Action research: towards excellence in teaching, assessment and feedback for clinical consultation skills

Dr Janet Lefroy

PhD by published works

June 2018

Keele University

Abstract

Background   Consultation skills are the core competencies required at graduation of the doctor as a practitioner. Every medical school has its own system of teaching and assessing consultation skills. These are generally amalgams of previous curricula and not rigorously developed. We took the opportunity presented by a new undergraduate medical curriculum to systematically develop the consultation skills curriculum from classroom teaching to OSCE assessment and formative workplace-based assessment and feedback.

 Methods   The consultation skills curriculum and assessment system were developed by action research. Data were collected using mixed methods involving questionnaires, focus groups, participant interviews, student reflective summaries and routine monitoring of usage of an app which we developed for generating feedback summaries in the clinical workplace. Participants were teachers and students at Keele University school of medicine. In addition, clinical tutors from seven other UK medical schools participated in a Delphi study of undergraduate medical consultation skills competencies.

Results   A case study of curriculum development by action research is presented in nine published papers.

Conclusion   This work has contributed to medical education knowledge as follows: an instrument for assessment of consultation skills has been developed and validated; and a set of strategies for improvement of these consultation skills have been developed and validated. It has added to understanding about transfer of learning from the classroom to the workplace; the impact of assessment grades on medical students' learning and self-perception; and the value of a system of formal workplace-based assessment.

Additionally this work was one of the first applications of realist methods in medical education research, and it has developed guidance on feedback in the workplace for individual tutors and educational institutions.

Key words

Undergraduate medical education; Curriculum development; Competency Based Medical Education; Consultation skills; Workplace assessment; Feedback; Feedback relationship; Feedback culture; Technology-enhanced learning; Action research; Realist evaluation

Contents

Acknowledgements

The published works arose from the efforts of several teams of action researchers. Each publication is authored by the team members who were involved in the design of the study; who assisted in data collection and/or analysis and who also contributed to the correction of drafts of the papers. Other researchers who contributed at the design stage, helped with interviewing or focus group facilitation, or peer reviewed the draft version are acknowledged in each paper. The many student and tutor participants have also been thanked in the publications. I would like to reiterate here how Keele medical school staff and students have understood their role in action research and been ready with evaluation and suggestions. The school have then trusted our findings and have taken up the proposed changes with heartening enthusiasm.

My contribution to these cycles of action research have been as the convener of the research team for each cycle. I wrote the initial drafts of research protocols and put the protocols through the ethical approval process. I organised data collection either by independent researchers whom I supervised or by team members (including myself in two studies). I wrote the initial draft of seven of the nine papers, and contributed to the revisions of the other two.

I was not the most senior researcher in the team and owe a huge debt of gratitude to Bob McKinley for suggestions about team membership and research strategy, and to Sarah Yardley for her clear and perceptive analysis of data. I am also grateful to Adrian Molyneux who designed the IT systems for the feedback instruments we developed. I always found that my ideas were improved by team discussion and thank all the team members who rallied to my invitations, didn't mind my being bossy, and participated in research meetings, data analysis and editing of draft papers usually in their own time after work. It has been an almost universal pleasure to work on each of these studies. Never having done research solo I have no comparator but I can commend team research unconditionally. I am also surprisingly grateful to the reviewers of journals which rejected drafts of some of these papers and thereby assisted us to improve them significantly. Peter Yeates has been co-supervisor of this PhD and has brought perspective and clarity to my thinking on drafts of this thesis.

Finally I dedicate this work to my husband Jeremy who has encouraged me every step of the way, and to my mother Valerie MacKay whose careful scholarship in researching for a biography of my father these past ten years has inspired my efforts.

Janet Lefroy

Chapter 1

Introduction

Summary of the critical commentary

The critical commentary is in two parts (Chapters 1 and 11) with the published works sandwiched between (Chapters 2-10). In Chapter 1 I describe how this work came about. In Chapter 11 I critique the research methodologies and methods used. I re-consider what each paper has added to knowledge at the time, and any evidence of subsequent impact. Finally, I consider future research questions.

1.1 Context for this research

> *"The intended output of a competency-based programme is a health professional who can practise medicine, at a defined level of proficiency, in accord with local conditions, to meet local needs"*
>
> W.C. McGaghie et al. Competency based curriculum development in medical education. An introduction. Public Health Papers No 68 World Health Organisation (1978) pg18.

In 2007 when we started the curriculum development project that lies behind these publications, Competency-Based Medical Education (CBME) was being adopted as the guiding principle of new and revised undergraduate and postgraduate curricula in the UK. This CBME focus was promoted by the GMC in Tomorrow's Doctors 2003 (1), in the USA with the Accreditation Council for Graduate Medical Education (ACGME)'s Outcomes Project (2), in Canada with the CanMEDS project (3), and in continental Europe via the Tuning Project (4). CBME challenged the assumption that competencies such as communication, professionalism and leadership are acquired naturally as trainees pass through clinical attachments. In CBME competencies are the defined outcomes of training programmes, with a local approach to the level of each competency required to meet the needs of the population served. Panels of local experts in various countries developed schedules of required competencies to guide the training and assessment of their doctors (5). By the time Keele University's faculty were developing our desired outcome competencies for Keele medical school graduates, the UK General Medical Council's (GMC) updated guide to undergraduate medical education had been published (Tomorrow's Doctors (2009) which emphasised the importance of the 'Doctor as a practitioner' and, in paragraphs 13 to 15, described the skills the medical graduate needs to acquire to consult with patients (6). These were a mix of cognitive, psychomotor, communication and interpersonal skills.

Competency frameworks from North America had much in common with each other, especially where they described competencies required for the core activity of the doctor – the clinical consultation. The 1999 Kalamazoo conference's consensus statement (7) identified "essential elements of communication in medical encounters" as agreed by participants from North American bodies: ACGME, the CanMEDS 2000 Project, the Educational Commission for Foreign Medical Graduates (ECFMG), and the Macy Health Communication Initiative. They focused on the communication tasks of the doctor in five models of doctor-patient communication which were in current use.

This emphasis on teaching and assessing communication in its clinical context had been taken further by some undergraduate medical schools who advocated the integration of learning and assessment of clinical communication with other clinical skills such as clinical reasoning and physical examination and also with procedural skills in order to improve the authenticity and therefore the usefulness of the learning (8–11).

At Keele University the school of medicine used the Manchester MBBS curriculum from 2002 until 2007 following which we developed and implemented our own curriculum. Our challenge was to optimize the existing communication skills curriculum, identifying what was working and what needed to change in the light of our own experience and that of others. This was important because consultation skills are core to the practice of medicine. There was a problem of poorly-aligned curricular elements - classroom teaching, placement expectations of student consulting, and assessment in Objective Structured Clinical Examinations. Such problems are common to many medical schools nationally and internationally, and therefore worthy of study.

As the lead for communication skills teaching at that time, I drew on the expertise and viewpoints of a group of stakeholders forming an action research group to develop the communication skills curriculum. The problems reported in teaching team debriefs were addressed first, viz: it was difficult to facilitate group learning of communication skills for students with diverse prior experience; some first year students found roleplay patient interviews terrifying; year 3 to 5 students who were already in clinical practice sometimes felt that our classes were unnecessary; students felt a disconnect between classroom teaching and placement consulting. The first cycle of this action research project was reported in my MMedEd dissertation (12) and formed the starting point for further research.

Our main challenge was to align curricular objectives with teaching and learning and with assessment in the new curriculum. The communication model we were using (Calgary Cambridge) was for communication only and didn't include other competencies. On placements students were practising the skills of the doctor in a more integrated way and needed teaching aligned with this. We were assessing consultation skills but had no validated assessment instrument for our assessments. Keele's communication skills curriculum was renamed the "consultation skills curriculum" as our reflective revisions of the curriculum adopted the integrated approach (8,9). We drew on four consultation frameworks current at that time – Tomorrow's Doctors 2009 (6), The Calgary-Cambridge Guide to the Medical Interview (13,14), The Kalamazoo consensus statement (7) and the Leicester Assessment Package (11).

Alongside the challenge of developing our core consultation skills curriculum competency framework and assessment instrument, we were also facing the question of how students could get the most helpful feedback on their consultation skills. We had noticed that students were generally not giving each other critical-constructive feedback in skills classes. Our student satisfaction surveys indicated a perception of inadequate feedback. Workplace assessment of consultation skills with feedback can be a powerful method of learning (15). The task of developing a system of formative workplace assessment and feedback for the school of medicine to take learning of the curriculum into placements was therefore the next problem addressed. As we implemented this system, we evaluated it with the users and went through several cycles of change. This change had two facets - the system itself underwent change and also the clinical

tutors using the system also changed how they gave their feedback. The publications in chapters 6 and 7 show how some of the problems in designing a useful system were addressed. Two questions we were asking were how to devise good quality, acceptable bespoke feedback, and should we include grading? Once the system was rolled out and clinical tutors adopted it with more or less enthusiasm, it was evaluated and was deemed to be contributing positively towards the school's aim of graduating excellent clinicians. There were, however, unexpected findings on evaluation which we turned into problems for further study. One puzzle was that while students said they were highly satisfied with the feedback they were getting from their clinical tutors, paradoxically we discovered that some were not actually accessing their feedback summaries in the electronic feedback portal. Another puzzle for the action research group was that the app for students' mobile devices - devised to make feedback both student-led and easier (by use of the voice recognition feature of mobile devices) - was largely neglected in favour of the computer version. The publications in chapters 8 and 9 outline how our understanding was changed by exploring and reflecting on these unexpected evaluation findings.

In order to help other workplace clinical tutors and their institutions to provide better feedback, I led a literature review commissioned by *Perspectives in Medical Education* producing novel evidence-based guidelines about feedback in the clinical workplace for tutors and institutions (16). This is outlined in the third section of the literature review which follows below.

In summary, the problems we wished to address were:
How to improve our teaching of consultation skills in the classroom to equip students for placements and future practice; how to ensure that our students get better feedback once they are in the clinical workplace; and how to align teaching and placement feedback with medical school assessment of consultation skills. The vision for the new Keele curriculum was "Graduating excellent clinicians". Since the consultation is the core activity of most doctors, it was vitally important to work out and agree upon our core consultation competencies, and to find how our students and our large and disparate faculty might be encouraged to adopt them.

1.2 Rationale for choosing an action research approach to curriculum development

---

A definition of action research

"It can be argued that three conditions are individually necessary and jointly sufficient for Action Research to be said to exist: firstly, a project takes as its subject-matter a social practice, regarding it as a form of strategic action susceptible of improvement; secondly, the project proceeds through a spiral of cycles of planning, acting, observing and reflecting, with each of these activities being systematically and self-critically implemented and interrelated; thirdly, the project involves those responsible for the practice in each of the moments of the activity, widening participation in the project gradually to include others affected by the practice, and maintaining collaborative control of the process"

Carr & Kemmis 1986 (165-166) (17)

We chose an action research approach to the problems we faced in our previous curriculum because we understood that the reflect-plan-act-observe cycle of curriculum action research would help us to develop systematically (18,19). Drawing on the definition of action research from Carr & Kemmis quoted above, teaching and learning is a complex social practice and therefore better suited to quality improvement by problem-solving and innovation involving critical reflection rather than by educational research based around simple hypothesis-building and testing. Indeed, one of the two major traditions of action research has been developed in education, linking research to improvement of practice (the other links research to bringing about social change) (20). Knowing that our curriculum was capable of improvement, we wanted to reflect self-critically at the start of the process of curriculum change. We decided to use existing knowledge from the literature (the testimony of others) alongside our own experiences to evaluate critically what we were doing and make changes likely to bring benefit to our students (21). As we made changes to teaching, assessment and systems for workplace learning, we wished to evaluate critically the outcomes – intended or other, in order to improve again. We also wished to reflect self-critically on our teaching (a process of professional development or 'action learning' (22)). One hazard of curriculum change is that the written curriculum becomes increasingly divorced from what students actually experience because change is brought about by those at the 'top' of the system without involving teachers in designing it, and therefore fails to get implemented as intended because teachers have not understood or believed in it (22,23). We considered that a collaborative action research approach to change was most likely to succeed and for this we felt that our action research team should involve student and simulated patients as well as interested members of faculty. There could be added benefits to making good curricular changes - faculty engagement in the observation and reflection processes would be good for team engagement and development; and if we did find something worked well, the findings could be disseminated to help colleagues facing similar difficulties elsewhere (22).

1.3 Critical literature review – what was already known about teaching and learning of consultation skills, assessment of consultation skills and feedback in the clinical workplace

Each published study in this thesis contains its own literature review, specific to the focus and timing of the study. The following brief literature overview aims to give the evolving context for the whole body of research reported in this thesis. It describes how we located our understanding of the issues, and offers a critical analysis of how we applied the theories of others to our practices or identified gaps in the literature. Since the body of work spans ten years, some of the contextual literature was published during that time. Rather than write a review of the literature as it was in 2007 and an additional review of advances to 2017, I have summarised knowledge in 2017. Our own work is omitted from this contextual review but is cited where relevant in the published studies. The literature review is divided into three sections, the first covering teaching and learning of consultation skills, the second covering assessment of consultation skills, and the third reviews the literature we drew on for our guidelines for feedback in the clinical workplace.

Critical engagement with the literature is encouraged in action research. McNiff describes this as interrogating your own thinking and that of others (24). In critically appraising the evidence for medical education outcomes, in addition to examining quality of reporting, methodological rigour and conceptual depth and breadth, we must be cognisant that medical education research is essentially a social science. Some of the theory and evidence drawn on therefore comes from social groups outside of medical education. While there is much that will transfer from one educational context to another, some caution is required in transferring what we know from research in school children or in non-clinical higher education settings into the medical education setting.

1.3.1 What was known about teaching and learning of consultation skills

Our team developing the consultation skills curriculum adopted the understanding that skills are probably learned by a process of: initial awareness – awkwardness – conscious skill – integrated skill (13). High-fidelity simulation with feedback to assist learners to move through these stages has been used in various clinical simulation settings and is said to be an effective learning method (25,26) although much of the evidence is low-level evaluative research. Simulation is used in medical learning to allow students to experiment safely, to rehearse skills and to get immediate feedback in a supportive environment (8,27). We considered that having had this experience, students would then be able to move on safely to consultation practice in the clinical workplace. What happens next also appears to matter if students are to achieve excellence in the newly acquired skills. Failing to practice a new skill soon after learning it can lead to the skill's being lost, as indicated for example by a systematic review of training in the skills of resuscitation (28). The fact that this review was conducted under the auspices and to the standards set by Best Evidence Medical Education (BEME) gives it credence, as far as resuscitation skills are concerned at least. What about other clinical skills? Based on studies in musicians, sciences and sport, Ericsson concludes that deliberate repetitive practice is a key aspect of attaining expertise in a range of technical and cognitive skills and applies this to medical expertise also (29). We used this as guidance for the organisation our curriculum: each skill is taught at a time when students will be able to practice it or at least witness its use.

Learning should also build on existing knowledge and skills (30,31). We found Vygotsky's work very helpful in our conceptualisation of the way learning is constructed. Although he studied group learning in children, his theories have been developed in adult education. We can therefore be confident that they apply in medical education as a theoretical grounding (32). Learners differ in how far they are able to move beyond existing competence – in Vygotsky's term the "width of their zone of proximal development"(33). One of the tutor's roles is to help the learner progress to the next level of achievement by appropriate steps, not to expect him/her to bridge an unbridgeable gap. This has influenced our development of teaching methodology evaluated in the study in Chapter 2.

When preparing lesson plans for group learning we were influenced by the "social cognitive learning theory" (34). Learning of social skills is thought to be based on watching and imitating others. Learners are thought to gauge their capabilities not only through their own performance but also by observing similar others perform, by persuasion from others ("you can do it") and by experiencing their own physiological indexes (heart rate, sweating). We were of course aware

that students (and doctors) differ in their self-confidence, risk-taking, metacognition (self-awareness of reasoning), mastery and performance goals and this affects how individuals learn and how much support and challenge each requires (35–38). Williams and Deci, the authors of self-determination theory reviewed studies of teaching which supports autonomy. Having found similar results in four medical schools to those from other higher education and school settings they concluded that the theory holds in all these settings. The ability of such longitudinal studies to show causality is limited but they were reported to indicate a modest and sustained effect of autonomy-supportive instruction on student development of autonomy and on observer-rated autonomy supportive interviewing of simulated patients. (39). We wished to apply this theory to our group skills learning exercises.

The clinical workplace has been reported to be a good place to hone consultation skills in the early years of medical education and Dornan and others have studied early clinical experience in depth. In a systematic review which used the BEME process 73 empirical studies of early clinical experience in the basic education of health professionals met the selection criteria and were examined (40) While half of these studies were found to be methodologically poor, in the other half evidence was found for several benefits including on student consultation skills but it was not clear how authentic early experience can be made most educationally effective . Dornan went on to conduct his own research which indicated that 'supported participation' is a core condition for such workplace learning (41). In a methodologically strong study, year 3 students when first introduced to the clinical environment consistently reported that they needed tutor support to make sense of their learning environment (42). We wanted to understand these perceptions of difference better, and whether our students could be helped to transfer learning. This led to the study in Chapter 3.

1.3.2 What was known about assessment of consultation skills

The medical education literature did not provide us with a validated generic consultation skills assessment instrument for our undergraduate medical curriculum. The Calgary Cambridge guide to the medical interview is used by many UK medical schools including Keele as the basis of their communication skills curricula (13,14). It does not, however, address the additional cognitive skills required for making a diagnosis or identifying appropriate management options and, although some congruent assessment schedules have been developed, they are context specific and have not been widely evaluated (13). In search of a more comprehensive consultation competency assessment instrument, the Leicester Assessment Package (LAP) (11) was considered. It has been used in both formative and summative assessment of undergraduates (43,44) and by established practitioners (45,46) in the UK and internationally. It does not, however, map onto a particular model of the consultation.

The mini-CEX assessment instrument was developed in 1995 for workplace assessment of first year US residents (47). Its seven domains (Medical interviewing skills, Physical examination skills, Humanistic qualities (professionalism), Clinical judgement, Counselling skills, Organisation/efficiency, and Overall clinical competence) provided a manageable generic framework for workplace assessment and feedback on consultation skills. It has the strongest validity evidence of all the tools for workplace assessment (48) and is now used in workplace assessment and feedback for undergraduate medical students as well as for doctors. It assumes

knowledge of the competencies in each domain, however, so does not provide the granularity needed by junior students learning these competencies.

1.3.3 What was known about feedback in the clinical workplace

Important literature about formative feedback in the clinical workplace up to 2016 was reviewed to produce guidelines (Do's, Don'ts and Don't Knows) (Chapter 10). This contextual literature influenced our action research group when devising classroom and workplace feedback processes. It has also influenced the institutional feedback culture. The key points are outlined here and further developed in Chapter 10.

Understanding the nature of feedback:
We felt that the commonly-used definition for feedback: 'specific information about the comparison between a trainee's observed performance and a standard, given with the intent to improve the trainee's performance' (49) was lacking the element of social interaction. Other medical educators have drawn on the importance of socialisation into the community of practice. Sargeant et al for example, in an overview of their series of six in-depth qualitative studies of family physician experiences of multi-source feedback, described feedback as part of a social interaction between colleagues which helps the individual to understand their own performance by comparison of their self-perception with the views of the other (50). We adopted the view that feedback might be better perceived as a conversation rather than as a commodity.

Credibility matters:
A number of qualitative studies have shown that learners value feedback that they deem to be credible, but may dismiss feedback that they perceive to lack credibility (51–54). In a Cochrane review of 70 high quality studies on the effect of audit and feedback on healthcare professional practice, Ivers et al (55) found that feedback may be most effective when it is provided by a supervisor or colleague; more than once; in both verbal and written formats; and when it includes both explicit targets and an action plan. This provided the model for the workplace feedback system which we developed.

The timing of feedback affects its usefulness:
Providing feedback long after the event is rarely perceived by learners as useful (51,56). The concept of optimal timeliness however is nuanced and appears to depend on the task and the level of feedback. Hattie and Timperley (57) provided evidence from school and higher education that different levels of feedback deserve different timing. In medical education there is also evidence. For example, a well-designed study of simulation training of procedural skills randomised 30 novice endoscopists to have feedback during or after each of 12 simulated colonoscopies. They found that feedback at the end of the task performance was more effective for learning than feedback during the task performance (58). Awareness of these studies made us consider the timing of workplace feedback. Was it is better for students to get feedback during or immediately after each consultation? Or better to have feedback on their most important strengths and weaknesses from assessment of a number of consultations – at the end of the

week, for example? It certainly made us aware that giving feedback after the student has forgotten the event is likely to be wasted effort.

What is the best feedback method?

Several approaches to feedback have been described in the medical education literature as outlined in Chapter 10 (13,59–62), each with a convincing theory-based rationale, but no head-to-head studies have been done to demonstrate superior effectiveness of one over another. We have adopted a form of agenda-led outcome-based analysis for our small group feedback on consultation skills (13)

Evidence about the content of feedback:

Eva et al in a theoretically informed, methodologically rigorous international study conducted focus groups in 8 institutions with undergraduate and postgraduate medical and midwifery participants. The study concluded that feedback has no recipe but needs to be tailored to the individual trainee's perceptions (63). Van de Ridder et al conducted a rigorous meta-review of 46 studies on feedback across disciplines on the variables which affect the impact of feedback (64). They found evidence that feedback works when it is directed at unsatisfactory elements of performance and is linked to specific learning aims, when it does not threaten the receiver's self-esteem, and when the message is encouraging, specific, detailed and frequently given.

Lack of specificity has been identified as a common weakness of feedback in medical training. In an online survey of UK doctors in one foundation school, the majority of respondents did not feel that the mini-CEX exercise was a useful part of their training and commented that this was especially so if it was a tick box exercise with little effort put in by their assessors, although some did value the feedback comments made and the majority felt the mini-CEX should continue to be (better) used (65). Caution is needed in interpretation of these results with a response rate of 20% making reporter bias likely. There is evidence that the quality of feedback improves after specific training of clinical faculty (65–67).

The focus of feedback also matters. Kluger and DeNisi's rigoursly developed feedback intervention theory which was derived from their meta-analysis of over 130 studies of feedback interventions in various settings (mostly non-clinical), suggests that feedback becomes less effective as attention shifts away from the task and toward the individual; in short, feedback that is threatening to self-esteem is unlikely to be effective (68).

Who should initiate feedback?

A Cochrane review of the effects of audit and feedback on professional practice found moderate quality evidence that poorer performers appear to benefit most from feedback (55). Poorer performers may however be the least likely to seek feedback because the complex psychological influences on feedback-seeking include fear of harm to self-esteem and the desire to impress, not only the desire for information about how to improve (69–71). The assessment made by the learner of the potential value of feedback information is influenced by goal orientation: individuals with an incremental (learning) goal orientation being more likely to seek feedback than those with a static (performance) goal orientation (72,73). This might suggest that the tutor should take the initiative to give feedback to those students who they feel need it and are unlikely to ask. Experimental studies show however that although goal orientation is a fairly stable concept, a learning goal orientation can be fostered by the supervisor and feedback is probably

more effective when it has been solicited (74,75). This influenced our thinking when designing the system of workplace feedback.

Evidence about the motivating impact of feedback:
Psychological orientation not only influences the value attached to feedback and therefore the likelihood of feedback-seeking, it also affects the response to feedback. Feedback works best for learning when the trainee has learning goals rather than performance goals (69,76). We noted feedback intervention theory (68) which suggests that feedback which threatens self-esteem is much less likely to be effective. Eva (63) demonstrated the importance of allowing the learner to maintain their self-concept when receiving feedback, in a study of the interpretation and acceptance of feedback. It seemed that the supportive feedback relationship was an important concept to foster.

Evidence about comparison with a standard:
Grades are a clear and non-nuanced form of feedback which, by the psychological mechanisms outlined above, can trigger both promotion and prevention responses in trainees (57,62,68,77). This literature, largely in non-medical education, raised the question of whether comparison with a standard is useful in workplace assessment. Here, although grading was the norm and all too often grades were given with the comment boxes left blank (15), there was not a lot of evidence about the value or harms of grading or other forms of comparison with a required standard. This triggered the study in Chapter 7.

