self with support from their mentor and other faculty. Remediation is essentially different from resits or supplemental examinations. Remediation is based on the diagnostic information emanating from the on-going reflective processes (i.e. from mentor meetings, from intermediate evaluations, and from the learner him/her self) and is always personalised. Therefore, the curriculum must has to provide sufficient flexibility for the learner to plan and complete remediation. There is no need for developing (costly) remediation packages. Engage the learner in making decisions on what and how remediation should might be carried out, supported by an experienced mentor. Ideally, remediation is made a responsibility of the learner who is provided with sufficient support and input to achieve this.

Just as the curriculum should have routinizedrequires routinized monitoring to identify problems and an improvement cycle to address them, so the assessment programme should could be monitored and evaluated. This monitoring should,could, inter alia, consider the learning effects of the programme and take steps to address problems and encourage improvement. Such an approach is well established with regard to the curriculum but rarely followed for assessment.

Of course the assessment evaluation should feedfeeds into curriculum evaluation. Assessment may serve three functions: to promote learning, to promote good decisions on whether learning outcomes are achieved, and to evaluate the curriculum. In programmatic assessment, the information richness is a perfect basis also for curriculum evaluation. The assessment data gathered, for example, in the e-portfolio, not only provides an X-ray of the competence development of the learners, but also on the quality of the learning environment.

As should be clear from the previous guidelines, programmatic assessment impacts at all levels: students, examiners, mentors, examination committees, assessment developers, and curriculum designers. Programmatic assessment is, therefore, the responsibility of the whole educational organisation. When implemented, frequent and on-going communication between the different stakeholder groups is essential in the process.

There is no doubt that programmatic assessment requires a culture change in thinking about assessment that is not easy to achieve in an existing educational practice. Traditional assessment is typically modular, with summative decisions and grades at the end of modules.

When passed, the module is completed. When failed, repetition through resits or through repetition of the module is usually the remedy. This is all very appropriate in a mastery learning view on learning. However, modern education builds on constructivist learning theories, starting from notions that learners create their own knowledge and skills, in horizontally and/or vertically integrated programmes to guide and support competence. Programmatic assessment is better aligned to notions of constructivist learning and longitudinal competence development through its emphasis on feedback, use of feedback to optimise individual learning and remediation tailored to the needs of the individual student. This radical change often leads to fear that such assessment systems will be soft and vulnerable to gaming of students, whereas the implementation examples demonstrate the opposite effect (Bok et al. 2013). Nevertheless, for this culture change in assessment a change strategy is required, since many factors in higher education are resistant to change (Stephens & Graham 2010). A change strategy needs to be made at the macro-, meso- and micro levels.

At the macro level, national legal regulations and university regulations are often strict about assessment policies. Some universities prescribe grade systems to be standardised across all training programmes. These macro level limitations are not easy to influence, but it is important to know the ‘‘wriggle room’’ these policies leave for the desired change in a particular setting. Policy-makers and administrators need to become aware of why a different view on assessment is needed. They also need to be convinced on the robustness of the decision-making in an assessment programme. The qualitative ontology underlying the decision-making process in programmatic assessment is a challenging one in a positivist medical environment. Very important is to explain programmatic assessment in a language that is not jargonistic and which aligns with the stakeholder’s professional language. For clinicians, for example, analogies with diagnostic procedures in clinical health care often prove helpful.

At the meso level programmatic assessment may have consequences for the curriculum. Not only should the assessment be aligned with the overarching competency framework, but with the curriculum as well. Essential are the longitudinal lines in the curriculum requiring a careful balance of modular and longitudinal elements. Individual stakeholders and committees need to be involved as early as possible. Examination rules and regulations need to be constructed which are optimally transparent, defensible, but which respect the aggregated decision-

making in programmatic assessment. The curriculum also needs to allow sufficient flexibility for remediation. Leaders of the innovation need to be appointed, who have credibility and authority.

Finally, at the micro level teachers and learners need to be involved in the change right from the start. Buy-in from teachers and learners is essential. To create buy-in the people involved should understand the nature of the change, but more importantly they should be allowed to see how the change also addresses their own concerns with the current system. Typically, teaching staff do have the feeling that something in the current assessment system is not right, or at least suboptimal, but they do not automatically make the connection with programmatic assessment as a way to solve these problems.

The development of programmatic assessment is a learning exercise for all and it is helpful to be frank about unexpected problems to arise during the first phases of the implementation; that is innate to innovation. So it is therefore good to structure this learning exercise as a collective effort, which may exceed traditional faculty development (De Rijdt et al. 2013). Although conventional faculty development is needed, involving staff and students in the whole design process supports the chance of success and the creation of ownership (Konings et al. 2005) and creates a community of practice promoting sustainable change (Steinert 2014).

Changing towards programmatic assessment can be compared with changing traditional programmes to problem-based learning (PBL). Many PBL implementations have failed due to problems in the implementation (Dolmans et al. 2005). When changing to programmatic assessment, careful attention should be paid to implementation and the management of change at all strategic levels.

Programmatic assessment has a clear logic and is based on many assessment insights that have been shaped through research and educational practice. Logic and feasibility, however, are inversely related in programmatic assessment. To introduce full-blown programmatic assessment in actual practice all stakeholders need to be convinced. This is not an easy task. Just like in PBL, partial implementations are possible with programmatic assessment (i.e. the increase in feedback and information in an assessment programme, mentoring). Again as in PBL, this will lead to partial success.

The problems which led to the devising of programmatic assessment are not unique to medical education. We would suggest that the approaches outlined above could with profit be considered for any programme of professional education.