Evidence about what trainees do as a result of feedback:
Research into learners' experiences of feedback has highlighted the value placed on feedback that is actionable (78). Sargeant et al (79) found that doctors responded more positively and assimilated multi-source feedback better if they had facilitated reflection. We have assimilated this concept and are facing the challenge of how to make it work successfully.

The learning culture supporting or constraining the exchange of meaningful feedback:
Institutions can create opportunities for longitudinal teacher-learner relationships to flourish, such as extended placements (80–82). Bates et al. (83) in a small study across two medical schools explored medical students' perceptions of assessment and feedback in longitudinal integrated clerkships. They found that students were able to interpret even challenging or corrective feedback as supportive within such a relationship.

We have been encouraged by this evidence of the value of the feedback relationship and have invested heavily in the same. Supervision of a trainee can have built-in routines of supervisor observation of trainee performance followed by feedback (84) and expectations of recurrent feedback following multiple assessment tasks over time (75).
In the movement towards programmatic assessment, progress and learning from feedback is emphasized and built into the system (85,86). This has been successful in promoting informed self-assessment at McMaster and Maastricht medical schools (87,88). It has also encountered some difficulties in implementation, such as how to aggregate the data from multiple assessments into a holistic judgement (81), and the finding that when summative judgements are seen to be based on the formative assessments, the feedback given may be less critical (89). This

is a challenge we anticipate in a revised, more programmatic model of assessment, and has caused us to decide for the time being not to incorporate formative workplace assessment scores into judgements about progression.

1.4 Timeline of the action research

Three separate action research projects have contributed to our Keele curriculum development project – the first from 2007 to 2009, the second from 2010 to 2012 and the third from 2013 with cycles still ongoing.
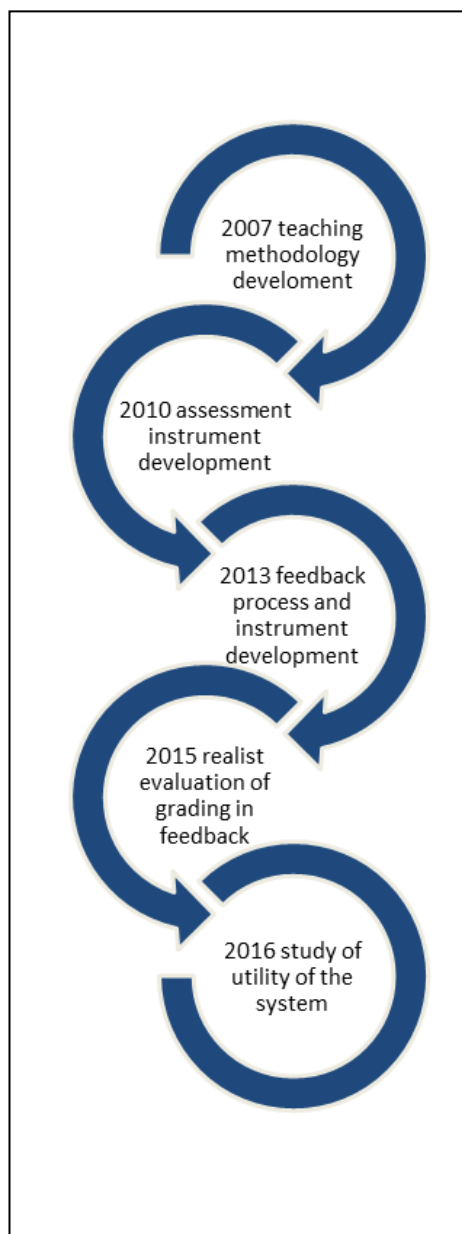
The first three publications report findings of general interest about classroom teaching arising from action research for development of the first year of the Keele consultation skills curriculum. These were:

A study of an innovation in classroom consultation skills teaching (Chapter 2); researching the gap between the classroom and real clinical practice (Chapter 3); and the benefits of integrated learning of end of life care skills (Chapter 4).

A competency framework for assessment of consultation skills "GeCoS" (Generic Consultation Skills) was then developed using a multi-centre modified Delphi study (Chapter 5).

Using a similar methodology, but this time including students in the process, we subsequently developed and validated tools to support clinical tutors to provide specific, behaviourally-orientated feedback (Chapter 6).

The implications of including comparative grading in formative workplace-based assessment were addressed by a realist evaluation of the effect on students of feedback with and without grades (Chapter 7).

Next, by secondary analysis of our existing data from the previous study we examined whether there was any value added to verbal feedback by the process of summarising it in writing (Chapter 8).

Having developed assessment and feedback tools, we needed to increase their usability. I supervised the development of an app to support assessment and feedback in the clinical workplace and evaluated its utility (Chapter 9).

2007 teaching methodology develoment

2010 assessment instrument development

2013 feedback process and instrument development

2015 realist evaluation of grading in feedback

2016 study of utility of the system

In parallel with the final action research project, I led a team of international medical educators in amalgamating the literature on formative feedback into a set of guidelines – the practical Do's, Don'ts and Don't Knows for individual clinical supervisors and for the institutions that support clinical learning (Chapter 10).

References

1. GMC. Tomorrow's Doctors [Internet]. GMC. 2003 [cited 2007 Feb 2]. Available from: http://www.gmc-uk.org/education/undergraduate/undergraduate_policy/tomorrows_doctors.asp

2. Swing SR. The ACGME outcome project: retrospective and prospective. Med Teach. 2007;29(7):648–54.

3. Frank JR, Jabbour M, Fréchette D, Marks M, Valk N, Bourgeois G. The CanMEDS 2005 Physician Competency Framework. Better standards. Better physicians. Better care. Framework. 2005. 1-40 p.

4. Cumming A, Ross M. The Tuning Project for Medicine--learning outcomes for undergraduate medical education in Europe. Med Teach. 2007;29(7):636–41.

5. Carraccio C, Wolfsthal SD, Englander R, Ferentz K, Martin C. Shifting paradigms: from Flexner to competencies. Acad Med. 2002 May;77(5):361–7.

6. GMC. Tomorrow 's Doctors The duties of a doctor registered with the General Medical Council. 2009.

7. Makoul G. Essential elements of communication in medical encounters: the Kalamazoo consensus statement. Acad Med. 2001;76(4):390–3.

8. Kneebone RL, Scott W, Darzi A, Horrocks M. Simulation and clinical practice: strengthening the relationship. Med Educ. 2004 Oct;38(10):1095–102.

9. Windish DM, Price EG, Clever SL, Magaziner JL, Thomas PA. Teaching medical students the important connection between communication and clinical reasoning. J Gen Intern Med. 2005;20(12):1108–13.

10. McKinley RK, Strand J, Ward L, Gray T, Alun-Jones T, Miller H. Checklists for assessment and certification of clinical procedural skills omit essential competencies: a systematic review. Med Educ. 2008 Apr;42(4):338–49.

11. Fraser R, McKinley R, H Mulholland. Consultation competence in general practice : Establishing the face validity of prioritized criteria in the Leicester assessment package. Br J Gen Pract. 1994;44:109–13.

12. Lefroy JE. How can the existing communication skills curriculum be optimised, in the light of experience here and elsewhere, and in congruence with the rest of the curriculum? Keele; 2010.

13. Kurtz S, Silverman J, Draper J. Teaching and Learning Communication Skills in Medicine Second Edition. Oxford: Radcliffe Publishing Ltd; 2005.

14. Kurtz SM, Silverman JD. The Calgary-Cambridge Referenced Observation Guides: an aid to defining the curriculum and organizing the teaching in communication training programmes. Med Educ. 1996 Mar;30(2):83–9.

15. Norcini J, Burch V. Workplace-based assessment as an educational tool: AMEE Guide No. 31. Med Teach. 2007;29(9):855–71.

16. Lefroy J, Watling C, Teunissen P, Brand P. Guidelines: the do's, don'ts and don't knows of feedback for clinical education. Perspect Med Educ. 2015;4(6):284–99.

17. Carr W, Kemmis S. Becoming Critical: Knowing Through Action Research. Lewes: Falmer; 1986.

18. McKearnan J. Curriculum Action Research. 2nd ed. London: Kogan Page; 1996.

19. Norton L. Action Research in Teaching & Learning. Oxford: Routledge; 2009.

20. Masters J. The History of Action Research [Internet]. Hughes I, editor. Action Research Electronic Reader, the University of Sydney, on-line. 1995 [cited 2018 Feb 26]. Available from: http://www.behs.cchs.usyd.edu.au/arow/Reader/masters.htm

21. McNiff J. Writing and doing action research. London: SAGE Publications Ltd; 2014.

22. Kember D. Action Learning, Action Research: Improving the Quality of Teaching and Learning. London: Kogan Page; 2000.

23. McKernan J. Curriculum action research; a handbook of methods and resources for the reflective practitioner. Second. London: Kogan Page; 1996.

24. McNiff J. Action Research: all you need to know. London: SAGE Publications Ltd; 2017.

25. Bradley P. The history of simulation in medical education and possible future directions. Med Educ. 2006 Mar;40(3):254–62.

26. Kneebone R, Kidd J, Nestel D, Asvall S, Paraskeva P, Darzi A. An innovative model for teaching and learning clinical procedures. Med Educ. 2002;36(7):628–34.

27. McGaghie WC, Siddall VJ, Mazmanian PE, Myers J, Committee AC of CPH and SP. Lessons for continuing medical education from simulation research in undergraduate and graduate medical education: effectiveness of continuing medical education: American College of Chest Physicians Evidence-Based Educational Guidelines. Chest. 2009 Mar;135(3 Suppl):62S–68S.

28. Mosley C, Dewhurst C, Molloy S, Shaw BN. What is the impact of structured resuscitation training on healthcare practitioners, their clients and the wider service? A BEME systematic review: BEME Guide No. 20. Med Teach. 2012 Jan;34(6):e349-85.

29. Ericsson K. Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. Acad Med. 2004;79(10 suppl):S70-81.

30. Cobb P, Bowers J. Cognitive and situated learning perspectives in theory and practice. Educ Res. 1999;28(2):4–15.

31. Schunk DH. Learning Theories: an educational perspective. Vol. 5th. NJ: Pearson Education

Inc; 2008.

32.   Taylor DCM, Hamdy H. Adult learning theories: Implications for learning and teaching in medical education: AMEE Guide No. 83. Med Teach. 2013;35(11):e1561-72.

33.   Vygotsky L. Thought and Language. MIT; 1986.

34.   Bandura A. Social Learning & Personality Development. Holt, Rinehart & Winston, INC: NJ; 1975.

35.   Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am Psychol. 2000;55(1):68.

36.   Lievens F, Coetsier P, De Fruyt F, De Maeseneer J. Medical students' personality characteristics and academic performance: A five-factor model perspective. Med Educ. 2002;36(11):1050–6.

37.   Hays RB, Jolly BC, Caldon LJM, McCrorie P, McAvoy P a, McManus IC, et al. Is insight important? measuring capacity to change performance. Med Educ. 2002 Oct;36(10):965–71.

38.   Bandura A, Locke E. Negative Self-Efficacy and Goal Effects Revisited. J Appl Psychol. 2003;88(1):87.

39.   Williams GC, Deci EL. The importance of supporting autonomy in medical education. Ann Intern Med. 1998;129(4):303–8.

40.   Dornan T, Littlewood S, Margolis SA, Scherpbier A, Spencer V. J& Y. How can experience in clinical and community settings contribute to early medical education? A BEME systematic review. Med Teach. 2006;28(1):3–18.

41.   Dornan T, Boshuizen H, King N, Scherpbier A. Experience-based learning: a model linking the processes and outcomes of medical students' workplace learning. Med Educ. Hope Hospital, School of Medicine, University of Manchester, Manchester, UK. tim.dornan@manchester.ac.uk; 2007 Jan;41(1):84–91.

42.   Dornan T, Hadfield J, Brown M, Boshuizen H, Scherpbier A. How can medical students learn in a self-directed way in the clinical environment? Design-based research. Med Educ. 2005;39(4):356–64.

43.   McKinley RK, Fraser RC, Van V D, Hastings AM. Formative assessment of the consultation performance of medical students in the setting of general practice using a modified version of the leicester assessment package. Med Educ. 2000;34(7):573–9.

44.   Hastings AM, McKinley RK, Fraser RC. Strengths and weaknesses in the consultation skills of senior medical students: Identification, enhancement and curricular change. Med Educ. 2006;40(5):437–43.

45.   McKinley RK, Dean P, Farooqi A. Reactions of volunteer general practitioners to educational assessment of their consultation performance: A qualitative study. Educ Prim

Care. 2003;14(3):293–301.

46. Fraser RC, Lee RSY, Yiu YK, Lam CLK, McKinley RK, Van der Vleuten CPM. Regulatory assessment of the consultation competence of Family Physicians in Hong Kong. Hong Kong Pract. 2004;26:5–15.

47. Norcini JJ, Blank LL, Duffy FD, Fortna GS. The Mini-CEX : A Method for Assessing Clinical Skills. Ann Intern Med. 2003;138(6):476–81.

48. Kogan J, Holmboe E, Hauer K. Tools for Direct Observation and Assessment A Systematic Review. JAMA. 2009;302(12):1316–26.

49. van de Ridder JMM, Stokking KM, McGaghie WC, ten Cate OTJ. What is feedback in clinical education? Med Educ. 2008;42(2):189–97.

50. Sargeant J, Mann K, van der Vleuten C, Metsemakers J. "Directed" self-assessment: practice and feedback within a social context. J Contin Educ Health Prof. 2008;28(1):47–54.

51. Bing-You RG, Paterson J. Feedback falling on deaf ears: Residents' receptivity to feedback tempered by sender credibility. Med Teach. 1997;19(1):40.

52. Sargeant J, Mann K, Ferrier S. Exploring family physicians' reactions to multisource feedback: perceptions of credibility and usefulness. Med Educ. 2005 May;39(5):497–504.

53. Sargeant J, Mann K, Sinclair D, Van Der Vleuten C, Metsemakers J. Challenges in multisource feedback: Intended and unintended outcomes. Med Educ. 2007;41:583–91.

54. Watling C, Driessen E, van der Vleuten CPM, Lingard L. Learning from clinical work: the roles of learning cues and credibility judgements. Med Educ. 2012 Feb;46(2):192–200.

55. Ivers N, Jamtvedt G, Flottorp S, Jm Y, Sd F, Ma OB, et al. Audit and feedback : effects on professional practice and healthcare outcomes. Cochrane Database of Systematic Reviews. John Wiley & Sons, Ltd; 2012.

56. Watling C, Kenyon C, Zibrowski E, Schulz V, Goldszmidt M, Singh I, et al. Rules of engagement: residents' perceptions of the in-training evaluation process. Acad Med. 2008;83(10 Suppl):S97-100.

57. Hattie J, Timperley H. The Power of Feedback. Rev Educ Res. 2007;77(1):81–112.

58. Walsh C, Ling S, Carnahan H. Concurrent versus terminal feedback: it may be better to wait. Acad Med. 2009;84(10 Suppl):S54-57.

59. Dohrenwend A. Serving up the feedback sandwich. Fam Pract Manag. 2002;9(10):43–6.

60. Pendleton D, Schofield T, Tate P, Havelock P. The new consultation Developing doctor-patient communication. Oxford University Press; 2003.

61. Cantillon P, Sargeant J. Giving feedback in clinical settings. BMJ. 2008;337:1292–4.

62. Kluger A, Van Dijk A. Feedback, the various tasks of the doctor, and the feedforward

alternative. Med Educ. 2010;44:1166.

63.   Eva KW, Armson H, Holmboe E, Lockyer J, Loney E, Mann K, et al. Factors influencing responsiveness to feedback: on the interplay between fear, confidence, and reasoning processes. Adv Health Sci Educ Theory Pract. 2012;17(1):15–26.

64.   van de Ridder JMM, McGaghie WC, Stokking KM, ten Cate OTJ. Variables that affect the process and outcome of feedback, relevant for medical training: a meta-review. Med Educ. 2015;49(7):658–73.

65.   Weston PSJ, Smith CA. The use of mini-CEX in UK foundation training six years following its introduction: Lessons still to be learned and the benefit of formal teaching regarding its utility. Med Teach. 2014;36(2):155–63.

66.   Minehart R, Rudolph J, Pian-Smith M, Raemar D. Improving Faculty Feedback to Resident Trainees during a Simulated Case. Anesthesiology. 2014;120(1):160–71.

67.   Sox CM, Dell M, Phillipi CA, Cabral HJ, Vargas G, Lewin LO. Feedback on Oral Presentations During Pediatric Clerkships: A Randomized Controlled Trial. Pediatrics. 2014;134(5):965–71.

68.   Kluger A, DeNisi A. The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. Psychol Bull. 1996;119(2):254.

69.   Teunissen PW, Bok HGJ. Believing is seeing: How people's beliefs influence goals, emotions and behaviour. Med Educ. 2013;47(11):1064–72.

70.   Higgins E. Promotion and prevention: regulatory focus as a motivational principle. In: Zanna MP, editor. Advances in Experimental Social Psychology. San Diego, CA: Academic Press; 1998. p. 1–46.

71.   Tuckey M, Brewer N, Williamson P. The influence of motives and goal orientation on feedback seeking. J Occup Organ Psychol. 2002;75(2):195–216.

72.   Morrison E, Cummings L. The impact of feedback diagnosticity and performance expectations on feedback-seeking behaviour. Hum Perform. 1992;5(4):251–264.

73.   Teunissen PW, Stapel DA, van der Vleuten CPM, Scherpbier AJJA, Boor K, Scheele F. Who Wants Feedback? An Investigation of the Variables Influencing Residents' Feedback-Seeking Behavior in Relation to Night Shifts. Acad Med. 2009;84:910.

74.   Molloy E, Boud D. Seeking a different angle on feedback in clinical education: the learner as seeker, judge and user of performance information. Med Educ. 2013 Mar;47(3):227–9.

75.   Boud D, Molloy E. Rethinking models of feedback for learning: the challenge of design. Assess Eval High Educ. 2013 Sep;38(6):698–712.

76.   Bok HGJ, Teunissen PW, Spruijt A, Fokkema JPI, van Beukelen P, Jaarsma D a DC, et al. Clarifying students' feedback-seeking behaviour in clinical clerkships. Med Educ. 2013

Mar;47(3):282–91.

77.   Higgins R, Hartley P, Skelton A. The Conscientious Consumer : reconsidering the role of assessment feedback in student learning. Stud High Educ. 2002;27(1):53–64.

78.   Watling C, Driessen E, van der Vleuten CPM, Lingard L. Learning culture and feedback: An international study of medical athletes and musicians. Med Educ. 2014;48:713–23.

79.   Sargeant JM, Mann K V., van der Vleuten CP, Metsemakers JF. Reflection: a link between receiving and using assessment feedback. Adv Heal Sci Educ. 2008;14(3):399–410.

80.   Veloski J, Boex JR, Grasberger J, Evans A, Wolfson DB. Systematic review of the literature on assessment, feedback and physicians' clinical performance: BEME Guide No. 7. Med Teach. 2006;28(2):117.

81.   Bok HG, Teunissen PW, Favier RP. Programmatic assessment of competency-based workplace learning: when theory meets practice. BMC Med Educ. BMC Medical Education; 2013;13(1):123.

82.   Hägg-Martinell A, Hult H, Henriksson P, Kiessling A. Students perceive healthcare as a valuable learning environment when accepted as a part of the workplace community. Educ Heal Chang Learn Pract. 2014;27(1):15–23.

83.   Bates J, Konkin J, Suddards C, Dobson S, Pratt D. Student perceptions of assessment and feedback in longitudinal integrated clerkships. Med Educ. 2013;47:362–74.

84.   Archer JC. State of the science in health professional education: effective feedback. Med Educ. 2010;44(1):101–8.

85.   van der Vleuten CPM, Schuwirth LWT. Assessing professional competence: from methods to programmes. Med Educ. 2005 Mar;39(3):309–17.

86.   Schuwirth LWT, Van der Vleuten CPM. Programmatic assessment: From assessment of learning to assessment for learning. Med Teach. 2011 Jan;33(6):478–85.

87.   Driessen EW, van Tartwijk J, Govaerts M, Teunissen P, van der Vleuten CPM. The use of programmatic assessment in the clinical workplace: a Maastricht case report. Med Teach. 2012 Jan;34(3):226–31.

88.   Li S, Sherbino J, Chan TM. McMaster Modular Assessment Program (McMAP) through the years: Residents' experience with an evolving feedback culture over a 3-year period. AEM Educ Train. 2016;5–14.

89.   Bok HGJ, Jaarsma D a. DC, Spruijt A, Van Beukelen P, Van Der Vleuten CPM, Teunissen PW. Feedback-giving behaviour in performance evaluations during clinical clerkships. Med Teach. 2016;38(1):88–95.

Chapter 2

Some like it hot: medical student views on choosing the emotional level of a simulation

Janet Lefroy[1]

Caragh Brosnan[2]

Sam Creavin[1]

[1]*Keele University School of Medicine, Keele, Staffordshire, Newcastle-U-Lyme, UK*

[2]*Centre for Biomedicine and Society, King's College London, London, UK*

Abstract

Objectives

This study aimed to determine the impact of giving junior medical students control over the level of emotion expressed by a simulated patient (SP) in a teaching session designed to prepare students to handle emotions when interviewing real patients on placements.

Methods

Year 1 medical students at Keele University School of Medicine were allowed to set the degree of emotion to be displayed by the SP in their first 'emotional interview'. This innovation was evaluated by mixed methods in two consecutive academic years as part of an action research project, along with other developments in a new communications skills curriculum. Questionnaires were completed after the first and second iteration by students, tutors and SPs. Sixteen students also participated in evaluative focus group discussions at the end of the Year 1.

Results

Most students found the 'emotion-setting switch' helpful, both when interviewing the SP and when observing. Student-interviewers were helped by the perception that they had control over the difficulty of the task. Student-observers found it helpful to see the different levels of emotion and to think about how they might empathise with patients. By contrast, some students found the 'control switch' unnecessary or even unhelpful. These students felt that challenge was good for them and preferred not to be given the option of reducing it.

Discussion

The emotional level control was a useful innovation for most students and may potentially be used in any first encounter with challenging simulation. We suggest that it addresses innate needs for competence and autonomy. The insights gained enable us to suggest ways of building the element of choice into such sessions. The disadvantages of choice highlighted by some students should be surmountable by tutor 'scaffolding' of the learning for both student-interviewers and student-observers.

Introduction

Self-determination in learning

Self-directedness is important for adult learning in a group setting and learners should be encouraged to have choice and control whenever possible (1). Adult learners also value self-esteem, and it is important that they do not fail dismally (2). Furthermore, students differ in their self-confidence, risk-taking, self-awareness (metacognition), mastery and performance goals and this affects how individuals learn and how much support and challenge each requires (3-5).

Teaching in medical schools which supports autonomy has been found to produce a more humanistic approach to the patient as well as to lower student anxiety, raise self-esteem, and enhance learning as evidenced by better test grades (6).

Learning should also build on existing knowledge and skills (7;8). Year 1 medical students differ in their pre-existing skills and in terms of how far they are able to extend themselves beyond their existing competence, a characteristic defined by Vygotsky as the 'width of their zone of proximal development' (9). One of the tutor's roles is therefore to help the learner progress to the next level of achievement by appropriate steps and not to expect the learner to leap an unbridgeable gap.

Learning of social skills is thought to be based on the modelling of others, as per 'social cognitive learning theory'. Learners are thought to gauge their capabilities not only through their own performance, but also by observing similar others perform, by persuasion from others ('You can do it!') and by their physiological indices (heart rate, sweating) (10).

Student control in simulation

Simulation is used in medical learning to allow students to experiment, to rehearse skills and to obtain immediate feedback in an environment that is safe for both student and patients (11-13). Two features of the 'simulated (standardised) patient' (SP) which enable experimentation and re-rehearsing are the 'pause' and 'rewind' functions. The pause function (which enables the interviewer to pause the interview in order to discuss with observers what is happening and then to proceed with the interview at a pace which suits him- or herself) and the rewind/replay facility (which allows the interviewer or a peer to try different approaches and see what happens) are both examples of context in which students are given control to help their learning (13). However, although students have typically been given control over the *pace* of interviews, there are no studies reporting sessions that give students control over the *content* of interviews, nor the effects of this extra layer of control on student learning.

For the past 2 years we have given junior medical students control over the level of emotion expressed by an SP in their first 'emotional interview'. Simulated patients are usually trained to set the emotional temperature as the script or facilitator instructs (13). The innovation here was that the student-interviewer had access to a 'control switch' to regulate the degree of emotion. This paper describes the impacts of offering such control to students on their learning, behaviour, and peer and self-perception. This is discussed within a theoretical framework which builds on the zone of proximal development and its implications for small-group teaching.

Institution-specific background : the communication skills programme and emotion-setting exercise

The communication skills programme within Keele University's new medical curriculum was designed by an action research group. One session was designed to prepare Year 1 medical students to handle patients' emotions by conducting interviews with emotional content about chronic illness with SPs. The design group included a medical student member (author SC), who proposed that the student-interviewer be allowed to set the level of emotion expressed by the SP. SC argued that an emotionally charged interview is a frightening experience for a Year 1 medical student, especially because he or she is performing in front of peers and does not want to fail. It was proposed that giving the students control over the emotional temperature would allow each student to face a challenge appropriate to his or her ability. The underlying hypothesis was that this, in turn, would enable students to learn more effectively, consistent with the theory of the 'zone of proximal development' (9).

In this session each group of eight or nine students discussed the skills involved in eliciting a patient's story and how they might handle the challenges they might face on their next placement, in which each student would be required to interview a patient with a chronic condition. Each group then engaged in three simulated interviews in which SPs and student-interviewers rotated. Students interviewed in pairs or alone (their choice), were observed by the rest of the group and were subsequently given feedback. The tutor invited interviewers to choose the level of initial emotion to be expressed by the SP (mild/ medium/ strong). The SPs had been trained to 'feel' and display this level of emotion when the topic of the chronic condition was broached, but also to respond to the student within the interview.

Methods

Study population recruitment

The study population comprised the first two cohorts of Year 1 undergraduate medical students in Keele University's new undergraduate medical curriculum (intakes of September 2007 and September 2008). At the end of their session on handling an emotional interview, all attending students were asked to complete a routine evaluation of their four-session introductory communication skills course and to give their optional informed consent for their responses to be used in research.

A subset of students were recruited to join focus group discussions to evaluate the first year of the new curriculum via an announcement made at a lecture and on the university's virtual learning environment. Students were asked to indicate (by returning an information slip or by e-mail) their interest in participating in a focus group to include evaluation and research questions or evaluation questions only. Groups were arranged so that as many students as possible could participate within their timetables (it was possible to allocate 26 out of 31 volunteers to groups), to ensure as even a gender split as possible, to separate problem-based learning (PBL) groups and to ensure adequate numbers in each group. Five focus groups were held. Two groups included students who wished to participate only in evaluation. The data discussed here come from the other three (research and evaluation) groups. Figure 1 shows how the sample of students related to the year cohorts.

Framework for questioning and analysis of responses

The routine evaluation questionnaire comprised 16 questions on the communication skills course and provided space for written comments (Appendix 1).

Focus group implementation

The focus groups took a modified grounded theory approach (14;15).

The moderator (author CB) and assistant were known to the students as the course evaluators and were not their tutors. The focus groups explored a range of issues relating to the new curriculum. As part of this broader evaluation, each group was asked to discuss their experiences of being allowed to set the emotional level of an SP interview.

Focus groups were audio-recorded and the material transcribed with the written consent of the participants. Thematic analysis of transcript data was performed using NVivo 2.0 [QSR International (UK) Limited, Southport, UK]. Tentative interpretations were developed at the time of data collection and the relevant literature was scanned to widen the interpretations.

Assumptions were discussed by the action research group in light of the findings; this discussion highlighted exceptions and sought explanations for apparent disagreements.
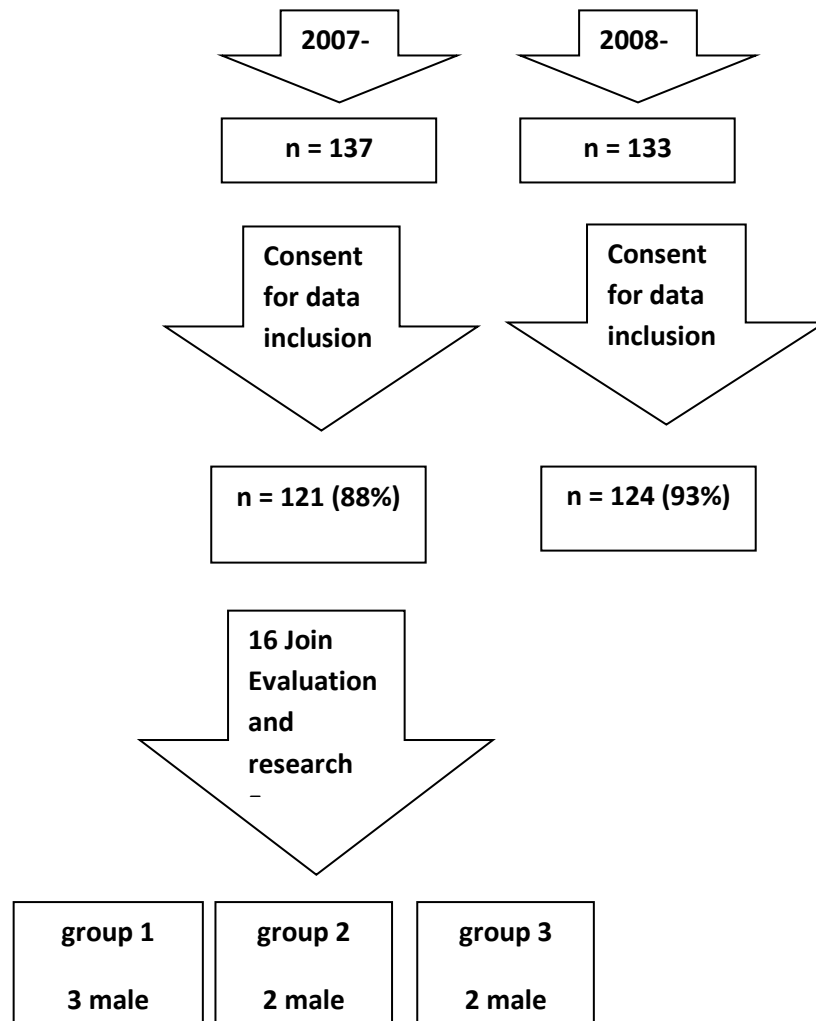
Fig 1: Flow chart to illustrate the evaluation cycle and how the sample of students in the study related to the year cohort.

Results

A total of 121 of the 137 students in the first cycle (88%) and 124 of the 133 students in the second cycle (93%) consented to the inclusion of their routine evaluation data in this study. Both year groups included a majority of female students and were similar in regard to proportions of graduates (Table 1).

Cohort Year 1 in:	Female	Male	Graduates	**Total study participants**	Total Year group
2007-8	69 (57%)	52 (43%)	14 (11.6%)	**121**	137
2008-9	75 (60.5%)	49 (39.5%)	18 (14.5%)	**124**	133

Table 1 : Frequency table of gender and graduate status of the student study participants

Quotations from questionnaire respondents are annotated with their responses on a Likert scale of Strongly agree-Agree-Disagree-Strongly disagree, with an identification number (e.g. T55 indicates a member of group T5 in 2007; 345 indicates a member of group 34 in 2008), and with the respondent's gender and graduate status (when positive).

At the end of the first year of the course, 16 students (seven male, nine female) took part in three focus groups. Two of the female students held previous degrees. Focus group members are identified by an initial, their gender and graduate status (when positive).

Students' evaluation of the 'emotion-setting switch'

The proportions of students choosing each level of emotion as recorded by group tutors were similar in 2007 and 2008, with around 55% of students requesting that strong motion be displayed (Fig 2).

Fig.2 The level of emotion requested in 87 interviews in 2007 and 2008

	mild	moderate	strong
2007	8	11	23
2008	7	13	25
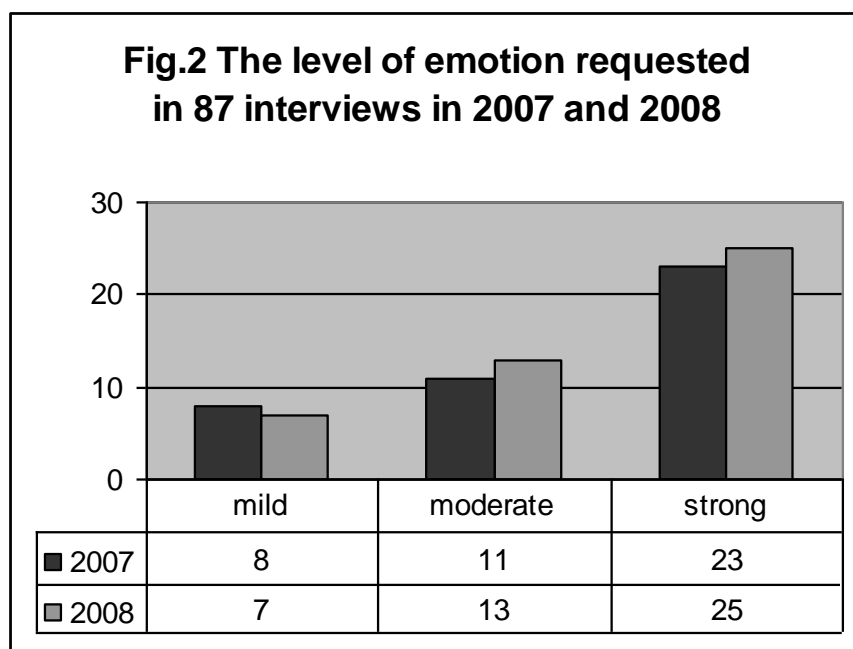
Of the 243 students who responded to the questionnaire item 'Being able to select the level of emotion in session 4 was very helpful', 209 (86%) agreed or strongly agreed with the statement (Fig. 3).

Students' written comments on this item were analysed in terms of expected and contrasting views. The expected view, supported by several student comments, was that giving control to the student enabled students to feel comfortable, prepared and self-confident and that too much challenge would be negative. However, a contrasting view also emerged from other student comments, indicating that some students believed challenge to be beneficial and that, as such, it would be better not to be too prepared.

Fig. 3 Students' responses to the statement 'Being able to select the level of emotion in session 4 was very helpful' in 2007 and 2008

	Strongly agree	Agree	Disagree	Strongly disagree
■ 2007 n=121	46	53	19	3
▨ 2008 n=122	48	62	12	0
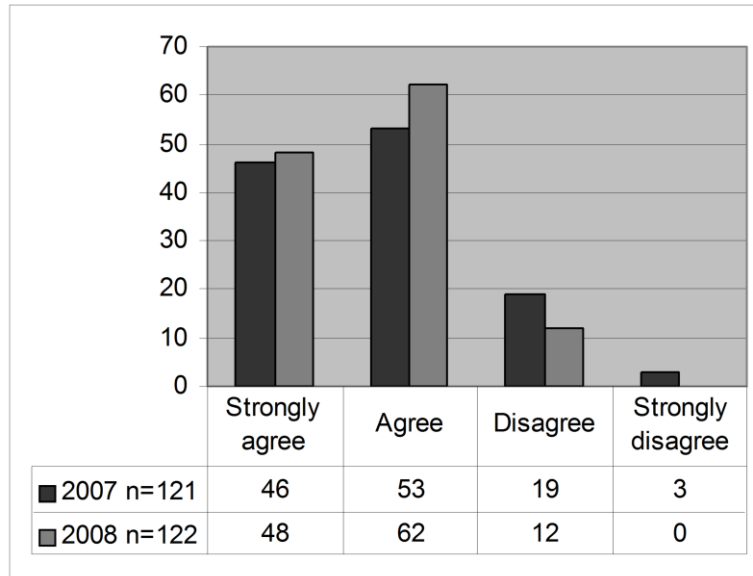
Other themes that emerged in both cycles referred to perceptions that seeing the range of emotions was educational and that the session topic itself was good preparation for placements. These themes are illustrated in Table 2 and developed below.

Table 2. Examples of comments expressing student agreement or disagreement with the statement: 'Being able to select the level of emotion in session 4 was very helpful'

[Agree] This meant that you were more confident and comfortable – thinking that you had decided on a level you can handle. (T55 Female)

[Agree] as we weren't just thrown in the deep end. (T57 Female)

[Strongly agree] Our group selected one of each level – this was very good for comparative purposes (345 male graduate)

[Agree] Although having a varied or unknown level is realistic, at least one strong emotional case is useful (336 male)

[Disagree] [The emotion switch] [w]ould only help those who were less confident. I felt most people wanted the challenge of high though.
(T13 Female)

[Disagree] …may have been more of a challenge to not know the emotion level.
(F15 Male)

Students' experiences of being in control

The focus groups shed further light on students' experiences of having control. Contrasting themes emerged. Firstly, having control enabled students to feel comfortable and to adjust the exercise to their learning needs. However, group dynamics also shaped the way control was used and perceived.

Control as a route to comfortable learning for individuals

Students liked being able to control both the pace of the interview and the level of emotion. This student referred to the standard SP facilities of 'pause' and 'rewind':

'I think the "pause" and the "rewind" kind of commands were really useful, because you could stop and talk to the group and things like that and that helped a lot rather than carrying on to fail and then talking about how badly you failed. It gave you a chance to correct what you were doing if you were making a mistake.' (R, male, group 1)

The next student took the discussion onto the additional emotional level control:

'I'm glad we could choose the levels because I was scared, and I thought if they just came in and they chose themselves that they were going to be really angry or upset, I wouldn't know how to deal with that, but if we just chose it would be a mild one, you knew how to deal with that better.' (S, female, group 1)

Here the 'control switch' clearly enabled the student to bring the task into her range of capability, rather than being disabled by the degree of the challenge. Other students in this focus group agreed that the 'control switch' made learning more comfortable.

Students also learned from observing the interviews of others. The emotion-setting exercise added value through observation of the handling of different levels of emotion.

Student comments on both the questionnaire and in focus groups about subsequent placement interviews indicated that students transferred this learning to skills used in placement interviews with real patients.

The influence of the group on the interpretation of control

A second major theme identified in the focus group data was the idea that challenge is good (and therefore comfort is bad). The fact that this was expressed more often in the focus group discussions (two of the three groups) than in the questionnaire comments may indicate that emotional memories had faded somewhat by the time the focus groups were conducted, but may also reflect the tendency of groups to find consensus, which, in this case, veered towards a group perception of themselves as challenge-seeking. Alternatively, it may be that the students who volunteered for the focus groups were among the more confident or extrovert of the cohort. Nevertheless, what these discussions illustrated was that in selecting their emotional level, students may have been influenced by the presence of their peers as well as by their individual learning needs.

Focus group discussion revealed that challenge-seeking was regarded by some as a male attribute:

'All the guys went for the top ones [levels].' (D, male, group 2)

We do not have data to triangulate with this view on a gender difference, but it is worth noting that the male-female difference in the rating of the helpfulness of the 'control switch' was not statistically significant. However, the comment indicates that there may have been an element of

bravado in students' choice of level, as is evident in the same student's reflection on the risk that the exercise might come to be perceived as more about controlling the SP than about learning to communicate:

'Instead of focusing on our communication skills, we would be joking around afterwards about how we made a SP cry or did you see the SP, focusing on the SP rather than what we were actually doing – that was the only danger I could see from doing it' (D, male, group 2)

Elsewhere in the focus group discussions, students referred to competition within their PBL groups. It appears that competition may also have shaped students' perceptions of the emotion-setting exercise. This student felt that to choose for comfort was cowardly, and that it was good to be challenged by a difficult interview:

'I don't think you should be able to pick to be honest. I don't think you should be allowed to wimp out of it, which I think a lot of people took the easy route and said "Give me a little bit". You don't learn anything from that. You need to go from normal to quite extreme otherwise you are not getting the full benefit of the session.' (W, female, graduate, group 2)

An element of peer pressure to choose the most difficult level was implied by the comments of more than one student:
'It is surprising how many people did for go for the higher setting. No-one really chickened out and went for the lower one.' (A, male, group 3)

It took prompting from the moderator ('what did you think, F?') to get another student to admit that she had indeed been daunted by the challenge. This student used the same judgemental term ('chicken') to describe herself while defending her need to choose the mild emotional level:

'I found it quite hard, I didn't really know how to deal with patients. I chickened out and went for the low one, but I think it is good to introduce it to us. We will be dealing with patients who are upset and angry - any range of emotions.' (F, female, group 3)

The choosing of the emotional level was therefore interpreted by some students within a normative framework in which a decision to face stronger emotion was more highly valued.

Discussion

Our study suggests that giving students choice over the level of emotion expressed by an SP in their first emotional interview is helpful for two main reasons. Firstly, student comments demonstrate the intended effect of the 'control switch' in line with the constructivist principles of enabling learning in the zone of proximal development (10). Students were able to build their new learning as far out from their existing level of competence as they wished, making use of the teaching at the beginning of the session, as well as the help offered by the group if they paused the interview because they were struggling. Secondly, because the 'switch' enabled students to tailor the learning experience to their abilities, some chose to be challenged beyond the level which we, as tutors, would have set the emotional temperature. This broadened the learning experience for the whole group and equipped them better for the range of experiences they

encountered on placements. This fits with social constructivist views such as those of Vygotsky (9), which stress that social group learning is useful. As students model for and observe one another, they not only teach skills, but experience higher self-efficacy for learning (8;9;16). This benefit to observers provides the group with an opportunity for interviewers to legitimately choose a range of levels between them.

Giving learners choice in a group setting is not always straightforward and some disadvantages to the provision of the 'emotion-setting switch' emerged. By contrast with the intended effect, choice also emphasises differences between students which can reduce self-esteem and create peer pressure. This indicates the complexity of the processes by which students make decisions about the challenges they choose to face.

Many medical students are competitive (4). This may make it difficult for some to choose a comfortable level of challenge. Both the challenge-seeking and the challenge-averse may choose a level for optimal learning that is inappropriate to their zone of proximal development.

An explanation (other than peer pressure) for our students' antipathy to choice is that some students observing an interview in which the interviewer chose a mild emotional level may have wished the emotional temperature to have been higher for their own learning as observers. They may have rationalised this by thinking it would be 'good' for their interviewing colleague(s). The benefit of seeing interviews with different emotional levels and how to handle them was certainly perceived as an added bonus of the 'emotion-setting switch'.

Strengths and weaknesses

The strengths of this study include its mixed-methods approach, which involved the administration of a class-wide questionnaire and the conducting of focus groups that enabled the in-depth exploration of concepts and the further refinement of group members' thoughts (17). Its weaknesses include the fact that the focus group evaluation was required to be part of the broader curriculum evaluation and may not have achieved saturation of ideas on all themes. Focus groups can overemphasise consensus (18). There was also a delay of six months between the teaching session and the focus groups.

Conclusions

Students' choices are driven by a complex web of peer pressure, challenge seeking and fear of failure and the tutor's task in such sessions is demanding. Constructivist learning theories state that in order to work within their zone of proximal development, individual learners may need the content of learning to be 'scaffolded' by teachers (10;19). Our tutor notes to this purpose are shown in Table 3.

Autonomy-supportive teaching is to be commended in medical education because of its many positive outcomes for both students and their patients. The findings of this study probably apply to medical educators who are interested in giving students choice in any group setting. The next stage in our exploration of emotional level control is to re-evaluate its use with explicit scaffolding by tutors to determine whether the negative impacts of competitiveness and peer pressure can be reduced.

Table 3 Suggestions for the 'scaffolding' of simulation in a group setting
The tutor can 'scaffold' the learning of both interviewers and observers by:

Orientating participants respectively to the task for the interviewer and the learning opportunities for the observers

Enabling self-assessment of their capabilities, an understanding of the "zone of proximal development" and sensible choice by interviewers.

Encouraging a supportive group response to differing educational needs to enable choice and train better medical teachers of the future.

Diverting any focus on peripheral aspects of the situation such as the acting prowess of the simulated patient. |

Reference List

(1) Knowles MS. The adult learner - a neglected species. 4th ed. Gulf Publishing Company; 1990.

(2) Kernis MH.   Measuring Self-Esteem in Context: The Importance of Stability of Self-Esteem in Psychological Functioning. *Journal of Personality* 2005;**73**(6):1569.

(3) Handbook of Competence and Motivation. New York: Guilford Press; 2005.

(4) Knights JA, Kennedy BJ. Medical school selection: Screening for dysfunctional tendencies. *Med Educ* 2006;**40**(11):1058-64.

(5) Lievens F, Coetsier P, De Fruyt F, De Maeseneer J. Medical students' personality characteristics and academic performance: a five-factor model perspective. *Med Educ* 2002;**36**(11):1050.

(6) Williams GC, Deci EL. The importance of supporting autonomy in medical education. *Ann Intern Med* 1998 Aug 15;**129**(4):303-8.

(7) Cobb P, Bowers J. Cognitive and situated learning perspectives in theory and practice. *Educational Researcher* 1999;**28**(2):4-15.

(8) Schunk DH. Learning Theories: an educational perspective. 5th ed. NJ: Pearson Education Inc.; 2008.

(9) Vygotsky L. Thought and Language. MIT; 1986.

(10) Bandura A. Social Learning & Personality Development. Holt, Rinehart & Winston, INC: NJ; 1975.

(11) Aspegren K. BEME Guide No. 2: Teaching and learning communication skills in medicine - a review with quality grading of articles. *Med Teach* 1999;**21**(6):563-70.

(12) Kneebone R. Simulation in surgical training: educational issues and practical implications. *Med Educ* 2003 Mar;**37**(3):267-77.

(13) Kurtz S, Silverman J, Draper J. Teaching and Learning Communication Skills in Medicine Second Edition. Oxford: Radcliffe Publishing Ltd; 2005.

(14) Strauss A, Corbin J. Grounded Theory Methodology; an overview. In Denzin N.K and Lincoln Y.S. Handbook of Qualitative Research. London: Sage; 1994.

(15) Pope C, Ziebland S, Mays N. Analysing qualitative data. In: Mays N, Pope C, editors. Qualitative research in health care. London: BMJ Books; 1996.

(16) Bandura A. Self-efficacy: the exercise of control. NY: Freeman; 1997.

(17) Merton R, Fiske M, Kendal P. The focused interview. Glencoe Illinois: The Free Press; 1956.

(18) Barbour RS. Making sense of focus groups. *Med Educ* 2005 Jul;**39**(7):742-50.

(19) Wood D, Bruner J, Ross G. The role of tutoring in problem-solving. *Journal of Child Psychology and Psychiatry* 1976.

Appendix 1. Student Evaluation of communication skills introductory course

Section A: communication skills programme evaluation				
Please tick the answer that best matches your opinion	**Strongly agree**	**Agree**	**Disagree**	**Strongly disagree**
The tutor(s) helped me to understand and improve communication skills	☐	☐	☐	☐
Our group gave me useful, constructive feedback	☐	☐	☐	☐
The simulated patients were realistic in their roles	☐	☐	☐	☐
The simulated patients gave useful, constructive feedback	☐	☐	☐	☐
I can now use descriptive feedback	☐	☐	☐	☐
I feel confident in my ability to initiate a patient interview	☐	☐	☐	☐
I feel confident in my ability to explore a patient's problems	☐	☐	☐	☐
The communication skills classes prepared me well for placements	☐	☐	☐	☐
The communication skills classes and placements integrated well with our PBL	☐	☐	☐	☐
Please comment overleaf, especially if you have disagreed with any of these statements				
Section B: evaluation of specific features of the communication skills programme				
Session 1 on active listening and feedback skills was necessary before we started group practice with SPs	☐	☐	☐	☐
Comments:				
I would prefer to interview the SP myself rather than with another student	☐	☐	☐	☐
Comments:				
Tutor continuity is very important in this group activity	☐	☐	☐	☐
Comments:				
Being able to select the level of emotion in session 4 was very helpful	☐	☐	☐	☐
Comments:				
I would like to have been videoed and to see my interview	☐	☐	☐	☐
Comments:				
What do you think of the agenda-led balanced feedback process we are using in the group sessions with SPs?				

Any alterations you might suggest for

- the learning activities?

 -the briefings in the handbook?

Any other comments:

This information will be used for routine programme evaluation.
Can we also use your answers and your OSSE communication skills data in anonymous form for a publication?