References

Bok HG, Teunissen PW, Favier RP, Rietbroek NJ, Theyse LF, Brommer H, Haarhuis JC, van Beukelen P, van der Vleuten CP, Jaarsma DA. 2013. Programmatic assessment of competency-based workplace learning: When theory meets practice. BMC Med Educ 13:123.

Cilliers FJ, Schuwirth, LW Adendorff HJ (2010) The mechanism of impact of summative assessment on medical students' learning Advances in Health Sciences Education Issue 5, pp 695–715

Cilliers FJ Schuwirth LWT Herman N (2012) A model of the pre assessment learning effects of summative assessment in medical education Advances in Health Issue 1 39-53

Dannefer EF, Henson LC. 2007. The portfolio approach to competency- based assessment at the Cleveland Clinic Lerner College of Medicine. Acad Med 82:493–502.

De Rijdt C, Stes A, van der Vleuten C, Dochy F. 2013. Influencing variables and moderators of transfer of learning to the workplace within the area of staff development in higher education: Research review. Educ Res Rev 8; 48-74

Dijkstra J, Galbraith R, Hodges BD, McAvoy PA, McCrorie P, Southgate LJ, van der Vleuten CP, Wass V, Schuwirth LW. 2012. Expert validation of fit-for-purpose guidelines for designing programmes of assessment. BMC Med Educ 12:20.

Dolmans DH, De Grave W, Wolfhagen IH, van der Vleuten CP. 2005. Problem-based learning: Future challenges for educational practice and research. Med Educ 39:732–741.

Driessen EW, Heeneman S, van der Vleuten CPM. 2013. Portfolio assessment. In: Dent JA, Harden RMM, editors. A practical guide for medical teachers. Edinburgh: Elsevier Health Sciences. pp 314–323.

Driessen EW, Overeem K. 2013. Mentoring. In: Walsh K, editor. Oxford textbook of medical education. Oxford: Oxford University Press. pp 265–284.

Driessen EW, van Tartwijk J, Govaerts M, Teunissen P, van der Vleuten CP. 2012. The use of programmatic assessment in the clinical workplace: A Maastricht case report. Med Teach 34:226–231.

Ericsson KA. 2004. Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. Acad Med 79: S70–S81.

Frye AW, Hemmer PA. 2012. Program evaluation models and related theories: AMEE Guide No. 67. Med Teach 34: e288–e299.

GMC (2009) Assessment in Undergraduate Medical Education London UK

Govaerts MJB, van der Vleuten CPM. 2013. Validity in work-based assessment: Expanding our horizons. Med Educ 47:1164–1174.

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 12:239–260.

Harrison CJ, Ko ¨nings KD, Molyneux A, Schuwirth LW, Wass V, van der Vleuten CP. 2013. Web-based feedback after summative assessment: How do students engage? Med Educ 47:734–744.

Hattie J, Timperley H. 2007. The power of feedback. Rev Educ Res 77: 81–112.

Hodder RV, Rivington R, Calcutt L, Hart I. 1989. The effectiveness of immediate feedback during the objective structured clinical examination. Med Educ 23:184–188.

Konings KD, Brand-Gruwel S, Merrie¨ nboer JJ. 2005. Towards more powerful learning environments through combining the perspectives of designers, teachers, and students. Br J Educ Psychol 75:645–660.

Muijtjens AM, Timmermans I, Donkers J, Peperkamp R, Medema H, Cohen- Schotanus J, Thoben A, Wenink AC, van der Vleuten CP. 2010. Flexible electronic feedback using the virtues of progress testing. Med Teach 32: 491–495.

Ruhe V, Boudreau JD. 2013. The 2011 program evaluation standards: A framework for quality in medical education programme evaluations. J Eval Clin Pract 19:925–932.

Sargeant JM, Mann KV, van der Vleuten CP, Metsemakers JF. 2009. Reflection: A link between receiving and using assessment feedback. Adv Health Sci Educ: Theory Pract 14:399–410.

Schuwirth LWT, van der Vleuten CPM. 2011. Programmatic assessment: From assessment of learning to assessment for learning. Med Teach 33: 478–485.

Schuwirth LW, van der Vleuten CP. 2012. Programmatic assessment and Kane's validity perspective. Med Educ 46:38–48.

Shute VJ. 2008. Focus on formative feedback. Rev Educ Res 78: 153–189.

Steinert Y. 2014. Medical education and faculty development. The Wiley Blackwell Encyclopedia of Health, Illness, Behavior, and Society. pp 1344–1348.

Stephens JC, Graham AC. 2010. Toward an empirical research agenda for sustainability in higher education: Exploring the transition management framework. J Cleaner Prod 18:611–618.

van der Vleuten CPM. 1996. The assessment of professional competence: Developments, research and practical implications. Adv Health Sci Educ 1:41–67.

van der Vleuten CP, Schuwirth LW. 2005. Assessing professional competence: From methods to programmes. Med Educ 39:309–317.

van der Vleuten CP, Norman GR, De Graaff E. 1991. Pitfalls in the pursuit of objectivity: Issues of reliability. Med Educ 25:110–118.

van der Vleuten CP, Schuwirth LW, Driessen EW, Dijkstra J, Tigelaar D, Baartman LK, van Tartwijk J. 2012. A model for programmatic assessment fit for purpose. Med Teach 34:205–214.

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.

Watling C, Driessen E, van der Vleuten CP, Lingard L. 2012. Learning from clinical work: The roles of learning cues and credibility judgements. Med Educ 46:192–200.