Yes ☐ **Signed …………………………………**

Chapter 3

Minding the gap between communication skills simulation and authentic experience

Sarah Yardley[1]

Alison Irvine[2]

Janet Lefroy[2]

[1]Research Institute for Primary Care and Health Sciences, Keele University, Keele, UK

[2]Clinical Education Centre, Keele University School of Medicine, University Hospital of North Staffordshire NHS Trust, Stoke-on-Trent, UK

Abstract

Context   Concurrent exposure to simulated and authentic experiences during undergraduate medical education is increasing. The impact of gaps or differences between contemporaneous experiences has not been adequately considered. We address two questions. How do new undergraduate medical students understand contemporaneous interactions with simulated and authentic patients? How and why do student perceptions of differences between simulated and authentic patient interactions shape their learning?

Methods   We conducted an interpretative thematic secondary analysis of research data comprising individual interviews (n=23), focus groups (three groups, n=16), and discussion groups (four groups, n=26) with participants drawn from two different year cohorts of Year 1 medical students. These methods generated data from 48 different participants, of whom 17 provided longitudinal data. In addition, data from routinely collected written evaluations of three whole Year 1 cohorts (response rates ≥ 88%, n=378) were incorporated into our secondary analysis dataset. The primary studies and our secondary analysis were conducted in a single UK medical school with an integrated curriculum.

 Results   Our analysis identified that students generate knowledge and meaning from their simulated and authentic experiences relative to each other and that the resultant learning differs in quality according to meaning created by comparing and contrasting contemporaneous experiences. Three themes were identified that clarify how and why the contrasting of differences is an important process for learning outcomes. These are preparedness, responsibility for safety, and perceptions of a gap between theory and practice.

Conclusions   We propose a conceptual framework generated by reframing common metaphors that refer to the concept of the gap to develop educational strategies that might maximise useful learning from perceived differences. Educators need to 'mind' gaps in collaboration with students if synergistic learning is to be constructed from contemporaneous exposure to simulated and authentic patient interactions. The strategies need to be tested in practice by teachers and learners for utility. Further research is needed to understand gaps in other contexts.

Introduction

Contemporaneous experience of simulated and authentic patient-student interactions occurs in medical curricula across the world and is required by the General Medical Council (1) in the UK. We define 'simulated patients' (SPs) as lay people who are trained to act as patients in medical interviews and give feedback from a patient perspective. Commonly these people are participating in role-plays based in teaching environments remote from clinical practice. 'Authentic early experience' denotes human contact that occurs in clinical or social workplaces for the purpose of learning (2).

Developing communication skills through SP and authentic patient interactions

The advantages of simulated interactions include a reduction in the risk for harm (including psychological distress) to patients or students, the ability to control 'patient' supply and demand, partial control of the content of interactions, and the provision of opportunities for students to 'practise' different scenarios and responses in order to develop appropriate knowledge, skills and behaviours (3-6), with accompanying patient-referenced feedback. By contrast, authentic early experiences are a form of workplace-based learning, and are intended to encourage students to contextualise the curriculum and ease the transition into clinical learning during later years (7-10). Differences are seen by students when good practice ideals promoted in the classroom are not replicated by the healthcare professionals they see consulting in the workplace (11,12) and when authentic patients respond differently from SPs. As with other forms of workplace-based (also called 'experience-based') learning, students require support to maximise the learning potential of these variable yet authentic experiences (11,13,14).

The importance of physical and psychological fidelity during simulation is debated in the literature (4,15), largely in relation to simulator equipment; less attention has been paid to that within SP encounters. It is helpful in this debate to take the social character of simulation into account by acknowledging that participants and organisers enter into a 'fiction contract', which allows them to treat the simulation as if it were real in order to practise transferable skills (16). How critical physical or psychological fidelity are to supporting the fiction contract depends on whether the desired educational goals are to learn psychomotor dexterity, procedural knowledge, decision making, interpersonal skills or team working norms and values, or a combination of these. The simulation of communication skills is often conceptualised as representing preparation for the workplace, 'bridging the gap' between the classroom and clinical practice (17,18). This is based on the premise that authentic experience sequential to simulated experience is safer for both students and patients (4,17-20) and that simulation offers an opportunity to instil understanding of ideal practice (21) prior to experiences of pragmatism in authentic workplaces (22).

Few studies directly compare the two educational settings (23), consequently how new medical students handle contemporaneous experiences of interactions with simulated and real patients has not been adequately considered (24). Students exposed sequentially to SPs in the early years of training and then to real patients in the later years report that real patients are more focused on students' understanding of medical content than on their ability to communicate (25). Students can be suspicious that SPs have been told to withhold information by faculty staff, but may still describe simulated interactions as useful preparation for real encounters or for practising skills in worst case scenarios (25).

There is also evidence that learners commonly struggle to transfer knowledge between contexts (26-28). Transfer will be impaired if there is a perceived gap between what is taught in medical school and the reality of medicine as it is practised in the workplace (29). This creates potential for dissonance between student experiences of simulation and authentic practice (4,30). Unless there is understanding of how and why students conceptualise their experiences, and particularly of how they handle these differences, we cannot seek to improve patient care through integrated simulation-based and authentic workplace-based education (4,31,32).

Methods

This is an exploratory study to clarify student perceptions of contemporaneous interactions with simulated and authentic patients so that we can better understand the consequences for the development of communication skills. In this paper we present an interpretative thematic secondary analysis of data before developing a conceptual framework for educational strategies in order to make sense of, and learn from, gaps or differences.

Methodological framework

Qualitative secondary analysis is attracting increasing interest in social science research disciplines (for examples, see (33-35)) and amongst research funders. Although there is no universally agreed definition of secondary analysis, this term is used to describe situations in which the researchers conduct further analysis of one or more datasets for purposes not defined or predicted in the original study design (see Heaton (36) and Thorne (37) for an overview of different types of secondary analysis). The attractions of secondary analysis as a methodology include: (i) the facilitation of data analysis across datasets (e.g. when each individual dataset provides relevant and complementary data with which to explore a particular question), and (ii) the further exploration of unrelated novel questions or unexpected findings generated as by-products of studies with different foci. It has also been suggested that the secondary analysis of data from different sources may improve the generalisability of qualitative findings (38). Many of the potential criticisms of secondary analysis (aside from issues that apply to any qualitative approach, whether primary or secondary), such as loss of contextual information, are negated or resolvable when researchers from the original studies are involved in a rigorous process of secondary analysis. We have re-examined data originally generated from students in Years 1 and 2 during research studies conducted by each of the authors (Table 1). All three of the original studies, albeit that they addressed different research questions, were situated within an interpretative constructionist paradigm and so shared commonalities in theoretical perspective.

Objectives

The studies from which our dataset originates were all designed to look at aspects of student interactions with either simulated or authentic patients (Table 1). During the primary analysis of one dataset (11), an unexpected finding was observed by the first author of this paper: not only did students compare learning in different settings, but this comparison led students to make value judgements about what was valid knowledge. Students were generating knowledge and meaning from their simulated and authentic experiences relative to each other. Although one might reason that expanded learning could emerge from students' comparisons of simulated and authentic experiences, with each offering complementary aspects of learning, it is also possible that learning from either sort of experience might be reduced as students contrasted experiences when making value judgements. The impact of comparison and contrast on learning from concurrent simulated and authentic experiences has not previously been studied in detail. The other two authors of the present paper had also separately identified a similar need to better understand the impact of students' comparing and contrasting of simulation with clinical experience in their own masters' studies (39,40) and observations during teaching (41). Secondary analyses of qualitative data look at the data through a different 'lens' and with fresh research questions. Our objective, therefore, was to explore and clarify effects on the contemporaneous provision of both types of experience through two research questions. (i) How do new

undergraduate medical students understand contemporaneous interactions with simulated and authentic patients? (In the study context, 'new' refers to students in the first 2 years of medical school.) (ii) How and why do student perceptions of differences between simulated and authentic patient interactions shape their learning? The second of these questions emphasises our objective of developing understanding of learning outcomes or consequences arising from exposure to difference.

Setting

All three of the studies from which we drew our dataset were conducted in a single UK medical school with an integrated curriculum for undergraduates. The curriculum uses a hybrid model incorporating problem-based learning, experiential learning within the medical school, laboratory sessions, lectures and authentic early experience placements.

Students interact with simulated and authentic patients from the start of their studies. Authentic patients are encountered predominantly in clinical placements, although patients are also used in classroom teaching. In their first term, students have four classroom-based, tutor-facilitated communication skills teaching sessions. The first explains and explores the use of role-play in teaching and the principles of feedback. The following three sessions use SPs. The first clinical placement occurs between the third and fourth sessions, and is supported by a student briefing at the end of the third session and debriefing at the start of the fourth session, which, respectively, prepare students for and enable them to reflect upon their first authentic patient experiences. Simulated interactions in the early years involved no simulation of the environment, but only of the 'patient' role. The general stated purpose of classroom sessions with SPs was to offer students practice prior to their interactions with authentic patients; each session has its own specific written learning objectives. During authentic early experience placements students were supervised (but not directly observed) by nominated professionals within workplaces. Usually (among other activities), the supervisor would set up an encounter with a patient, whom the students would then interview in pairs.

All three original studies (11,39,41) were subject to independent peer review and prospective ethical approval was gained from Keele University School of Medicine Research Ethics Committee for the elements of work in each study that exceeded normal procedures for evaluation of the curriculum (for which ethical approval is not currently required in this setting). All participants gave informed consent for the data contained in this paper to be used in research. The methodological framework of our secondary analysis was also peer reviewed.

The dataset

We conducted secondary analysis on data generated from research methods and routine evaluations. The complete dataset comprised research data from individual interviews (n=23), focus groups (three groups, total participants n=16), and discussion groups (four groups, total participants n=26) taken from two sequential year cohorts (entering in the 2007/2008 and 2008/2009 academic years) of undergraduate medical students. The interviews had been audio-recorded with undergraduate students in Years 1 and 2 participating in a study of meaning making and knowledge construction from authentic early experience (11). Students from these year cohorts later (in Years 2 and 3 of their training) participated in audio-recorded discussion groups, which were transcribed verbatim (11). The focus groups were also audio-recorded and transcribed verbatim in a study of student experiences with SPs (39). Overall, these methods

generated data from 48 different participants, of whom 17 provided longitudinal data as a result of their sequential participation (a feature of one of the original studies (11) meant 14 students participated in sequential interviews and discussion groups; three of the students who participated in this study also participated in one of the others (39)). We have ensured that no individual student's views are over-represented in our secondary analysis dataset by cross-checking transcriptions. In addition, routine written evaluation data obtained from three Year 1 cohorts of students (n=378, as described in Table 1) were incorporated into the secondary analysis dataset. Given that response rates for all forms of evaluation data were ≥ 88%, we would expect to find that students who participated in other forms of data generation were also represented in the routine evaluation data, but because of the anonymising of the evaluation data, we cannot confirm this. Details of the conduct of each original study from which the datasets were drawn are outlined in Table 1, in which we have summarised the theoretical framework, original research questions, setting, recruitment, sampling and participation, and methods of each study. We have sequentially recorded verbatim quotations (rather than retaining the different original dataset classification systems) for the purposes of this paper in order to support readability. Selected quotations were drawn from different participants.

Table 1	Summary of primary studies from which data sets were drawn		
Study	**Yardley (2011)**	**Lefroy (2010)**	**Irvine and Lefroy (2011)**
Theoretical framework	Orientated towards the principles of constructionism, interactionism and interpretivism. Socio-cultural theories were used to interrogate empirical data and empirical data was used to refine and develop these theories within the field of Medical Education	Action research to design and evaluate the curriculum during its first year of delivery drawing on adult learning theories within a constructivist approach to learning	Interpretative analysis of evaluation data (free text)
Original research questions	'How and why do students construct useful knowledge and meaning making from authentic early experience?' and 'How and why do students make authentic early experiences work for them?'	In the communication skills curriculum what curricular content should be taught, by what methods, in what locations, by whom, to achieve which desired learning outcomes with respect to the first year of the curriculum?	The survey asked for 6 open text responses: Comments on learning activities Unit 1: Emergencies: What did you enjoy the most? What aspects of this unit have you found most challenging? Unit 2: Infections and Immunity: What did you enjoy the most? List the 2 best things about Semester 1. List the 2 things that could be improved about Semester 1 and suggestions for how they could be improved.
Setting	UK medical school established in 2002 and implementing a new	The same UK medical school as Yardley (2011) in the 2 years	As described for Yardley (2011). The cohort of Year 1

	locally designed curriculum from 2007/2008. Students participated in simulated patient interactions (during classroom communication skills training) and in real patient interactions (during authentic early experience placements in workplaces) from the start of their first year. Students were provided with a paper briefing prior to each session in the classroom or workplace outlining intended communication skills learning outcomes.	surrounding the launch of its new curriculum. The action research team included stakeholders such as communication skills tutors, clinical tutors from workplace settings, a fourth year medical student representing the student perspective and simulated patients to represent informed lay perspectives.	students are divided into 15 tutor-led small groups for each communication skills teaching session. Five groups run at a time and each session is followed by a tutor debriefing involving the 5 tutors reporting back to the lead tutor who compiles a written lead tutor report.
Study	**Yardley (2011)**	**Lefroy (2010)**	**Irvine and Lefroy (2011)**
Recruitment, sampling and participation	Students were recruited from academic years commencing in 2007/2008 and 2008/2009. Participation was voluntary following recruitment via e-mail and lecture announcements. Students were sampled from both Years 1 and 2 of the undergraduate degree as the authentic early experience programme spanned both years. 2007/2008 cohort: n = 4 (individual interviews alone), n = 8 (individual interviews and discussion group), n = 3 (discussion group alone) 2008/2009 cohort n = 5 (individual interviews alone), n = 6 (individual interviews and discussion group), n = 9 (discussion group alone).	The study population comprised Year 1 students whose undergraduate degree commenced in 2007/2008 and 2008/2009, plus the tutors and simulated patients working with these student cohorts. A subset of students in the 2007/2008 cohort was also recruited to join focus group discussions (3 groups, n = 16 in total). This subset volunteered following announcements in a lecture and online.	All 133 Year 1 students in the cohort commencing in 2010/2011 were sent an electronic evaluation using 'SurveyMonkey TM' software at the end of semester 1. 100% responded. In addition all students were invited to give verbal feedback on their first placement experience to their small group tutor in the fourth communication skills session, at the end of Semester 1. The feedback from 15 tutors, taken over the 3 sets, after this session was summarised into 3 lead tutor reports (students' views were hence subject to reinterpretation).
Methods	Interviews and discussions groups (when results from prior interviews were shared with students divided by year and previous participation) were used to generate data between January 2009 and March 2010. Students had been on between 2 and 4 placements in their current year of study, and the second year students had completed up to 6 placements in their first year of study. Participants were asked to provide examples of their experiences and encouraged to explain their own interpretations of these during semi-structured interviews (52). Interviews included discussion of the expectations, processes and consequences of authentic early	Routine evaluation data 2007/ 2008 cohort: 121/137 Year 1 students (88% of cohort). 57% of responders were female, 11.6% graduates and 9.9% repeating the year. 2008/2009 cohort: 124/133 Year 1 students (93%), 60 (45% of respondents) were female, 14.5% graduates and 5.6% repeating the year. Analysis of routine evaluation data for communication skills programme collected through a student questionnaire containing 16 questions and space for free text. For further details including the evaluation forms see Lefroy, Brosnan and Creavin(3). Students self-identified comparisons between communication skill sessions and	Students' open text comments to the 6 survey questions were transcribed verbatim. Thematic analysis identified 72 comments from 54 students of relevance to simulated patient teaching and authentic patient interactions. Thematic analysis of the elements of the 3 lead tutor reports relevant to comparison and contrast between simulated and authentic patient encounters.

	experiences. The discussion groups (53–55) were designed to allow student participants to comment on developing findings and to enhance understanding of the student perspective through discussion of views amongst peers to identify areas of consensus or difference, and collective meaning-making (56). Data were audio-recorded and independently transcribed. Three complementary overarching data types were generated; phenomenological themes, narrative (content and structure/language) and presented meaning. Analysis was conducted in discussion with 3 other researchers using mixed qualitative methods that drew on strategies from thematic analysis, narrative and discourse analysis and interpretative phenomenological analysis. Further details of the original study including the full study design and methodology are available in Ref. 57.	their authentic early experiences during the process of evaluation. 16 students took part in 3 focus groups at the end of the first cycle of Year 1 (2007/2008 cohort) – 7 male and 9 female of whom 2 were graduates. Allocation to focus groups was by availability and in order to disperse PBL group members as much as possible. Efforts were also made to obtain an even gender balance within groups. Student focus groups were analysed using a modified grounded theory approach. Data was transcribed and coded by the author into themes. The moderator and assistant were known to the students as the course evaluators and were not their tutors. The focus groups explored a range of issues relating to the new curriculum, and as part of this broader evaluation each group was asked to discuss their experiences of communication skills sessions and placements. Focus groups were audio-recorded and transcribed with written consent from participants. Thematic analysis of transcript data was performed using NVivo 2.0 software. Tentative interpretations were developed at the time of data collection and the relevant literature was scanned to widen the interpretation. Assumptions were discussed by the action research group in light of findings, highlighting exceptions and seeking explanations for apparent disagreement. For further details see Lefroy (39).	
Original data re-used in secondary analysis	Interview and discussion group verbatim transcripts.	Free text from questionnaires. Transcripts from focus groups.	Verbatim transcriptions from survey (tutor reports provided contextual information).
Unique participants contributing research data to secondary analysis data set	35 Unique participants contributed through either an individual interview or discussion groups.	13 Unique (in addition to Yardley 2011) participants contributed through focus groups.	Not applicable.

Routine evaluation data included in secondary analysis data set	Not applicable	n=245	n=133

Analysis

The data from all three sources were combined before an interpretative thematic analysis addressing the research questions outlined in this paper was conducted. All text was read and coded for type of experience, subject matter, comparison or contrast narratives, and comments on similarities and differences between simulated and authentic patients. Similarities and differences between each dataset were sought. Themes in the data were identified through discussion of these codes by all three authors. Data extracts are presented in the Results section to illustrate specific points within the analysis. Attention was paid to the social construction of the data and the language used. Our interpretation was developed through a rethinking of existing metaphors of 'gap'. This reframing produced an alternative conceptual model for using difference and contrast to potentiate learning and enabled the development of our proposed educational strategies (see Discussion).

Results

Both forms of learning were well received amongst the student body as evidenced by levels of satisfaction reported in contemporary written routine evaluations. In this section, we present three key cross-cutting themes derived from secondary analysis of data in which students compare or contrast communication differences between simulated and authentic patient interactions:

1. Preparedness for being a student on placement or for becoming a doctor;
2. Responsibility for the safety of the patient and student, and
3. Student perceptions of a gap between theory and practice.

The initial analysis is presented and our interpretation is further developed.

Preparedness for being a student on placement or for becoming a doctor

On evaluating their introductory communication skills course, 118 (99%) Year 1 students in November 2007 and 121 (98%) in November 2008 agreed that communication skills classes prepared them well for placements (39). However, in the interviews and discussion groups conducted a few months later by SY (11), although students reported that their expectations of simulation had been met, they also argued that it would not have been possible to fully prepare them for their experiences with authentic patients:

> '…although we were adequately prepared for placements, I didn't feel that prepared because I hadn't actually gone out and spoken to patients yet because ... what I mean is the actual development of getting better at talking to patients is by talking to more patients and, so I think I really needed to develop the confidence, really… get out in the real world before I felt adequately prepared for placements.'(S1) (11)

Students participating in focus groups (39) also reported satisfaction with the realism of the SPs, but questioned whether learning arising from these sessions could really be directly considered as 'preparedness':

> 'I thought it [simulated sessions with student choice of level of patient emotion] was really useful, but a couple of weeks after that I had a placement where a patient did actually start crying. Even though it was useful and I knew more what to expect, you still feel completely overwhelmed when you are sat there in a room with two of you. Maybe because it was a male patient I felt that there was a guy sat there crying and you're there like what do we do, what have we done?....... because you know with the simulated patient that you can't offend them if you upset them, it's not actually…it doesn't prepare you that much for real emotion, you are still completely overwhelmed by it.' (S2)(39)

Both the students quoted imply that simulation is useful in acquiring skills but it is less useful for preparing the learner for how he or she is going to feel when faced with reality. For such students, the fiction contract (16) that is in place during simulation does not extend to consideration of their own feelings. Instead, students in simulation focus on personal performance, or the reactions of their peers and tutors.

Students particularly value the educational role of the SP

> '... what is more helpful with the simulated patients is the feedback that they give you afterwards, because they've obviously done it plenty of times before – they know what they're looking for, they know what... they know what a good history is all about, so they can give constructive feedback which is invaluable really – simulated patients are really invaluable in that respect.' (S3)(11)

However, students in discussion groups framed interactions with SPs as more awkward or antagonistic than with authentic patients, as they were felt to require prescribed student behaviours to 'unlock' phases of the patient script:

> 'Yes the simulated patients like, it's like they've been primed, they've only been told that they can say certain things if you ask a question in the correct way. If you don't say it in the correct way, they don't give you that bit of information that you need to then ask your next question whereas a normal patient you can just ask them one question and they can go on forever and you can pick up loads of points to then ask them.' (S4) (11)

The finding that students felt they were participating in a script during simulated sessions is not unique to this setting, as illustrated by general debate, within the field of medical education, surrounding the hidden curriculum in multiple spheres of learning, and about fidelity issues within simulations. It does, however, suggest that students might need more support to engage in the 'fiction contract' (16). Nonetheless, the 'artificial' aspects of simulated interactions did provide students with learning opportunities that would otherwise perhaps not have occurred. For example**,** the option to 'pause' and seek advice mid-interaction promoted student learning:

> 'I think the pause and the rewind... commands were really useful, because you could stop and talk to the group and things like that and that helped a lot rather than carrying on to fail and then talking about how badly you failed. It gave you a chance to correct what you are doing if you were making a mistake.' (S5) (39)

The use of 'gospel' as a metaphor by the following student could suggest a perception that the medical school, unlike the student, believes there is a single correct way to communicate. This is supported by the use of 'right' and 'wrong' when describing feedback:

> '…if they just gave us communication skills and left it at that, it would just be learning a set of theories or a set of questions... you can't take this rigid structure as gospel anyway, it's meant to be a framework which you work from because not every patient's gonna be the same... But it's... invaluable to have the grounding first... with... simulated patients...with a tutor there to guide you where you're going wrong and to tell you when you're going right... then actually going out and doing it.' (S6)(11)

Tutors, SPs and students are, in fact, instructed to facilitate feedback in terms of what worked and offering alternatives rather than judgements of what is right and wrong, although we do not know if these instructions were always followed during the study sessions. Taken to a logical conclusion, these findings suggest that students could be feeling pressure to behave in one way in classrooms and another in workplaces. Their comments indicated that many students conceived the purpose of simulation sessions as being limited to the short term goal of coping with authentic experiences as students. By contrast, at least some of the students viewed authentic experiences as preparation for future practice. For these students the impact was considerable:

> 'Placements - all three were very different, memorable experiences that encouraged me through giving me a vision of what I could be doing in five years' time. They helped me understand the patient experience and communicate with patients' (S7) (41)

With respect to learning content and practically applicable knowledge for the future, the unpredictable agendas of authentic patients were reported as providing valuable opportunities to learn and derive meaning. Students were able to identify potential learning beyond the faculty-designed objectives when interacting with authentic patients, such as in understanding the patient's life:

> 'They might come out with... a lot of things which you don't expect or which you never asked but somehow it came out... they came out with something totally unrelated but still a good insight to their lives.' (S8) (11)

Responsibility for safety: patient and student

Placements can be disappointing, especially if providers seem unprepared or unwelcoming, and the expected educational opportunities do not materialise. Some students in the 2010-2011 cohort reported, for example, that '…the provider didn't even know we were coming!' (S9) (41), despite there being clear administrative processes to book and confirm placements well in advance of student visits. Some providers seemed unclear about students' intended educational objectives:

> 'When I went to placements, I felt I was abandoned sometimes and I didn't know what to do apart from interviewing patients. Someone should be beside me while I was interviewing the patient and I should be given feedback at the end. Therefore, I could learn from mistakes and I could improve my communication and interviewing skills' (S10) (41)

This comment also suggested simulation might create student dependence upon a level of supervision that is not always available in clinical practice. The student body had taken to heart concerns of some faculty about risk and the potential harm that may arise from authentic interactions. The faculty intention as expressed in briefings was to reassure students that they should not be pressured into acting above their competencies. For some students, at least, this resulted in anxieties that limited them:

> 'He (a workplace supervisor) just said out of the blue, 'would you like to take a history off the patient?' and I just thought right, well, I'd rather not do it terribly and, you know, potentially make the patient worse off because of it – why put her through a history that's not going to be properly taken…' (S6) (11)

> 'I'd say with reference to the communication skills, being able to get the practice in with simulated patients before was definitely beneficial rather than just getting straight out and interviewing a patient because the potential for mistakes is quite high'. (S11) (39)

Some students believed that authentic patients might not detect underperformance (because they expect competency); this created a sense of responsibility by contrast with the 'safe' experience afforded by interacting with SPs, which created a sense of performance. The following example shows how a student's self-confidence is affected by the performance she perceives the patient to expect:

> '…you know you can do it and you know that the patient's not going to know if you've done it wrong… when it's a normal patient… well they expect me to know what I'm doing, so… it's easier to have the confidence because there isn't somebody there to scrutinise you' (S12)(11)

Other students were more cautious, voicing concerns about upsetting patients and crossing the expected norms of lay interactions, which might produce unpredictable reactions from patients:

> 'You can't harm simulated patients… you can't really make them upset… whereas a real patient… they perceive us as doctors.' (S13) (11)

> 'there's a lot more to think about when you're with a real patient…you really are delving into their personal, private lives... whereas the simulated patients are told to react in a certain way, these patients could act any which way they want to… and you have to...go…a bit more cautious' (S14)(11)

For some students these unknowns are exciting and challenging; for others they are unsettling. A combination of simulation and authenticity was sometimes created by inviting authentic patients into classroom settings. These sessions were valued by the students and appeared to be viewed as less risky:

> 'They were really useful. The fact that they were very willing to talk about their experiences and were willing if you asked them anything. Their answer would be fantastic. You didn't feel worried to ask them a question because of the environment we were in, it just felt very open and easy to talk to them' (S15) (39)

Student perceptions of a gap between theory and practice

Some comments reveal a substantial gap from student perspectives:

'skills acquired in EL [experiential learning] are impossible to be applied on placement. EL and placement are completely different situations (S16) (41)

'they don't do it the way you teach us to' (S17) (41)

'whereas a real patient obviously isn't [primed by the medical school]... – so it just feels more like a real conversation... whereas I think with an SP obviously you're doing things to try and tick off the right things... what you learn would be quite different. On simulated patients you are basically practising what you have been taught during that session... – what you should do with consent and so on... It's quite rigid.' (S18)(41)

The clearest example of a student-perceived gap between theory and practice related to the discussion of consent and confidentiality in the two types of interaction (41). In particular, the perception that the medical school was mistaken about the importance of consent and confidentiality was common amongst students. This was because although these aspects had been identified as important in the classroom, students had not seen placement providers explicitly talk about these issues at the start of every patient encounter in practice (41). Some concluded that SPs were following the medical school's rules rather than representing a valid patient perspective:

'I think simulated patients try to do things a lot more by the book, whereas real patients... they aren't as, you know, sort of straightforward as you might think – you wouldn't normally go through, confidentiality with them and then consent and that sort of stuff, 'cause they just... they don't see it as being important, whereas simulated patients will – that's only probably because they've been told to... by the medical school.'(S19) (11)

Despite the differing requirements for consent within clinical and primarily educational encounters (the latter type predominantly refer to early patient interviews involving novice learners), none of the student interviewees described considering such nuances. Very few students appeared to realise that often practitioners had continuing professional understandings with their patients, or that some patients might, in certain circumstances, see consent or confidentiality as of vital importance. A student may spontaneously draw the conclusion that real patients do not see confidentiality and consent as important, rather than considering alternative explanations such as, for example, that real patients believe the observation of good practice in these areas to be a given and, therefore, not to require discussion.

Discussion

To interpret the meanings of the three themes identified, we have developed a conceptual framework suggesting alternative meanings for metaphors which refer to the notion of the 'gap' that teachers, supervising clinicians and learners might find useful in developing educational strategies for making sense of, and learning from, gaps or differences. The three themes identified in our secondary analysis can be conceptualised as contributing to an overarching theory-practice gap between simulated and authentic patient interactions. Our key finding is not that SPs are

perceived differently from authentic patients (we suggest that this will be self-evident), but is a clarification of how students actively use their perceptions of difference to compare and contrast and so construct learning from their contemporaneous experiences. Our analysis identifies that students generate knowledge and meaning from their simulated and authentic experiences relative to each other, and that similarities or differences seen in the workplace reinforce or negate classroom learning in complex ways. When difference was identified during interactions with patients, students made meaning about what was 'real' in the workplace and what was important to the medical school faculty (identified through SPs and tutors who were perceived as agents of the medical school). Students found it difficult to suspend the sense of giving a performance in the classroom. Authenticity produced a contrasting sense of responsibility towards patients, whereas many students remained reluctant to be assertive about their learning needs. In authentic situations, students believed patients might not detect underperformance because they would expect competency. This meant that some students were actually more at ease during real patient experiences, but the associated responsibility caused others some discomfort.

We have interpreted student talk of the exemplar differences and resultant meaning making illustrated in the present data as representative of 'gaps' that require recognition and explanation if we are to maximise the learning opportunities to be derived from contemporaneous exposure to simulated and authentic early experiences. The 'gap' arises commonly in metaphors in both everyday language and the fields of medicine and medical education. For example, we may talk about the gap between theory and practice, that between expectations and achievement, or that between the teaching that is delivered and the learning that is generated. The notion of the 'gap' is present in communication skills literature (17,18) and in clinical and teaching practice. Use of the concept of a physical 'gap' is often associated with solutions to remove the gap, or eliminate its effects, illustrated through the common use of phrases such as 'bridging the gap' or 'closing the gap' in everyday life. These term suggest that gaps are conceptualised as sources of disconnection or risk, rather than as metaphorical spaces for development. This is by contrast with the work of Vygotsky, who conceptualised learning and meaning as social and cultural rather than individual processes (42). He describes a metaphorical gap or space (the zone of proximal development) to define the additional potential a learner has to understand understanding, through interaction with other agents and structures, beyond what might be achieved alone (42). To understand and explain gaps requires a critical approach to the purpose of metaphor, and consideration of whether different meanings could underlie the metaphor. Our interpretation reframes the meaning of metaphors of the gap to develop educational strategies for teachers and learners.

Figure 1 How educators can be 'gap-conscious' in the teaching of medical undergraduate communication skills

Simulated experience:

- Student centred
- Medical school teaching environment
- Levelled to the learner
- Structured, planned, supported, predictable, often highly integrated with curriculum
- Consistent delivery of teaching experiences
- Trained educators with protected time
- Simplified/standardised cases
- Patient illness is generally absent or simulated
- Supervised practice
- Safe place for mistakes
- Performance-related feedback
- Driven by learning objectives
- Seen as pretend/rehearsal and requires the 'fiction contract'

Authentic experience:

- Patient centred
- Workplace clinical environment
- Patient's conditions unpredictable, so not necessarily integrated with curriculum
- Different learning experiences
- Teachers may not be formally trained in medical education
- Complexity – of illness, team interactions, patients' needs
- Unwell, anxious patients
- Service commitments can compete with teaching
- Harm avoiding and risk averse culture
- Student performance may be unobserved without feedback
- Driven by patients' clinical needs
- Seen as 'real'

The educators' role

Don't ignore the gap

- Recognise that an educational conversation and a clinical encounter are different – be explicit about the value of a 'fiction contract'
- Prepare students intellectually and emotionally for an important transition
- Simulation *per se* is not the bridge, but make it as realistic as possible with contextualised scenarios, credible patients and clinically-oriented teachers
- Don't widen the gap by promoting unachievable idealism in classroom and assessment rather than what is attainable in clinical practice

Manage the gap

- Careful planning of placements and good cross-boundary communication
- Briefings for placement, students and tutors to define expectations
- Contextual orientation in placements
- Debriefings in placement and classroom
- Enable students to capitalise on the unplanned and unpredictable through educator-supported critical analysis

Be mindful of the gap

- Use differences as a substrate for learning
- Use reflection, peer discussion & tutor- facilitation to make sense of differences
- Bring complexity & uncertainty into the classroom
- Be wary of the potential dangers of simulation
- Be conscious of the impact of positive and negative role models and encourage debate of what is 'good' communication in different situations

Educational opportunities in the theory-practice gap

Both simulated and authentic patient-student interactions are social practices: contextual events which occur in space and time in which people interact with one another, artefacts and the environment for learning purposes (16). We have already drawn on the work of Dieckmann *et al.* (16) by building on their use of the term 'fiction contract' to describe how participants who suspend disbelief and conduct simulated interactions as if they are authentic may benefit more in terms of educational value. In addition, we suggest that the educational value of both simulated and authentic interactions may be synergistically increased through explicit attention to, and discussion of, difference. To date, few studies have directly compared the two educational settings. Our findings demonstrate that students continually make comparisons for themselves, and that the spontaneous meanings of difference which students construct can lead to a process of 'competitive contrast', in which the student rejects learning constructed from simulation that appears to conflict with the practice he or she observes in authentic workplaces. Exposure to both modes of teaching could be better used to expand overall learning by actively encouraging students to critically appraise their simulated and authentic experiences in comparison to each other, asking why difference occurs and seeking to assimilate and accommodate the resulting understanding into their evolving conceptual frameworks of good clinical practice.

Moving from 'competitive contrast' to 'constructive comparison' of difference

Theoretical and empirical evidence in other areas of medicine has previously shown that reasoning and meaning-making often involve the use of comparison and contrast (4, 23-25, 43, 44). In spontaneous processes of meaning-making, difference is more striking than similarity (45, 46). Figure 1 summarises our evolving conceptual framework of the two teaching environments and the physical, intellectual and emotional gaps between them. We propose that these gaps, or at least the 'solutions' to them, need to be reconceptualised to maximise the educational value of students concurrent engagement in, respectively, simulated and authentic patient interactions. Metaphors can be helpful in conceptualisation, but can also lead to assumptions of common understanding rather than to discussion of what different people perceive the work of the metaphor to be. Rather than seeking to 'close' or even to 'bridge' the gaps, we suggest, based on our findings, that educators – within medical schools and workplaces alike – in collaboration with their students, need to 'mind' these gaps, or to acknowledge them by thinking differently and critically about them.

In order to move students' learning from 'competitive contrast' of ideals with the pragmatic and nuanced realities of workplace learning (as the present data shows this reasoning to result in the rejection of these ideals), we need to develop educational strategies which allow students to make 'constructive comparisons', and to generate learning from differences. This finding is not dissimilar to that observed in general practice clerkships by van der Zwet *et al*. (47) who describe how developmental space is needed to learn and develop a professional identity. Space is created when context and interactions with others allow students opportunities to 'mind their learning' with educators' support (47).

Practical implications for educators

We suggest that educators need to be mindful of gaps between student experience of, respectively, simulated and authentic patient interactions. The educator has a role to play in driving a continual cycling of constructive comparison (indicated by the arrows in Figure 1 and the

panel describing the educator's role). We suggest the following strategies for putting mindedness into practice:

1. <u>Don't ignore a gap</u> as this risks paradoxical meaning-making, the rejection of ideals in face of contrast in reality, creation of dichotomies and misunderstandings. For example, Kneebone (48), who has written extensively and thoughtfully about the use of simulation in surgery, used the ha-ha wall as a metaphor for understanding the different perspectives of novices and experts in order to illustrate the dangers of ignoring a gap. A ha-ha is a hidden ditch in the grounds of a large country house (see Figures 2 and 3). Our photographs show the vertical ha-ha wall, and a seemingly seamless transition from the lawn of the house (expert's perspective) to the livestock in the park beyond, but an almost insurmountable boundary to climb from the park (novice's perspective).

Figures 2 and 3
The ha-ha view – a metaphor for the different perspectives of experts and novices

| **Figure 2. Novice view** | **Figure 3. Expert view** |
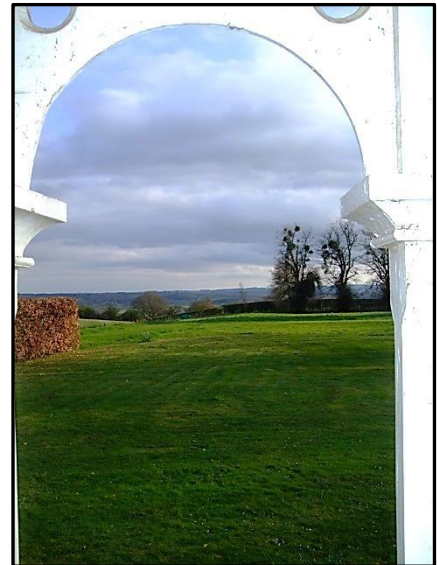
2. <u>Manage the gap</u>: educators who recognise and understand gaps can work collaboratively with students to discuss perceived differences and make constructive comparisons. This requires explicit expectation of difference, making the educator's role one of facilitating the student's making of meaning, which includes encouraging students to theorise about how and why identified differences occur. It also requires the educator to acknowledge that placements may require a level of adaptability and self-directedness over and above that which students may have needed in the classroom, and to provide the necessary support for students during the process and debriefing elements of their interactions, without relying solely on preparedness.

3. <u>Use it - being 'mindful' of the gap</u>: use of Epstein's (49) term is intentional. Mindfulness can be considered an element of students' reflective practice that leads to personal and professional development (50). Tutors also need to be mindful, however, of how they portray the other side of the gap, and of their potent effect as role-models (both positive and negative). Regardless of the quality, breadth and depth of the 'communication skill

toolkits' offered to students in classrooms**,** simulation cannot achieve the same potency as exposure to the daily professional practice of qualified clinicians.

Strengths and limitations

Qualitative research studies usually produce data that exceed the researchers' original purpose and that generate interesting findings beyond the specific research questions for which the study was designed. Seeking to interpret the data rather than simply to confirm expected findings represents a marker of robust and rigorous qualitative analysis. It is therefore important that any unexpected findings that are identified are given due consideration. Secondary analysis provides a mechanism for this. The differences noted by students impact on their learning in either setting and we found that students actively construct meanings to explain these differences. Our research questions focused not on whether students perceive difference, but, rather, on how students perceive differences and what effects these perceptions have on their learning. The congruence and replication of findings within our data can be considered as representative of a form of triangulation through the process of secondary analysis. However, there are also potential limitations to our work. A secondary analysis (or meta-analysis or systematic review for that matter) will depend, at least in part, on the quality of the original studies, although, as we returned to the original data, our analysis was not dependent on any pre-existing interpretation. Data from all three original studies (11,39,41) were derived from the same UK medical school. It is possible that elements of the findings represent the circumstances of that particular school. Yardley (11) and Lefroy (39) sampled the same cohort in 2007/2008 (the first cohort to undertake a new curriculum in the medical school) and therefore the views of this cohort may not be representative of those of other years once the curriculum had become embedded. We also do not know what impact was effected by the fact that two of the authors (AWI and JL) taught on the classroom communication skills programme at the time of the research. We are, however, reassured by the congruence of their findings with those of Yardley (11) who was then known to students only as an education researcher. It is possible to construct the sequential participation of some students within the data as either a strength or weakness. Some of the original studies deliberately generated data longitudinally (11,39). Within the secondary analysis, it could be argued, for example, that this overlap is a strength as the consistency of views across studies suggests students were not simply trying to please particular researchers or meet particular study expectations. As with any secondary analysis, we cannot know if our participants might have offered different perspectives or different explanations for their handling of perceived gaps had such questions been directly put to them in a primary qualitative study. This area requires further exploration.

Conclusions

'Minding the gap' is an interpretative metaphor that we offer on the basis of our analysis. We do so to suggest that students will construct meaning in the gaps they perceive between classroom and authentic practice out of an intrinsic human desire to reconcile or explain lived experiences. Metaphor is defined as understanding one conceptual domain (the target domain) in terms of another conceptual domain (the source domain), which leads to the identification of a conceptual metaphor (51). The metaphor itself may not be spoken out loud but may be apparent (e.g. in our data, the phrase 'they don't do it the way you teach us' (S17) clearly illustrates the presence of a

conceptual gap even if this is not explicitly named as such) or interpreted in interactions between people, such as teachers and learners. It is important to pay attention not just to what is said or not said, but to how and why it is said in order to more fully understand the meaning of the utterance for the speaker.

This research shows that learning context is significant, but also that different contexts can be positively contrasted by students to potentiate learner-created meaning. We have generated a conceptual framework that challenges people to think critically about the use of gap metaphors and what they personally mean when invoking the notion of the gap as a metaphorical tool. We hope our suggested educational strategies will be of practical use for teachers and supervisors engaging in simulated and authentic patient experiences with students by providing them with insight into students' perceptions and reasoning. In the medical school at which this work was conducted, Year 1 students are now explicitly briefed to think about the 'gap' between simulated and authentic patient interactions and are given guidance on how the recognition of differences may be an opportunity to extend their learning.

It is important to recognise that interactions in both simulated and authentic contexts can be subject to complex interpretations by students. We should neither reject simulation as lacking in reality nor be seduced into expecting it to resolve all the challenges of developing effective communication skills in practice. Instead, we should seek to find ways of minding the gap to increase the learning potential of concurrent simulation and authentic experiences.

In order to clarify how gaps between theory and practice influence learning, and whether more specific discussion of differences is beneficial, further research is required. We hope that the concept of 'minding the gap' might be considered more widely and in other contexts in order to further exploration of whether this concept has potential to encourage the development of transferable learning. Further studies might also usefully consider how student expectations of contemporaneous interactions with simulated and authentic patients are formed and whether interventions might target this process to further potentiate the development of communication skills. The conceptual framework and educational strategies we suggest need to be tested in practice by teachers and learners for utility. The outcomes and impact of using our conceptual framework and educational strategies for teaching and learning should be evaluated through further research. In addition, research to understand gaps between theory and practice in other contexts might also usefully contribute to understanding of the importance of differences for students in shaping their learning.

Contributors: SY conceived the idea for this paper based on her observations of difference, contrast and comparison in data generated as part of her doctoral thesis on authentic early experience. Her work led to the development of theory using the concept of a 'gap' that might be used to potentiate educational value. SY wrote the first draft of the paper, contributed to the analysis and integration of data from all three of the earlier studies for this purpose and finalised the submitted version. AWI developed strategies for 'minding the gap' from themes within the background reading for her master's dissertation. She analysed data, supplied references, and contributed to the drafting, redrafting and proof-reading of this manuscript. JL contributed to the

collation and analysis of research data, the formulation of the conclusions, and the drafting and revision of this paper. All authors approved the final manuscript for publication.

References

1 General Medical Council. Tomorrow's Doctors: outcomes and standards for undergraduate medical education. London: GMC 2009.

2 Littlewood S, Ypinazar V, Margolis SA, Scherpbier AJJA, Spencer J, Dornan T. Early practical experience and the social responsiveness of clinical education: systematic review. BMJ 2005;331:387–91.

3 Lefroy J, Brosnan C, Creavin S. Some like it hot: medical student views on choosing the emotional level of a simulation. Med Educ 2011;45:354–61.

4 Kneebone RL, Scott W, Darzi A, Horrocks M. Simulation and clinical practice: strengthening the relationship. Med Educ 2004;38:1095–102.

5 Bradley P, Postlethwaite K. Simulation in clinical learning. Med Educ 2003;37 (Suppl 1):1–5.

6 McGaghie W. C., Siddall V.J., Mazmanian P.E., Myers J., American College of Chest Physicians Health and Science Policy Committee. Lessons for continuing medical education from simulation research in undergraduate and graduate medical education: effectiveness of continuing medical education: American College of Chest Physicians evidence-based educational guidelines. Chest. 2009;135 (3 Suppl): 62–8.

7 Dornan T, Littlewood S, Margolis SA, Scherpbier AJJA, Spencer J, Ypinazar V. How can experience in clinical and community settings contribute to early medical education? A BEME systematic review. Med Teach 2006;28:3–18.

8 Hopayian K, Howe A, Dagley V. A survey of UK medical schools' arrangements for early patient contact. Med Teach 2007;29:806–13.

9 Yardley S, Littlewood S, Margolis SA, Scherpbier AJJA, Spencer J, Ypinazar V, Dornan T. What has changed in the evidence for early experience? Update of a BEME systematic review. Med Teach 2010;32:740–6.

10 Dornan T, Bundy C. What can experience add to early medical education? Consensus survey. BMJ 2004;329:834–9.

11 Yardley S. Understanding authentic early experience in undergraduate medical education. PhD dissertation. Keele: Keele University 2011.

12 Yardley S, Brosnan C, Richardson J. The consequences of authentic early experience for medical students: creation of student m_etis. Med Educ 2013;47:109–19.

13 Dornan T, Boshuizen H, King N, Scherpbier A. Experience-based learning: a model linking the processes and outcomes of medical students' workplace learning. Med Educ 2007;41:84–91.

14 Yardley S, Brosnan C, Richardson J, Hays R. Authentic early experience in Medical Education: a socio-cultural analysis identifying important variables in learning interactions within workplaces. Adv Health Sci Educ 2012. doi: 10.3109/0142159X.2012.73304510.1007/s10459-012-9428-2.

15 Demaria S Jr, Bryson EO, Mooney TJ, Silverstein JH, Reich DL, Bodian C, Levine AI. Adding emotional stressors to training in simulated cardiopulmonary arrest enhances participant performance. Med Educ 2010;44:1006–15.

16 Dieckmann P, Gaba D, Rall M. Deepening the theoretical foundations of patient simulation as social practice. Simul Healthc 2007;2:183–93.

17 Bligh J, Bleakley A. Distributing menus to hungry learners: can learning by simulation become simulation of learning? Med Teach 2006;28:606–13.

18 Malhotra A, Gregory I, Darvill E, Goble E, Pryce-Roberts A, Lundberg K, Konradsen S, Hafstad H. Mind the gap: learners' perspectives on what they learn in communication compared to how they and others behave in the real world. Patient Educ Couns 2009;76:385–90.

19 du Boulay C, Medway C. The clinical skills resource: a review of current practice. Med Educ 1999;33:185–91.

20 Gaba DM. The future vision of simulation in health care. Qual Saf Health Care 2004;13 (Suppl 1):2–10.

21 Hanna M, Fins JJ. Power and communication: why simulation training ought to be complemented by experiential and humanist learning. Acad Med 2006;81 (3):265–70.

22 Smith B. From simulation to reality – breaking down the barriers. Clin Teach 2006;3:112–7.

23 Bokken L, Rethans JJ, J€obsis Q, Duvivier R, Scherpbier AJJA, van der Vleuten CPM. Instructiveness of real patients and simulated patients in undergraduate medical education: a randomised experiment. Acad Med 2010;85:148–54.

24 Bokken L, Rethans JJ, Scherpbier AJJA, van der Vleuten CPM. Strengths and weaknesses of simulated and real patients in the teaching of skills to medical students: a review. Simul Healthc 2008;3:161–9.

25 Bokken L, Rethans JJ, van Heurn L, Duvivier R, Scherpbier AJJA, van der Vleuten CPM. Students' views on the use of real patients and simulated patients in undergraduate medical education. Acad Med 2009;84:958–63.

26 Norman G. Teaching basic science to optimize transfer. Med Teach 2009;31:807–11.

27 Dornan T, Hadfield J, Brown M, Boshuizen H, Scherpbier A. How can medical students learn in a self-directed way in the clinical environment? Design-based research. Med Educ 2005;39:356–64.

28 Brown J. Transferring clinical communication skills from the classroom to the clinical environment: perceptions of a group of medical students in the United Kingdom. Acad Med 2010;85:1052–9.

29 Eva KW. What every teacher needs to know about clinical reasoning. Med Educ 2004;39:98–106.

30 Okuda Y, Bryson EO, DeMaria S Jr, Jacobson L, Quinones J, Shen B, Levine AI. The utility of simulation in medical education: what is the evidence? Mt Sinai J Med 2009;76 (4):330–43.

31 Ellaway RH, Kneebone R, LaChapelle K, Topps D. Practica continua: connecting and combining simulation modalities for integrated teaching, learning and assessment. Med Teach 2009;31:725–31.

32 Kneebone RL, Kidd J, Nestel D, Barnet A, Lo B, King R, Yang GZ, Brown R Blurring the boundaries: scenario-based simulation in a clinical setting. Med Educ 2005;39:580–7.

33 Bishop L. Ethical sharing and reuse of qualitative data. Aust J Soc Issues 2009;44:255–72.

34 Broom A, Cheshire L, Emmison M. Qualitative researchers' understandings of their practice and the implications for data archiving and sharing. Sociology 2009;43:1163–80.

35 Valles M, Miguel S, Corti L, Tamboukou M, Baer A Qualitative archives and biographical research methods. An introduction to the FQS special issue. Forum Qual Soc Res 2011;12 http://www.qualitativeresearch.net/index.php/fqs/article/view/1755. [Accessed 13 February 2013]

36 Heaton J Reworking Qualitative Data. Sage Publications: London 2004.

37 Thorne S. Ethical and representational issues in qualitative secondary analysis. Qual Health Res 1998;8:547–55.

38 Hammersley M. Qualitative data archiving: some reflections on its prospects and problems. Sociology 1997;31:131–42.

39 Lefroy J. How can the existing communication skills curriculum be optimised, in the light of experience here and elsewhere, and in congruence with the rest of the curriculum? Dissertation. Keele: Keele University 2010.

40 Irvine AW. Evaluation of the use of peer physical examination (PPE) in the teaching of basic physical examination skills to medical undergraduates. MA dissertation. Stafford: Staffordshire University 2011.

41 Irvine AW, Lefroy JE. Addressing the classroom–clinical practice divide. In: Communication in undergraduate medical education conference; 24th March; University of Manchester. Manchester: UK Council of Communication Teaching in Undergraduate Medical Education 2011.

42 Kozulin A, Chaklin S, Karpov Y et al. (eds.) Vygotsky's Educational Theory in Cultural Context. Cambridge: Cambridge University Press 2003.

43 Cook , S . Mind the theory/practice gap in nursing. J Adv Nurs 1991;16:1462–9.

44 Norman G. Research in clinical reasoning: past history and current trends. Med Educ 2005;39:418–27.

45 Piaget J. The Psychology of Intelligence. London: Routledge; 1950.

46 Helmore G, Piaget J. Piaget: A Practical Consideration. Oxford: Pergamon Press 1969.

47 van der Zwet J, Zwietering PJ, Teunissen PW, van der Vleuten CPM, Scherpbier AJJA. Workplace learning from a socio-cultural perspective: creating developmental space during the general practice clerkship. Adv Health Sci Educ 2011;16:359–73.

48 Kneebone R. Perspective: simulation and transformational change: the paradox of expertise. Acad Med 2009;84 (7):954–7.

49 Epstein RM. Mindful practice. JAMA 1999;282 (9):833-9.

50 Shapiro J, Kasman D, Shafer A. Words and wards: a model of reflective writing and its uses in medical education. J Med Humanit 2006;27 (4):231–44.
51 Rees C, Knight LV, Wilkinson CE. 'Doctors being up there and we being down here': a metaphorical analysis of talk about student/doctor-patient relationships. Soc Sci Med 2007;65:725–37.

52 Riessman CK. Narrative Methods for the Human Sciences. Thousand Oaks, CA: Sage Publications 2008.

53 Alderson P, Farsides B, Williams C. Examining ethics in practice: health service professionals' evaluations of in-hospital ethics seminars. Nurs Ethics 2002;9:508–21.

54 Frey JH, Fontana A. The group interview in social research. Soc Sci J 1991;28:175–87.

55 Williams C, Ehrich K, Farsindes B, Scott R. Facilitating choice, framing choice: staff views on widening the scope of preimplantation genetic diagnosis in the UK. Soc Sci Med 2007;65:1094–105.

56 Morgan DL Focus Groups as Qualitative Research, 2nd edn.Newbury Park, CA: Sage Publications 1997.

57 Yardley S, Brosnan C, Richardson J. Sharing methodology: a worked example of theoretical integration with qualitative data to clarify practical understanding of learning and generate new theoretical development. Med Teach 2012, doi:10.3109/0142159X.2012.733045 (Online).

Chapter 4

Designing whole-task learning opportunities for integrated end-of-life care: a practitioner-derived enquiry

Sarah Yardley[1]

Claire Hookey[2]

Janet Lefroy[2]

[1]Primary Care and Health Sciences, Keele University

[2]Keele University School of Medicine

Summary

Introduction   Knowledge and skills to contribute to high-quality patient-centred end-of-life care are essential for newly qualified doctors. End-of-life care is a multifaceted complex task but learning opportunities are often fragmented in undergraduate curricula. Whole-task models provide a framework for delivery of learning activities which equips students to function in variable complex contexts.

Objectives   To create learning experiences that would help students to integrate the knowledge, skills and behaviours needed when encountering patients near the end-of-life, including during transitions between primary and secondary care settings.

Methods   We describe the development, implementation, content and evaluation of an educational intervention for undergraduate medical students. This comprised a study day offering whole-task learning opportunities for integrated end-of-life care combined with a longitudinal placement. Our research drew on two data sources: reflective summaries and end-of semester, online, anonymous student questionnaires.
Thematic analysis of student reflective writing demonstrated learning in multiple domains.

Results and discussion   Our intervention formed an important link between classroom learning and clinical practice due to its design according to whole-task models: learners were engaged in solving real-world problems, new knowledge was applied and integrated in practice, students built on existing knowledge longitudinally, and experienced professionals supported putting knowledge into action. Although set in the UK the issues we address are of relevance worldwide.

Introduction

In this paper we describe an educational intervention based in primary care, designed to equip final year undergraduate students with the necessary core knowledge, skills and behaviours to deliver and develop palliative care in generalist settings on commencing their postgraduate careers. Our objectives are to discuss the benefits of using whole-task learning theories in curriculum design, describe our educational intervention and to present outcomes and impact from the perspectives of students. While none of the individual elements of our intervention are unique, we believe the use of educational theory to inform how we have brought these elements together as part of our educational strategy for palliative care will be of interest to others seeking to ensure similar education is of high quality and utility.

End-of-life care: an education challenge

In England approximately 470 000 people die each year, of whom it is estimated 355 000 will require palliative care (1). Similar patterns are observed worldwide. This need will increase due to an aging population who are living longer with chronic diseases and greater morbidities. All doctors need a core understanding of end-of-life care provision across primary and secondary care settings, including issues related to transition from one setting to another. Although

palliative care has long been identified as a longitudinal theme within undergraduate curricula, and classroom-based teaching time has increased in many medical schools, it cannot be assumed that this leads to learning. A lack of meaningful exposure to dying patients has been postulated as explanation for newly qualified doctors feeling unprepared to care for patients at the end-of-life (2). This perception is echoed by palliative care teaching co-ordinators within medical schools (3). On graduation 'Tomorrow's Doctors' must be able to 'contribute to the care of patients and their families at the end-of-life, including management of symptoms, practical issues of law and certification, and effective communication and team-working'(4). Knowledge and skills needed to deliver patient-centred end-of-life care are often fragmented within undergraduate curricula and may not be explicitly linked to end-of life care. The reality of practice places newly qualified doctors in busy complex environments where it is seldom possible (or appropriate) to focus on one step of a task at a time. Interruptions are frequent and disruptive. Doctors must integrate their learning to work flexibly for delivery of individualised high quality end-of-life care.

Methods

Context

Our practitioner-derived enquiry was conducted in a single UK medical school with an integrated spiral five-year undergraduate curriculum using a hybrid model of learning opportunities (incorporating problem-based learning, experiential learning within the medical school, laboratory sessions, lectures and clinical placements). Students study different aspects of curriculum topics and build increasing depth of knowledge by returning to topics as they progress. End-of-life care learning opportunities are summarised in Table 1.

Table 1 End-of-life care learning opportunities	
Year 1	Problem-based learning cases and lectures on palliative care, end-of-life care, and a 'good death'. Start communication skills training
Year 2	Problem-based learning cases and lectures on sudden death, patient perspectives on dying and issues in patient care. Communication skills training in history-taking and giving information
Year 3	Clinical practice exposure to end-of-life issues. Communication skills training in breaking bad news and talking with relatives
Year 4-5	Clinical apprenticeship exposure to end-of-life issues, withdrawal of treatment, breaking bad news including our educational intervention described in this paper

In the final year, students are divided into two rotations: one rotation spends the first 15 weeks in general practice followed by 15 weeks in hospital practice while the other does the reverse. Our study day (see below) occurs midway through the general practice assistantship during which students have opportunities to contribute to the palliative care of a patient. To support this

learning, students write a reflective summary of their learning for formative assessment by their GP tutor.

Conceptual orientation: the theory of whole-task learning

The idea that competence for complex tasks can be gained through the development of context-free generic skills has long been criticised (5,6). Despite this there has been a longstanding trend in medical education to teach task elements separately and to teach skills in step-wise fashion (7). For example, students may learn about breaking bad news in communication skills sessions, the legalities of prescribing or death certification in lectures or problem-based learning cases, and then during clinical placements be disorientated by apparently arbitrarily different ways of managing the same symptoms depending on the setting in which a patient is seen (8,9). These factors may explain why developing transferable learning and achieving true integration of learning has taxed students and educators alike. Transfer will always be impaired if there is a perceived gap between the reality of the medical school and the reality of medicine as practiced in the workplace (10).

'Whole-task models' provide an alternative framework to develop activities designed to equip students to function in variable complex contexts. A whole-task model is one which seeks to 'deal with complexity without losing sight of the relationships between the elements' and which avoids fragmentation or compartmentalising parts of a multi-task goal (11). Such models support students who need to develop complex cognitive skills and professional competencies that they can apply in uncertain and unpredictable contexts. An example of the structure of a whole-task model is the Four-component Instructional Design Model (12). Its components are: meaningful learning tasks based on real life and requiring multi-tasking performance; supportive information bridging what students know with what they need to learn; procedural information; and part task practice. We were convinced of the potential usefulness of whole-task models to address the challenge of end-of-life care education after considering the work of van Merriënboer and Kester (11). Students value learning skills in an integrated way (13). Whole-task learning models provide a theoretical framework to build on their enthusiasm in order to develop skills that students can more easily transfer and apply in 'real life' practice. In this article we describe the development, implementation and evaluation of a whole-task model of learning to deliver an integrated approach to providing for end of-life care interactions with patients. Although set in the UK, the issues we address concern patients worldwide.

Educational intervention: opportunities for learning integrated end-of-life care

A study day was created, drawing on local expertise in palliative medicine, general practice and communication skills. This ran twice over the course of the final year, so that each student could attend near to the mid-point of their 15-week GP assistantship. The study day presented scenarios that demanded students integrate knowledge, skills and behaviours needed when encountering patients near the end of life (Box 1).

Students were advised to focus in depth on one patient encountered during the GP assistantship, selected with the help of their GP tutor. The patient was to have supportive and palliative care needs. The student's knowledge of this patient was used in discussions on the study day in an anonymised format and formed the basis of a reflective summary for the student portfolio. To formatively evaluate learning, this reflection was submitted to the medical school at the end of the GP assistantship in the first iteration of the new course. In subsequent years it is being submitted to and discussed with the GP tutor (Box 2).

The study day begins with a 60-minute plenary session led by a consultant in palliative medicine, delivered in a question and answer style format and including:
      • competencies required of Foundation year doctors relating to end-of-life care
      • key transitions for a patient approaching end of life

- diagnosis of dying
- use of the Liverpool Care Pathway14 in hospital and community settings
- symptom control at end of life
- ethical issues including 'Do Not Attempt Resuscitation' orders.

Students then move into small tutorial groups and work through two clinical scenarios for the remainder of the day. The scenarios integrate group discussion, practical tasks, and role play using simulated scenarios (see Boxes 3 and 4). The simulated patients were lay people trained by the medical school to act as patients for student consultations and to give feedback afterwards. Each scenario runs over three hours with three sections allowing for interim discussion and debriefings with a clinically experienced facilitator trained in communication skills small-group teaching.

Box 3 Greg and Laura Brown

Laura Brown is a 58-year-old lady with ovarian cancer who develops bowel obstruction. Initially Laura needs to be informed that she has obstruction and that there is no surgical treatment for this. The students have to prescribe symptom control medication including a syringe driver according to advice from a palliative care team, and communicate the diagnosis and treatment plan to Laura and her husband.
In the second part of the scenario, Laura is symptomatically better and wishes to go home. Her husband is concerned about this. The students discuss the practicalities of such a discharge, prescribe the medication she needs to take home, and complete a DNAR form. One student role-plays the conversation with Laura and her husband regarding discharge planning.
In the final scenario, Laura is now at home and has become more poorly. A student role-plays a discussion with Laura and her husband about her deterioration and what is likely to happen now (although Laura is very sleepy and most of the conversation is with Greg).

Scenarios were semi-scripted allowing for uncertainty and flexibility in the interaction between the simulated patients and the students. Students could 'pause' to discuss issues or seek support and ideas from their peers, the simulated patients and the tutor. At the end of each scenario all present, including the simulated patients, participated in feedback and reflective debriefings.

Methods of analysis

Evaluation of the students' learning about end-of life care drew on two sources of data: reflective summaries (a rich source of data about the impact of learning experiences) and end-of- semester online anonymous student evaluation questionnaires. The latter included questions about the study day and opportunity for free text comments (see Table 3). In the first year of the intervention (2011/12), all submitted reflective summaries were read by one researcher (JL) and a

Box 4   Stanley and Nancy Jenkins

Stanley is a 77-year-old gentleman with dementia, admitted to hospital with a broken hip following a fall at his nursing home. Stanley has had surgery for his broken hip and been transferred to a rehabilitation unit. Unfortunately he has not recovered well, he developed a chest infection and despite 3 days of intravenous antibiotics he is now semiconscious and agitated. The consultant has decided that it is now appropriate to prepare for Stanley's likely death. A student role plays the conversation with his wife explaining this. All the students complete a Liverpool Care Pathway and prescribe appropriate symptom control medication.

In the second part of the scenario, a sputum culture result has come back showing resistance to the antibiotics Stanley had been given and sensitivity to an alternative antibiotic. The students discuss possible approaches to management. The consultant suggests trying intravenous antibiotics. A student role-plays the ensuing discussion with Mrs Jenkins.

Stanley does improve temporarily following these antibiotics, but in the final scenario he had become more unwell again and died over the weekend. Mrs Jenkins has asked to speak to a doctor when she comes to collect the death certificate. The students discuss death certification and referral to the coroner and complete a death certificate. One student role-plays a conversation with a distressed Mrs Jenkins, including explaining that, because of the fall Stanley suffered in the nursing home, the death certificate cannot be issued and the case needs to go to the coroner.

thematic analysis was performed, paying particular attention to themes related to student perceptions of what they had learned, and connections or comparisons made between workplace and study day learning. A selection of the summaries (*n*=14 approx. 18% of the total submitted, *n*=79) containing reflections about the study day were also analysed by one of the other authors (SY) to confirm the findings. All three authors met to discuss the coding and themes within the analysis. This process emphasised the importance of whole-task learning, and personal involvement in patient care as well as verifying the thematic analysis.

Results

Direct evaluation of intervention
The new final-year curriculum is currently midway through its second cycle, therefore, the study day has currently run three times. Table 2 summarises the numbers of students participating in each rotation and providing data through their evaluations and reflective summaries.

The study day was very positively evaluated by students in all three rotations (*n*=129) in questionnaires at the end of their 15 weeks in general practice (see Table 3). One-hundred and twenty-four (96.1%) students agreed/strongly agreed that the study day has helped them understand issues related to dying with 118 (91.4%) agreeing/strongly agreeing that they had been able to apply their learning in other parts of the course including clinical practice (118 (91.4%) agreeing/strongly agreeing), development of communication skills (113 (87.5%) agreed/strongly agreed) and prescribing skills (119 (92.2%) agreed/strongly agreed). Several students provided free text comments. These included very few negative reactions. One student commented on the emotional challenges of the activities and expressed concern that they were

expected to undertake these tasks while another commented that they thought the day was too long. Positive reactions were received to the integration of communication and practical skills that required clinical knowledge about end-of-life care as well as two students requesting more time to learn about these issues. One student was critical of an assistantship supervisor's attitudes to DNAR decisions which contrasted to learning on the study day.

Table 2 Student participants

	2011/12 rotation 1 study day	2011/12 rotation 2 study day	2012/13 rotation 3 study day	Total
No of students participating	54	57	65	176
No answering evaluation questions	41	39	49	129(73%)
No submitting reflective summaries	47	32	N/A	79*

*Although students were informed that it was mandatory to submit reflective summaries in 2011/12 only 79 out of 111 did so and as these did not form part of the summative assessment of the year those who had not submitted by finals (despite reminders) were not barred from qualifying. The process has since been modified so students submit their reflective summaries to local GP tutors rather than centrally to the medical school.

Table 3 Questions asked in the online end of semester (anonymous) student survey

Likert scale responses on a five-point scale	Strongly disagree n (%)	Disagree n (%)	Unsure n (%)	Agree n (%)	Strongly agree n (%)	Total n
The end-of-life care day has helped me understand the issues related to dying	0	3(2.3)	2(1.6)	83(64.3)	41(31.8)	129
I have been able to apply what I learned in the end-of-life day to other parts of the course	0	5(3.9)	6(4.7)	84(65.1)	34(26.4)	129
The end-of-life care day has helped me develop my skills in communicating with dying patients and their carers	0	6(4.7)	10(7.8)	58(44.9)	55(42.3)	129
The end-of-life care day has helped me develop my prescribing skills for palliative care	0	1(0.8)	9(7.0)	80(62.0)	39(30.2)	129
I have been able to apply what I learned in the end-of-life care day to my clinical practice	0	5(3.9)	6(4.7)	84(65.1)	34(26.4)	129

Impact identifiable in written reflections

Seventy-nine (71.1% of the year cohort) students submitted reflective summaries. Thematic analysis identified outcomes arising directly from the study day:

> • *Skills practice* was the most helpful, *challenging, realistic and prepared them for practice*. Only one student thought that she would not use these skills apparently believing that only consultants do these tasks.
> • *Knowledge* was appreciated – diagnosis of dying, prescribing, syringe drivers, Liverpool Care Pathway, DNAR, Death certification.
> • *Awareness* of relatives and self to a lesser extent than in the placement encounters but one student felt that the EOL day was very difficult because of 'painful memories'.
> • Many commented on how helpful they felt this learning would be in their role as a junior doctor.

The reflective summaries also demonstrated that the students had used the information and experience gained on the study day to inform their reflection on real-life experiences with end-of-life patients. Learning identified following their study day and GP assistantship from their involvement with patients at the end-of-life included:

> • a gain of *knowledge* about community team-working, pre-planning with patients, best practice standards and the Liverpool Care Pathway making dying at home really an option
> • the importance of learning *whole tasks* including the complexities of *combining* communication skills with clinical decision making and practical tasks.
> • *role modelling* by GPs and palliative care nurses was prominently mentioned including sometimes in students' personal family experiences of dying (in these there was detailed recall of doctor behaviour)
> • enhanced *awareness of patient perspectives* from personal involvement in their care and some surprise from students at finding calm acceptance of diagnoses and dying. A variety of patient wishes were described.
> • enhanced *significance given to discussion with and consideration for relatives/carers*
> • *skills practised in end-of-life care on GP assistantship*: discussion of medication, services, feelings, supporting relatives, Do not attempt resuscitation (DNAR) discussions, negotiating on behalf of patient, diagnosing death.
> • *self-awareness and emotional intelligence* – sense of being an awkward intruder; emotionally attached; coping strategies; rawness of discussing DNAR.
> • *some coyness about mentioning the death of their patient* – of 47 only 12 specified that their patient had died, and the euphemism 'till the end' was used by one student. Those who did specify death did not always describe whether it was a *'good death'* (seven good deaths were mentioned, the concept discussed by some, plus two 'not so good' deaths)
> • *criticism of practice* (they were asked to criticise but were generally highly positive about the care given):
> — care given in the past especially around 'breaking bad news', withholding information, delayed diagnosis (only two patients were unhappy about their care, a few more students had picked up on defects in care even when the patient was content with their care)
> — waiting for the patient to ask for help (anticipatory care was suggested)
> — conflict of curative and palliative aims
> — inadequate support for the dying (this was in two cases only).

Discussion

A challenge for educators involved in curriculum design is that often the need and opportunity to redesign elements of a curriculum arises within very limited timeframes for implementation and delivery.

It is not always possible to set up a full programme of research within these timeframes. However, by using methods drawn from research practices it is possible to ensure that educational interventions are evaluated with rigour. Our thematic analysis of reflective writing has facilitated evaluation of the curriculum as well as deepening understanding of students' perspectives on learning experiences. In addition it has helped us to identify future research questions.

The development of a new curriculum at our medical school provided opportunities for improved learning experiences in generalist provision of palliative care. We developed this by drawing on appropriate educational theory to guide development and implementation of our educational intervention and evaluating the lived experiences of students when the intervention was delivered. We have illustrated how a whole-task model can be used to design and integrate educational interventions with clinical practice with the aim of preparing students for the realities of delivering end-of-life care.

Students reflected on the knowledge and skills they would need but also the impact of end-of-life care on patients, families and themselves. They were able to draw on the role modelling of their GP supervisors to develop conceptualisations of good medical practice. Our intervention integrates communication skills teaching in end-of-life care with practical tasks including prescribing, use of the Liverpool Care Pathway, writing a death certificate, and completing a DNAR form. This 'double integration', both within the study day and from classroom to practice, appears to be an effective model for end-of-life teaching for final-year medical students.

The UK End-of-life Care Strategy15 states '…a newly qualified doctor on the first day of their Foundation Year One (FY1) programme may be faced with a dying patient and a distressed relative. He or she may also have to deal with a patient who has just learnt that they have an advanced, incurable illness. These doctors need to have the necessary core competences to enable them to deal with these situations without adverse consequences for the patient, family members and themselves'.

To achieve this goal requires working knowledge of end-of-life care within the community and the role of GPs, who alongside other community healthcare professionals provide the majority of palliative care services outside of specialist settings. Although there was a theoretical risk that a student and their GP tutor might not identify a suitable patient during the 15-week placement this did not occur. *Dying Matters* estimate that around 1% of any GP's patients will die in any given year (16). Given that our requirement is to identify people with supportive and palliative care needs (for example, people identified for supportive and palliative care registers, rather than a more narrow category of those who are imminently and actively dying), and the length of placement in general practice (15 weeks), we do not anticipate problems with identifying patients

in the future. We recognise that any longer-term impact arising from the outcomes reported above cannot be identified within the timespan of the final undergraduate year.

Our intervention is now part of the established curriculum but further work is required to know whether (and how) students put learning into practice as they commence their postgraduate careers. We hope, nonetheless, that learning arising from participation in the intervention will give students the motivation to continue to develop their practice in end-of-life care.

References

1 Hughes-Hallett T, Craft A and Davies C (2011) *Palliative Care Funding Review: funding the right care and support for everyone. Creating a fair and transparent funding system: the final report of the Palliative Care Funding Review.* www.palliativecarefunding.org.uk (accessed 28/09/12). http://palliativecarefunding.org.uk/PCFRFinal%20Report.pdf

2 Gibbins J, McCourbrie R and Forbes K (2011) Why are newly qualified doctors unprepared to care for patients at the end-of-life? *Medical Education* 45: 389–99.

3 Gibbins J, McCoubrie R, Maher J, Wee B and Forbes K (2010) Recognizing that it is part and parcel of what they do: teaching palliative care to medical students in the UK. *Palliative Medicine* 24: 299–305.

4 General Medical Council (2009) *Tomorrow's Doctors. The duties of a doctor registered with the General Medical Council*. GMC: London.

5 Van der Vleuten CPM, Dolmans DHJM and Scherpbier AJJA (2000) The need for evidence in education. *Medical Teacher* 22: 246–50.

6 Norman G, Young M and Brooks L (2007) Non-analytical models of clinical reasoning: the role of experie*nce. Medical Education* 41: 1140–5.

7 Windish DM, Price EG, Clever SL, Magaziner JL and Thomas PA (2005) Teaching medical students the important connection between communication and clinical reasoning. *Journal of General Internal Medicine* 20: 1108–13.

8 Schmidt HG, Norman GR and Boshuizen HPA (1990) A cognitive perspective on medical expertise: theory and implications. *Academic Medicine* 10: 611–21.

9 Dornan T, Hadfield J, Brown M, Boshuizen H and Scherpbier A (2005) How can medical students learn in a self-directed way in the clinical environment? Design-based research. *Medical Education* 39: 356–64.

10 Eva KW (2004) What every teacher needs to know about clinical reasoning. *Medical Education* 39: 98–106.

11 van Merriënboer JJG and Kester L (2008) Whole task models in education. In: Spector JM and Harris PA (eds) *Handbook of Research on Educational Communications and Technology.* 3rd edn. Taylor & Francis, pp. 441–56.

12 van Merriënboer JJG (1997) *Training Complex Cognitive Skills.* Educational Technology Publications: Englewood Cliffs, NJ.

Authors' Addendum (July 2013)

The educational innovation and associated research data collection that this paper is based on was completed in January 2013. Readers may be aware that on the 15 July 2013 an independent review, chaired by Baroness Neuberger of the use of the Liverpool Care Pathway (LCP) in England was published. This review is supportive of the principles of the LCP, when implemented appropriately and by adequately trained healthcare professionals. It acknowledges, however, that the principles have not always been translated into practice. In the light of this it recommends the withdrawal of the LCP in hospitals over the next 6–12 months, replacing it with individualised end-of-life care plans (see www.gov.uk/government/uploads/system/uploads/attachment_ data/file/212450/Liverpool_Care_Pathway.pdf). It remains imperative that people with irreversible life-threatening conditions receive good palliative care including appropriate care if they are actively and imminently dying. The educational principles of the work reported in this paper therefore stand. We have now modified the study day content to account for the changing external context. Readers are encouraged to read the independent review and consider our work in the context of its emphasis on the need for better training. Those interested in the details of the modifications we have made are invited to contact the corresponding author.

Contributions (not included in the publication): SY and CH are Palliative Care doctors, JL is a General Practitioner. The authors worked together on devising the study day for learning integrated end-of –life care. JL and SY analysed the students' reflective summaries about this course. All authors discussed this analysis and the end of semester evaluations. CH wrote the outline draft of the paper. SY led the formulation of the conclusions, and the main body of this paper. All authors contributed to revisions and approved the final manuscript for publication.

Chapter 5

Development and face validation of an instrument to assess and improve clinical consultation skills

Janet Lefroy[*]

Simon Gay[*]

Sheena Gibson[*]

Stephen Williams[*]

Robert K McKinley[*]

[*]Keele University School of Medicine, Keele, United Kingdom

Abstract

Context: Development of medical students' consultation skills with patients is at the core of UK General Medical Council's (GMC) Tomorrow's Doctors guide (2009). Teaching and assessment of these skills must therefore be a core component of the medical undergraduate curriculum. The Calgary Cambridge guide to the medical interview and the Leicester Assessment Package (LAP) provide a foundation for teaching and assessment, but both have different strengths.

Objective: To develop and validate a comprehensive set of generic consultation competencies.

Design: The Calgary Cambridge guide to the medical interview was revised to include 'clinical reasoning', 'management', 'record keeping' and 'case presentation'. Each section was populated with competencies generated from Tomorrow's Doctors (2009), the LAP and the Calgary Cambridge guide to the medical interview. A Delphi validation study was conducted with a panel drawn from hospital and general practice clinical tutors from eight UK medical schools.

Main outcome measures: *A priori* consensus standards for inclusion (or exclusion) of an element were: at Stage 1 ≥70% agreement (or disagreement) that the item should be included; at Stage 2 ≥50% agreement (or disagreement) that the item should be included. If more than 10% of respondents suggested a thematically similar new item (or rewording of an existing item) in Stage 1, it was included in Stage 2.

Results: The design stage resulted in a set of 9 categories of consultation skills with 58 component competencies. In the Delphi study all the competencies reached 70% agreement for inclusion, with 24 suggested amendments, all of which achieved consensus for inclusion at Stage 2.

Conclusion: We have developed a **ge**neric **co**nsultation **s**kills assessment framework (GeCoS) through a rigorous initial development and piloting process and a multi-institutional and multi-speciality Delphi process. GeCoS is now ready for use as a tool for teaching, formative and summative assessment in any simulated or workplace environment in the hospital or community clinical setting.

Introduction

The UK General Medical Council's (GMC) Tomorrow's Doctors guide (2009) has laid new emphasis on the importance of the 'Doctor as a practitioner' and, in paragraphs 13 to 15, describes the skills the medical graduate needs to acquire to consult with patients [1]. These are a complex amalgam of cognitive, psychomotor, communication and interpersonal skills which, like any other set of high level skills, need sustained repeated deliberate practice [2, 3], with support from tutors through formative assessment. Such formative assessment should be congruent with both the curriculum and with summative assessment.

The Calgary Cambridge guide to the medical interview is used by many medical schools worldwide as the basis of their communication skills curricula [4, 5]. However, it does not address the additional cognitive skills required for making a diagnosis or identifying appropriate management options and, although some congruent assessment schedules have been developed, they are context specific and have not been widely evaluated [4]. Conversely the Leicester Assessment Package (LAP) [6, 7] was developed and used to support both formative and summative assessment of undergraduates [8, 9] trainees [10, 11] and established practitioners [12, 13, 14] in the UK and internationally, and has been utilised to promote congruence between assessment and the curriculum [8]. Furthermore, it contains a series of generic strategies for improvement of skills mapped onto each of its competencies which can be used by tutors as the basis for preparing feedback [15], thus addressing the problem of specificity of the content of feedback [15, 16]. It does not, however, map onto a particular model of the consultation and, as the published version is almost 20 years old, it may be dated.

We consider that the Calgary Cambridge guide to the medical interview and the LAP each have strengths which compliment the other's weaknesses, and that they could be usefully combined. We now describe a modification of the Calgary Cambridge guide to the medical interview and the development and face validation of a **ge**neric **co**nsultation **s**kills assessment tool (GeCoS) which would be evaluated for use in formative and summative assessment in both workplace and simulated environments, such as the 'clinical skills laboratory' and in OSCEs.

Methods

Development

Modified Calgary Cambridge framework for the consultation: Keele University School of Medicine has adopted an integrated model for consultation skills [16] (Figure 1) which brings together communication, physical examination, patient management, clinical problem solving, information management and procedural skills. With advice from Dr Jonathan Silverman (Cambridge University, UK) we adapted the Calgary Cambridge guide to the medical interview to the needs of our curriculum by adding a clinical reasoning stream (in the background throughout the consultation), recording the consultation and presenting the patient to colleagues. The visual representation of clinical reasoning emphasises its contribution to gathering information, performing the physical examination, choosing investigations, formulating a diagnosis, negotiating a management plan, making a clinical record and presenting the case. The framework also draws attention to the processes and content of each stage of the consultation. The final version of the framework can be seen in Appendix 1.
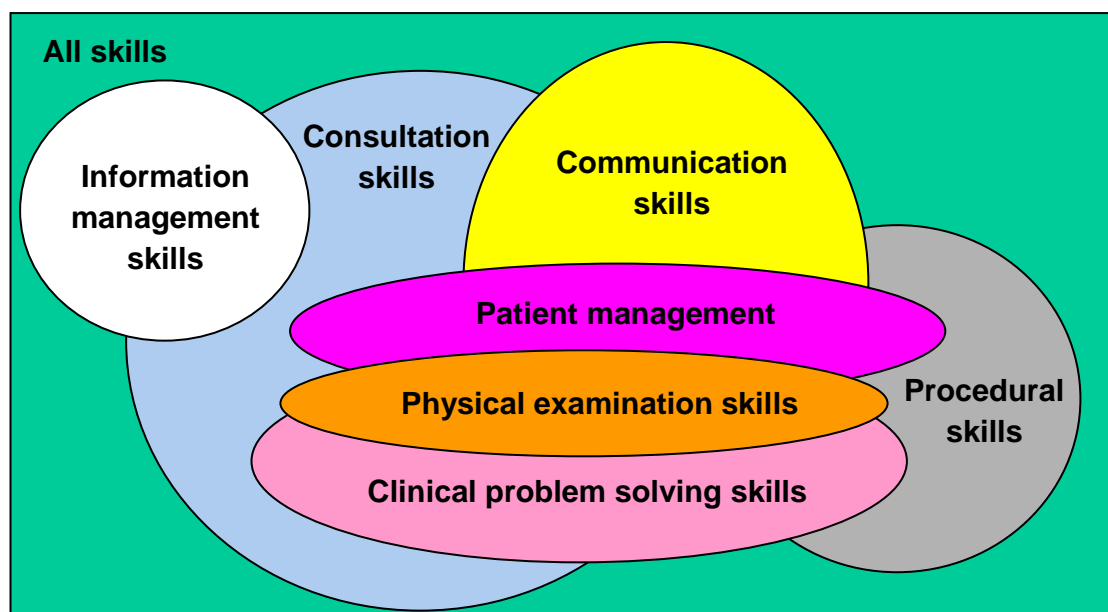
Figure 1: An integrated model for consultation skills [16]; At Keele University School of Medicine the skills used in encounters with patients are taught and assessed as an integrated skill set. For example, communication, physical examination and problem solving skills are taught and can be assessed together with clinical procedural skills

Generic consultation skills instrument (GeCoS): The development of GeCoS was undertaken by the authors (four general practitioners and one paediatrician) with advice from Dr Jonathan Silverman (Cambridge University, UK). We systematically identified similarities and differences between the 42 competencies in LAP [7], the 71 in the Calgary Cambridge guide to the medical interview and the GMC's Tomorrow's Doctors guide (2009) [1]. Component competencies identified from each were allocated to the categories in the revised framework, condensing them when possible to keep the list concise. The terminology of LAP was updated to match that in the Calgary Cambridge guide to the medical interview where this was felt helpful. The conventional term "Management" was chosen for the Calgary Cambridge stage "Explanation and Planning" as we felt it included aspects of selection of therapy. This was an iterative process involving each of the authors initially reviewing and condensing the list of skills, discussing their changes and reaching consensus with the rest of the team and then piloting of successive versions of the instrument in formative assessment of students in the skills lab with simulators, and in the workplace with real patients. This resulted in an instrument with 9 categories of consultation skills and 58 component competencies (Table 1).

Validation study
Questionnaires:
A two round modified Delphi process was used to establish the face validity of GeCoS. The first round Delphi questionnaire was based on that used for the original face validation of the LAP [7] and of other skills assessment tools [17, 18, 19], but we modified the response scale to that of McIlwaine et al [20] ("very relevant and succinct", "relevant but needs minor alteration", "unable to assess relevance without item revision or item in need of such revision that it would no longer be relevant" and "not relevant").

Table 1. results of Stage 1 validation questionnaire

	Not relevant (%)	Unable to assess relevance (%)	Relevant but needs minor alteration (%)	Very relevant and succinct (%)	Agreement (%)	Revision suggested (n)
Category 1: OPENING	**0.0**	**0.0**	**22.4**	**77.6**	**100**	**14**
1.1. Introduces self	0.0	0.0	18.8	81.3	**100.0**	12
1.2. Establishes identities of patient and third parties and preferred forms of address	0.0	1.6	12.5	85.9	**98.4**	6
1.3. Establishes agendas	1.6	9.4	31.3	57.8	**89.1**	23
Category 2: HISTORY	**0.0**	**1.5**	**25.4**	**73.1**	**98.5**	**16**
2.1. Enables patient to fully elaborate presenting problem(s)	0.0	1.7	13.6	84.7	**98.3**	9
2.2. Listens attentively	1.7	1.7	8.5	88.1	**96.6**	8
2.3. Skilled use of questioning	0.0	5.1	33.9	61.0	**94.9**	21
2.4. Clarifies words used and/or symptoms presented by patient as appropriate	0.0	0.0	10.2	89.8	**100.0**	5
2.5. Recognises and responds appropriately to verbal and non-verbal cues	3.4	1.7	8.5	86.4	**94.9**	7
2.6. Sequence of events	1.7	5.1	13.6	79.7	**93.2**	10
2.7 Symptom analysis	1.7	5.1	22.0	71.2	**93.2**	13
2.8 Effect on the patient	0.0	6.8	18.6	74.6	**93.2**	14
2.9 Patient's ideas, concerns and expectations	0.0	0.0	11.9	88.1	**100.0**	6
2.10 Background information including physical, social and psychological factors	1.7	1.7	18.6	78.0	**96.6**	11
Category 3: EXAMINATION	**0.0**	**0.0**	**9.0**	**91.0**	**100**	**7**
3.1. Obtains initial and ensures continuing consent	0.0	1.7	15.3	83.1	**98.3**	10
3.2. Displays competent practice of infection prevention	0.0	0.0	8.5	91.5	**100.0**	5
3.3. Displays sensitivity to patient's needs and dignity	1.7	0.0	6.8	91.5	**98.3**	5
3.4. Gives clear instructions and explanations of process	0.0	1.7	5.1	93.2	**98.3**	5
3.5. Performs examination competently	1.7	0.0	10.2	88.1	**98.3**	8
3.6. Elicits the physical signs	1.7	1.7	20.3	76.3	**96.6**	14
Category 4: PATIENT MANAGEMENT	**1.5**	**1.5**	**29.9**	**67.2**	**97**	**21**
4.1. Relates explanations to patient's perspective	0.0	1.7	18.6	79.7	**98.3**	10
4.2. Gives clear information in small chunks	1.7	0.0	15.3	83.1	**98.3**	9
4.3. Negotiates a mutually acceptable plan with patient and/or third parties	0.0	0.0	8.5	91.5	**100.0**	6
4.4. Reassures appropriately	1.7	0.0	13.6	84.7	**98.3**	8
4.5. Checks understanding	0.0	0.0	6.8	93.2	**100.0**	5
4.6. Gives key evidence-based information	1.7	0.0	25.4	72.9	**98.3**	15
4.7. Explores available options, risks and benefits	0.0	0.0	6.8	93.2	**100.0**	3
4.8. Gives appropriate advice on selfcare and lifestyle modification	1.7	0.0	6.8	91.5	**98.3**	6
4.9. Investigates appropriately	1.7	1.7	11.9	84.7	**96.6**	8
4.10. Prescribes rationally	1.7	0.0	20.3	78.0	**98.3**	14
4.11. Refers appropriately	1.7	0.0	10.2	88.1	**98.3**	6
4.12. Makes appropriate use of opportunities for health promotion	5.1	1.7	8.5	84.7	**93.2**	9
4.13. Agrees appropriate follow-up	1.7	1.7	13.6	83.1	**96.6**	10

Table 1. (continued)

	Not relevant (%)	Unable to assess relevance * (%)	Relevant but needs minor alteration (%)	Very relevant and succinct (%)	Agreement (%)	Revision suggested (n)
Category 5: PROBLEM SOLVING	**3.0**	**0.0**	**13.4**	**83.6**	**97**	**12**
5.1. Seeks relevant and specific information from patient's record or third parties	3.4	0.0	8.5	88.1	**96.6**	7
5.2. Generates appropriate working diagnoses or problem list	0.0	0.0	5.1	94.9	**100.0**	3
5.3. Seeks relevant and discriminating information from history, examination and investigations to help confirm or refute working diagnoses	1.7	0.0	6.8	91.5	**98.3**	4
5.4. Correctly interprets information obtained	3.4	0.0	3.4	93.2	**96.6**	2
5.5. Applies basic, behavioural and clinical sciences to solution of patient's problem	5.1	1.7	8.5	84.7	**93.2**	4
5.6. Recognises limits of competence and acts accordingly	0.0	1.7	6.8	91.5	**98.3**	4
Category 6: BUILDING THE RELATIONSHIP	**4.5**	**1.5**	**14.9**	**79.1**	**94**	**15**
6.1. Develops and maintains a professional relationship with patient	1.7	1.7	6.8	89.8	**96.6**	4
6.2. Respects the patient's ideas, beliefs and autonomy	1.7	1.7	6.8	89.8	**96.6**	5
6.3. Responds empathically	5.1	0.0	1.7	93.2	**94.9**	2
6.4. Fosters co-operation	8.5	5.1	20.3	66.1	**86.4**	18
Category 7: ORGANISATION	**0.0**	**0.0**	**16.4**	**83.6**	**100**	**10**
7.1. Optimises the setting	8.5	11.9	10.2	69.5	**79.7**	11
7.2. Uses third parties appropriately	5.1	6.8	20.3	67.8	**88.1**	14
7.3. Exhibits a well-organised approach to gathering and giving of information	0.0	1.7	8.5	89.8	**98.3**	7
7.4. Makes organisation of consultation overt to patient	11.9	8.5	11.9	67.8	**79.7**	11
7.5. Prioritises agendas appropriately	1.7	8.5	18.6	71.2	**89.8**	13
7.6. Summarises appropriately	1.7	0.0	10.2	88.1	**98.3**	6
7.7. Uses time appropriately	3.4	0.0	8.5	88.1	**96.6**	7
Category 8: RECORD-KEEPING	**1.5**	**0.0**	**19.4**	**79.1**	**98.5**	**14**
8.1. Makes concise and accurate notes without interfering with dialogue or rapport	1.7	0.0	8.5	89.8	**98.3**	6
8.2. Diagnoses/problems	0.0	5.1	13.6	81.4	**94.9**	10
8.3. Relevant history and examination	0.0	1.7	10.2	88.1	**98.3**	6
8.4. Outline of management plan, investigations, referral and follow up	0.0	0.0	5.1	94.9	**100.0**	2
8.5. Information, instructions and special precautions given to the patient	0.0	0.0	10.2	89.8	**100.0**	4
Category 9: CASE PRESENTATION	**3.0**	**0.0**	**9.0**	**88.1**	**97**	**6**
9.1. Engages and orientate colleague	5.1	0.0	23.7	71.2	**94.9**	14
9.2. Delivers relevant detail with clarity and logical order	1.7	0.0	6.8	91.5	**98.3**	3
9.3. Transparent interpretation of data	3.4	1.7	27.1	67.8	**94.9**	15
9.4. Purposeful conclusion	3.4	0.0	20.3	76.3	**96.9**	13

* Full wording of item "Unable to assess relevance without item revision or item in need of such revision that it would no longer be relevant"

The questionnaire offered participants the opportunity to express an opinion on the relevance of all nine categories and 58 component competencies, to suggest rewording of any element, to add categories and competencies, to move competencies between categories and to reorder categories. The questionnaire was loaded on a commercial questionnaire administration website [21], piloted amongst clinical staff at Keele University and modified where necessary. The questionnaire contained 79 questions and in piloting took participants between 20 and 45 minutes to complete.

The second round questionnaire accompanied the results of the first round questionnaire, which are outlined below. Elements which entered the second round were proposed rewordings of original elements or new elements. Respondents were asked to choose between inclusion or exclusion of new elements, or between the old and new wording of reworded elements using the same response format as in stage 1. This 27 item questionnaire was piloted amongst clinical staff at Keele University and modified as necessary.

Definition of consensus: We used the same *a priori* consensus standards as previous Delphi studies [17, 18, 19]: 70% or greater agreement (the "very relevant and succinct" or "relevant but needs minor alteration" responses) or disagreement (the "unable to assess relevance without item revision or item in need of such revision that it would no longer be relevant" and "not relevant" responses) for inclusion or exclusion respectively in the first round, and 50% or greater agreement for inclusion or exclusion respectively in the second round. If 10% or more of respondents suggested a thematically similar additional element or rewording of an existing category or component in the first round, it would be included in the second round.

Participants: The panel was drawn from hospital and general practice clinical tutors who are assessors of medical students, in order to include experts in a broad range of consultation types. To obtain a multi-institutional view of what should be assessed, clinical skills tutors from other undergraduate Medical Schools were invited to participate via a key contact at each school. We aimed to recruit from schools which used the Calgary Cambridge guide to the medical interview, the Leicester Assessment Package and schools which had no affiliation to either instrument.

Recruitment of panel members was by email invitation. The invitations, study participant information leaflets and consent forms were sent:

1. To selected expert clinical tutors at Keele University considered representative of the speciality groups. Recruitment continued until 50 had agreed to participate (25 from hospital specialities, 25 from general practice)
2. Via a contact person at each of the other Medical Schools asking them to recruit up to 10 clinical skills tutors with affiliation to the university, who were considered to be experts in the field, and would be willing to participate.

Potential participants were asked to contact one of the authors (JL), following which a web link to the questionnaire would be sent to them. All potential participants were sent three reminders, the final reminder being from their institutional contact person. Responses were anonymous unless the participant expressed a desire to receive the results, in which case they included their

email address and were also sent the link to the second round questionnaire with a request to continue to participate and, subsequently three follow-up reminders if necessary.

Data processing: All categorical data and the free text responses from each question were downloaded from the website. Categorical data was imported into SPSS for analysis; free text responses were sorted by question and printed for analysis.

Analysis: Response to the Agreement / Disagreement scale was analysed using simple descriptive statistics. Free text responses (suggested modifications to existing elements or additional elements) were closely thematically analysed by pairs of the research team. Subsequently the research team met to discuss each pair's analyses and to agree a consensus between the pair, and the rest of the team, on the themes identified by respondents. The number of respondents who suggested each theme was noted.

Results

Stage 1:

Of the 96 people who consented to participation and were sent the survey link, 82 (85%) started and 59 (61%) completed the questionnaire. Of these 48 (59%) were male, 48 (59%) described themselves as general practitioners, 1 as practising in both general practice and hospital and 10 gave no reply. 55 (67%) described themselves as undergraduate teachers, 19 (23%) as postgraduate teachers, whilst 8 gave no response. 45 requested the results of Stage 1 and were invited to participate in Stage 2.

Responses to the questions seeking opinions on the relevance of the categories and individual competencies are summarised in Table 1. The nine broad categories were considered either 'relevant but needs minor alteration' or 'very relevant and succinct' by 94 to 100% of respondents, with 'Building the relationship' having the lowest agreement (94%) with 4.5% of respondents considering it 'not relevant'.

Agreement as to the relevance of the individual competencies varied from 80% (for items numbered 7.1 and 7.4) to 100%. All but six competencies were considered relevant by more than 90% of respondents. These were items numbered 1.3 'Establishes agendas', 6.4 'Fosters co-operation', 7.1 'Optimises the setting', 7.2 'Uses third parties appropriately, 7.4 'Makes organisation of consultation overt to patient' and 7.5 'Prioritises agendas appropriately'. It is of note that four of these six were from Category 7: "Organisation".

There was no consensus for changing the order of categories or moving components between categories.

There were a total of 608 free text comments on the 67 categories and components, with a median of eight (range 0 to 23, interquartile range 5 to 13) comments. Our prior definition of consensus included the statement that if 10% of respondents suggested a thematically similar

change to the text of GeCoS we would include the change in a second round. With 59 respondents completing the questionnaire, we took a cut-off of five respondents making a similar suggestion as the threshold to include a suggestion. There were four suggestions made by five or more respondents (listed in Table 2). We considered that 17 other suggestions better encapsulated competencies than our original statements and these were also included in the second round (Table 3). Of these 21, three were for renaming Categories 4 'Patient management', 5 'Problem solving' and 6 'Building the relationship', and six were suggestions to increase the patient centred approach of the instrument (items numbered 2.8, 3.1, 3.3, 6.4, 7.2 and 7.3). An additional two competencies were suggested by more than 5 respondents (Table 4). Although there was no consensus to remove competencies in the main part of the study, three respondents had identified an overlap between items 4.8 'Gives appropriate advice on self care and lifestyle modification' and 4.12 'Makes appropriate use of opportunities for health promotion' so we offered Stage 2 respondents the opportunity to exclude the latter.

Table 2. Rewordings suggested by 10% or more of respondents and results of Stage 2 validation questionnaire.

Category	Original	Revision	Suggested by N respondents	N(%)of 27 respondents preferring revised wording
HISTORY: Process	*2.3:* Skilled use of questioning	Skilled use of questioning including open and closed questions	13	19(70)
HISTORY: Content	*2.8:* Effect on the patient	Effect on the patient's life	5	15(56)
EXAMINATION	*3.6:* Elicits the physical signs	Elicits normal and abnormal findings	5	23(85)
PATIENT MANAGEMENT	PATIENT MANAGEMENT	MANAGEMENT	5	20(74)

Table 3. Rewordings suggested by fewer than 5 respondents but which might encapsulate competencies better than the original statements and results of Stage 2 validation questionnaire

Category	Original	Revision	N(%)of 27 respondents preferring revised wording
HISTORY: Process	*2.7:* Symptom analysis	Details of symptoms	16(59)
HISTORY: Content	*2.10:* Background information including physical, social and psychological factors	Relevant background information including: Past Medical, Drug, Family and Social History; Systems review; Factors influencing health *or* Relevant background information	12(44) 7(26)
EXAMINATION	*3.1:* Obtains initial and ensures continuing consent	Obtains and maintains consent	21(78)
	3.2: Displays competent practice of infection prevention	Displays competent practice of infection control	25(93)
	3.3: Displays sensitivity to patient's needs and dignity	Displays sensitivity to patient's needs and dignity; offers chaperone if appropriate	20(74)
PATIENT MANAGEMENT	*4.10:* Prescribes rationally	Prescribes rationally and accurately	23(85)
PROBLEM SOLVING	PROBLEM SOLVING	CLINICAL REASONING	22(81)
	5.3: Seeks relevant and discriminating information from history, examination and investigations to help confirm or refute working diagnoses	Seeks discriminating information from history, examination and investigations to help confirm or refute working diagnoses	18(67)
BUILDING THE RELATIONSHIP	BUILDING THE RELATIONSHIP	BUILDING AND MAINTAINING THE RELATIONSHIP	23(85)
	6.4: Fosters co-operation	Fosters collaboration	25(93)
ORGANISATION	*7.1:* Optimises the setting	Considers and optimises the setting	16(59)
	7.2: Uses third parties appropriately	Involves third parties appropriately	25(93)
	7.3: Exhibits a well-organised approach to gathering and giving of information	Exhibits a well-organised approach to gathering and sharing of information	24(89)
RECORD KEEPING	*8.4:* Outline of management plan, investigations, referral and follow up	Outline of management plan; therapy, investigations, referral and follow up *or* Outline of management plan	15(56) 10(37)
CASE PRESENTATION	*9.2:* Delivers relevant detail with clarity and logical order	Delivers clear and relevant detail in a logical order	18(67)
	9.3: Transparent interpretation of data	Communicates interpretation of data transparently	21(78)
	9.4: Purposeful conclusion	Draws purposeful conclusion	18(67)

Table 4. Additional competencies suggested by fewer than 5 respondents but considered important to include in second round and results of Stage 2 validation questionnaire

Category	Suggested new competence	Suggested by N respondents	Agreement with inclusion N(%)of 27 respondents
ORGANISATION	7.8 Closes consultation appropriately	3	25(93)
RECORD KEEPING	8.6 Identification of the author and date of record	2	21(78)

Stage 2: In the Stage 2 questionnaire the 21 suggested rewordings and the original version of each item, one suggested amalgamation and two suggested additional items, were presented and participants asked for their opinion.

Of the 45 respondents in Stage 2, 27 completed the questionnaire (60%); 68% were male, 54% general practitioners and 61% undergraduate teachers. All the suggested changes were selected by a majority of respondents (Tables 2, 3 and 4) and 19 (70%) agreed with the deletion of item 4.8.

The validated version of GeCoS (9 categories and 59 component competencies) is shown at Appendix 1.

Discussion

What we found: We have modified the Calgary Cambridge guide to the medical interview as a consultation skills model by incorporating a 'Clinical reasoning' core which runs through the framework in parallel with the 'Organization' and 'Building and maintaining the relationship' pillars. We have developed an assessment framework (tool) from the LAP which maps onto the modified Calgary Cambridge guide through a rigorous initial development and piloting process and a multi-institutional and multi-speciality Delphi process and achieved consensus on the inclusion of all its elements. The level of agreement reached by stage 1 of the study was sufficient for GeCoS to satisfy the *a priori* consensus standards: all the broad headings and all their component competencies were considered "very relevant and succinct" or "relevant but needs minor alteration" by over 70% of respondents. Indeed, 91% of the elements were deemed relevant by over 90% of respondents. However, consideration of the free text suggestions has enabled us to further refine GeCoS through rewording and subsequently validating three of the broad category headings and 18 competencies.

Strengths and weaknesses: The initial development of GeCoS was rigorous with careful mapping of the competencies in the Calgary Cambridge guide and LAP to identify overlaps and gaps between each of them and Tomorrow's Doctors (2009) [1], a careful consensus between members, and initial piloting of the instrument before embarking on the Delphi study. The study used the same *a priori* definitions of consensus as previous studies. We took care to recruit the panel from a range of clinical specialties and Medical Schools which use one or neither of the

parent documents. The thematic analysis of the free text responses was similarly rigorous with each group of text being considered by pairs of the research team and the final decision reflecting the consensus of all. We remained open to further revision of the tool.

We set the limit for inclusion of any item in stage 2 at five (rounding down from 5.9 rather than up to six) similar responses and included any suggestion we felt represented an improvement. The Delphi method brings the advantages of obtaining a consensus from a panel of content experts whilst minimizing the influence of more forceful personalities [22]. The panel size was similar to that in other Delphi studies [7, 23] and the response rate was modest, but better than that in others [17, 19]. The stage 1 questionnaire was long, but despite this, 61% of respondents completed all 79 items in the survey and a median of eight free text comments were made about each item. We consider that this reflects a high level of engagement by respondents and that their responses are likely to have been considered.

Other literature: Variations of the Delphi method have been used previously for the identification and face validation of assessment criteria in health care [7, 17, 18, 19, 23] and other disciplines [24]. The LAP has been validated for teaching in general practice, but has never been formally validated for hospital teaching [7]. We have not been able to find another instrument which is designed for the assessment of generic (as opposed to context specific) consultation skills and is mapped to a clearly defined consultation skills curriculum.

How GeCoS can be used: GeCoS is now ready for use in formative and summative assessment of the consultation skills of medical students in any simulated or workplace, hospital or community clinical setting. Since it is generic, not all of its elements will be used in any one consultation. Some of the broad categories such as opening, building and maintaining the relationship, organization, record keeping and clinical reasoning will be pertinent to most consultations, even though not all the competencies within these categories will. The other categories (history, examination, management and case presentation) will not all be relevant to every consultation.

The GeCoS assessor judges which of the categories and components are relevant to each consultation and makes a global assessment of how the student responds to the specific challenge presented by the consultation in each category and, if desired, the case overall. Ideally, the assessor's judgment is made over a series of consultations so that all categories and most competencies are assessed. Assessment can result in both a global rating for each category of skills and also in a note of the specific skills which were done well or require improvement.

Being generic, GeCoS lends itself to providing a basis for the second stage of formative assessment, namely constructive feedback. We have also developed a GeCoS tutor / assessor support tool. This is a set of 'Strategies for Improvement' modeled on those for the Leicester Assessment Package [15] which contains suggested strategies for improvement of each of the GeCoS competencies. The assessor / teacher (and student) can use this to pick strategies which are likely to assist the student to develop the skills which s/he most needs to improve. A carefully worded "educational prescription" can be provided without the busy workplace-based assessor needing to re-think the wording of each piece of advice.

What next: Evaluation of the experiences of teachers and learners (and peer assessors) in using GeCoS will inform the refinement of the processes for formative and summative assessment. The development of software to support clinical teachers in formative assessment may be the next step in the development of GeCoS. A study of its reliability as an assessment instrument will be an important sequel.

Conclusion

We reviewed the Calgary Cambridge guide to the medical interview and the Leicester Assessment Package (LAP) and identified concepts common to both or only represented in one or the other. We revised the Calgary Cambridge guide to include concepts it did not contain ('Clinical reasoning', 'Management', 'Record keeping' and 'Case presentation') and populated it with competencies generated from the GMC's Tomorrow's Doctors guide, the LAP and the Calgary Cambridge guide. We validated this in a two-stage Delphi study across eight UK medical schools. The resulting instrument, the **Ge**neric **Co**nsultation **S**kills assessment framework (GeCoS), is ready for use in teaching, formative and summative assessment.

References

1. General Medical Council. (2009). Tomorrow's Doctors: Outcomes and Standards for undergraduate medical education. London, General Medical Council (GMC). Available at: http://www.gmc-uk.org/education/undergraduate/tomorrows_doctors_2009.asp [Accessed June 2010].

2. Norman G. (2002). Medical expertise and mashed potatoes. Medical Education. **36**(12):1167-1168.

3. Ericsson K A, Krampe R T, Tesch-Römer C. (1993). The role of deliberate practice in the acquisition of expert performance. Psychological Review. **100**(3):363-406.

4. Kurtz S, Silverman J, Draper J. (1998). Teaching and Learning Communication Skills in Medicine. Oxford, Radcliffe Medical Press Ltd.

5. Kurtz S M, Silverman J D. (1996). The Calgary-Cambridge Referenced Observation Guides: an aid to defining the curriculum and organizing the teaching in communication training programmes. Medical Education. **30**(2):83-89.

6. Fraser R C, McKinley R K, Mulholland H. (1994). Consultation competence in general practice: testing the reliability of the Leicester assessment package. British Journal of General Practice. **44**(384):293-296.

7. Fraser R C, McKinley R K, Mulholland H. (1994). Consultation competence in general practice: establishing the face validity of prioritized criteria in the Leicester assessment package. British Journal of General Practice. **44**(380):109-113.

8. Hastings A, McKinley R K, Fraser R C. (2006). Strengths and weaknesses in the consultation skills of senior medical students: identification, enhancement and curricular change. Medical Education. **40**(5):437-443.

9. McKinley R K, Fraser R C, van der Vleuten C, Hastings A M. (2000). Formative assessment of the consultation performance of medical students in the setting of general practice using a modified version of the Leicester Assessment Package. Medical Education. **34**(7):573-579.

10. Fraser R C, Sarkhou M E, McKinley R K, Van der Vleuten C. (2006). Regulatory end-point assessment of the consultation competence of family practice trainees in Kuwait. European Journal of General Practice. **12**(3):100-107.

11. McKinley R K, Middleton J, French A. (2004). A pilot study of diagnostic educational evaluation of junior general practice registrars. Education for Primary Care. **15**(2):242-250.

12. Fraser R C, Lee R S Y, Yui Y K, Lam C L K, McKinley R K, Van der Vleuten C P M. (2004). Regulatory Assessment of the Consultation Competence of Family Physicians in Hong Kong. The Hong Kong Practitioner. **26**(1):5-15.

13. Lau J K C, Fraser R C, Lam C L K. (2003). Establishing the content validity in Hong Kong of the prioritised criteria of consultation competence in the Leicester Assessment Package (LAP). The Hong Kong Practitioner. **25**(12):596-602.

14. McKinley R K, Dean P, Farooqi A. (2003). Reactions of volunteer general practitioners to educational assessment of their consultation performance: a qualitative study. Education for Primary Care. **14**(3):293-301.

15. Fraser R C. (1999). Clinical Method: A general practice approach, 3rd Edition. Oxford, Butterworth-Heinemann.

16. McKinley R K, Williams V, Stephenson C. (2010). Improving the content of feedback. The Clinical Teacher. **7**(3):161-166.

17. Redsell S A, Lennon M, Hastings A M, Fraser R C. (2004). Devising and establishing the face and content validity of explicit criteria of consultation competence for UK secondary care nurses. Nurse Education Today. **24**(3):180-187.

18. Redsell S A, Hastings A M, Cheater F M, Fraser R C. (2003). Devising and establishing the face and content validity of explicit criteria of consultation competence in UK primary care nurses. Nurse Education Today. **23**(4):299-306.

19. McKinley R K, Strand J, Gray T, Schuwirth L, Alun-Jones T, Miller H. (2008). Development of a tool to support holistic generic assessment of clinical procedure skills. Medical Education. **42**(6):619-627.

20. Mcllwaine L M, McAleer J P G, Ker J S. (2007). Assessment of final year medical students in a simulated ward: developing content validity for an assessment instrument. International Journal of Clinical Skills. **1**(1):33-35.

21. SurveyMonkey™. Available at: http://www.surveymonkey.com [Accessed June 2010].

22. Goodman C M. (1987). The Delphi technique: a critique. Journal of Advanced Nursing. **12**(6):729-734.

23. Baldwin P J, Paisley A M, Brown S P. (1999). Consultant surgeons' opinion of the skills required of basic surgical trainees. British Journal of Surgery. **86**(8):1078-1082.

24. Haslam I R. (1990). Expert assessment of the national coaching certification program (NCCP) theory component. Canadian Journal of Sport Sciences. **15**(3):201-212.

Generic Consultation Skills (GeCoS)

A FRAMEWORK FOR THE CONSULTATION INCORPORATING CONTENT AND PROCESS SKILLS

Organisation

CLINICAL REASONING

OPENING

HISTORY
PROCESS   CONTENT

EXAMINATION

MANAGEMENT
PROCESS   CONTENT

CLOSING

Building and maintaining the relationship

RECORD KEEPING
PROCESS   CONTENT

CASE PRESENTATION
PROCESS   CONTENT
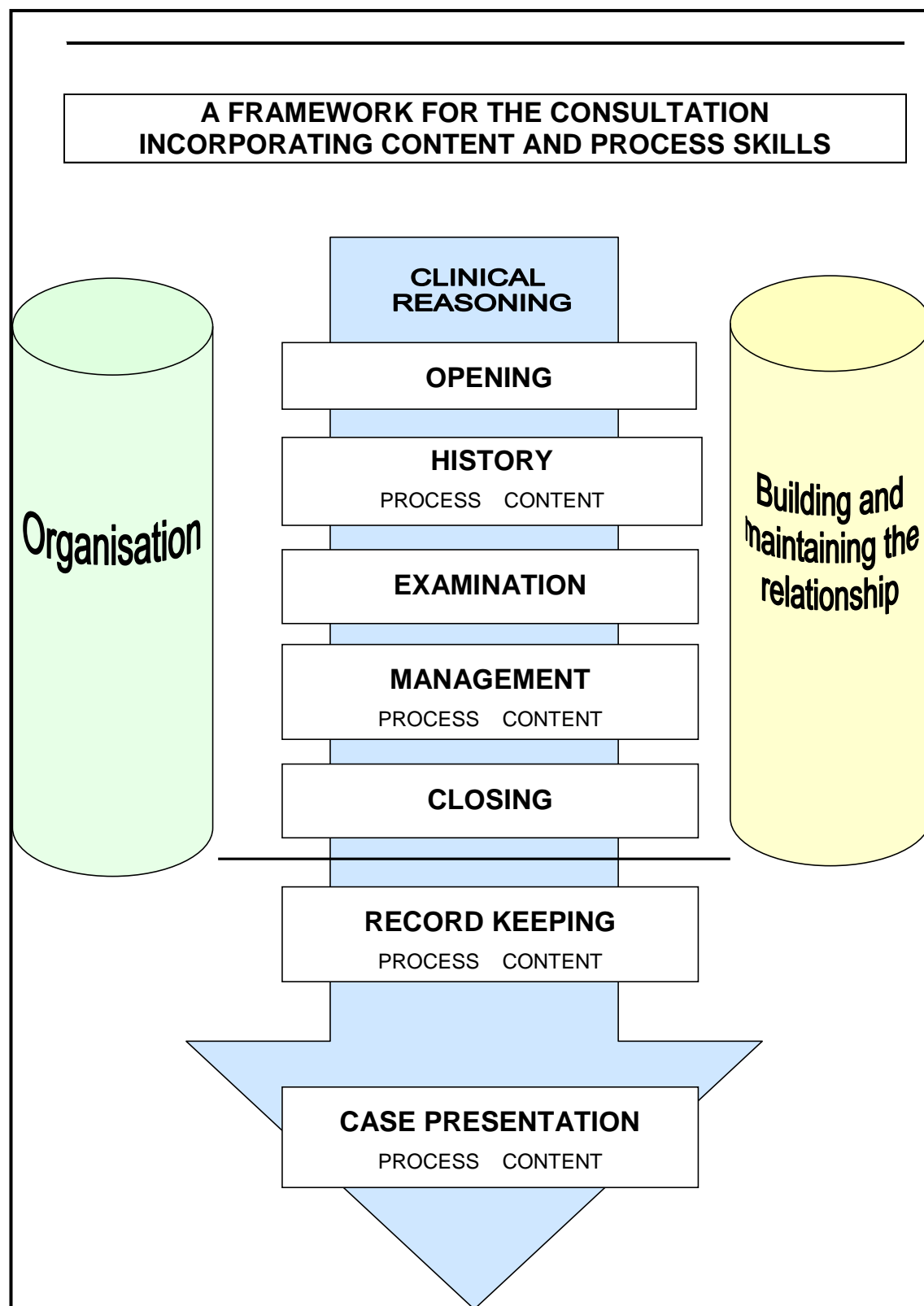
Generic Consultation Skills (GeCoS) - overview of skills to be assessed
Keele University School of Medicine

OPENING
- Introduces self
- Establishes identities of patient and third parties and preferred forms of address
- Establishes agendas

HISTORY
PROCESS
- Enables patient to fully elaborate presenting problem(s)
- Listens attentively
- Skilled use of questioning including open and closed questions
- Clarifies words used and/or symptoms presented by patient as appropriate
- Recognises and responds appropriately to verbal and non-verbal cues
CONTENT- obtains the following:
- Sequence of events
- Details of symptoms
- Effect on the patient's life
- Patient's ideas, concerns and expectations
- Relevant background information including: Past Medical, Drug, Family and Social History; Systems review; Factors influencing health

EXAMINATION
- Obtains and maintains consent
- Displays competent practice of infection control
- Displays sensitivity to patient's needs and dignity; offers chaperone if appropriate
- Gives clear instructions and explanations of process
- Performs examination competently
- Elicits normal and abnormal findings

MANAGEMENT
PROCESS
- Relates explanations to patient's perspective
- Gives clear information in small chunks
- Negotiates a mutually acceptable plan with patient and/or third parties
- Reassures appropriately
- Checks understanding
CONTENT
- Gives key evidence-based information
- Explores available options, risks and benefits
- Investigates appropriately
- Prescribes rationally and accurately
- Refers appropriately
- Makes appropriate use of opportunities for health promotion
- Agrees appropriate follow-up

CLINICAL REASONING
- Seeks relevant and specific information from patient's record or third parties
- Generates appropriate working diagnoses or problem list
- Seeks discriminating information from history, examination and investigations to help confirm or refute working diagnoses
- Correctly interprets information obtained
- Applies basic, behavioural and clinical sciences to solution of patient's problem
- Recognises limits of competence and acts accordingly

BUILDING AND MAINTAINING THE RELATIONSHIP
- Develops and maintains a professional relationship with patient
- Respects the patient's ideas, beliefs and autonomy
- Responds empathically
- Fosters collaboration

ORGANISATION
- Considers and optimises the setting
- Involves third parties appropriately
- Exhibits a well-organised approach to gathering and sharing of information
- Makes organisation of consultation overt to patient
- Prioritises agendas appropriately
- Summarises appropriately
- Uses time appropriately
- Closes consultation appropriately

RECORD KEEPING
PROCESS
- Makes concise and accurate notes without interfering with dialogue or rapport
MINIMUM CONTENT includes:
- Diagnoses/problems
- Relevant history and examination
- Outline of management plan; therapy, investigations, referral and follow up
- Information, instructions and special precautions given to the patient
- Identification of the author and date of record

CASE PRESENTATION
- Engages and orientates colleague
- Delivers clear and relevant detail in a logical order
- Communicates interpretation of data transparently
- Draws purposeful conclusion

Adapted from: the Calgary Cambridge Framework for the Medical Interview with the kind permission of Dr Jonathan Silverman, University of Cambridge; Fraser RC. *Clinical Method: a general practice approach. Third ed. Oxford Butterworth-Heinmann, 1999* and material provided by AM Hastings, Department of Medical and Social Care Education, University of Leicester.

Chapter 6

Development and face validation of strategies for improving consultation skills

Janet Lefroy

Adam Thomas

Chris Harrison

Stephen Williams

Fidelma O'Mahony

Simon Gay

Ruth Kinston

Robert McKinley

Affiliation: Keele University School of Medicine

Abstract

Context: While formative workplace based assessment can improve learners' skills, it often does not because the procedures used do not facilitate feedback which is sufficiently specific to scaffold improvement. Provision of pre-formulated strategies to address predicted learning needs has potential to improve the quality and automate the provision of written feedback.

Objectives: To systematically develop, validate and maximise the utility of a comprehensive list of strategies for improvement of consultation skills through a process involving both medical students and their clinical primary and secondary care tutors.

Methods: Modified Delphi study with tutors, modified nominal group study with students with moderation of outputs by consensus round table discussion by the authors.

Results: 35 hospital and 21 GP tutors participated in the Delphi study and contributed 153 new or modified strategies. After review of these and the 205 original strategies, 265 strategies entered the nominal group study to which 46 year 4 and 5 students contributed, resulting in the final list of 249 validated strategies.

Conclusions: We have developed a valid and comprehensive set of strategies which are considered useful by medical students. This list can be immediately applied by any school which uses the Calgary Cambridge Framework to inform the content of formative feedback on consultation skills. We consider that the list could also be mapped to alternative skills frameworks and so be utilised by schools which do not use the Calgary Cambridge Framework.

Introduction and need for this study

A key set of skills for all medical practitioners is being able to talk to patients who present with medical concerns, examine them, decide on the likely diagnosis and negotiate an agreed plan of investigation and treatment: in short, being able to conduct a 'consultation' with a patient. We usually teach these skills separately but expect learners to hone and integrate them by conducting multiple consultations in the workplace. To learn from this activity learners need to reflect on their performance and receive external feedback upon it. Both of these tasks, self-reflection and external feedback, can be helped by a comprehensive conceptual framework or scaffold which encompasses the tasks required so long as they can be readily understood and held in mind by both learners and tutors. Having an accepted scaffold for consultation skills allows much more specific reflection and advice regarding how elements of the tasks can be achieved.

It is clear that assessment of workplace-based activity can improve a learner's consultation skills[1] and improvement is predominantly mediated by discussion between learner and tutor.[2] Whether

this discussion is called feed-forward or feedback, its goal should be to provide the learner with "specific information about the comparison between the learner's observed performance and a standard, given with the intent to improve the learner's performance."[3] Feedback needs to provide the learner with guidance which is specific enough to enable the learner to enhance achievement.[4] At its simplest, feedback provides verification as to whether an observed performance was optimal or not. However, if tutors wish to help learners to improve their performance, learners will require further explanation as to why the observed performances were or were not optimal. Furthermore, explanation is insufficient on its own: additional structure or scaffolding is needed to support learners to make use of feedback by providing more explicit instructions and strategies to remediate the sub-optimal performances. This often requires that tasks are deconstructed to make them more achievable[5] and to set specific goals. In addition, tutors need to set their learners an appropriate and realistic level of challenge.[6] By doing so, the gap between observed and intended performance can be narrowed.

In practice, current workplace assessments can result in a greater focus on assessment than feedback with the design of the forms often contributing to limited feedback.[1 1] Learners who perceive that they have passed an assessment may feel little incentive to make use of feedback that is available.[7] There is often a failure to set specific learning goals.[8] Tutors often have different concepts of the standard required and therefore are likely to have different notions of the level of challenge that should be set by a learning goal.[9] We have observed that while tutors can readily identify what was done well and what learners need to improve, the feedback given is often non-specific and does not assist learners to address the deficit between the observed and desired performance.[10]

Tutors are often discouraged from providing appropriate feedback because of time constraints.[8] One way to reduce time pressure is to assist tutors to scaffold learners' learning when sub-optimal performances have been observed. This can be done by providing pre-written strategies for use in the event of various possible predicted sub-optimal performances. This has potential additional benefits such as acquainting tutors with the standard required; the use of terms familiar to the learners from their curriculum; and helping tutors to set appropriate challenges for their learners. We now report a study in which we aimed to systematically develop, validate and maximise the utility of a comprehensive list of strategies for improvement through a process involving both medical students and their clinical tutors from primary and secondary care.

Methodology

Context: Serial workplace-based assessment is a key component of the consultation skills development programme for students at Keele University School of Medicine. This is an integrated spiral programme in primary and secondary care running from year 3 to year 5. This uses the **Ge**neric **Co**nsultation **S**kills (GeCoS)[11] assessment tool which was designed to support the formative and summative assessment of consultation skills in various settings. GeCoS has been developed by a group of clinical tutors in Keele from the Leicester Assessment Package (LAP),[12;13]

the Calgary Cambridge Guide to the Medical Interview[14] and Tomorrow's Doctors 2009,[15] and validated by a panel drawn from hospital and GP tutors, half from Keele and half from seven other UK medical schools.[11] It can be seen in Appendix 1.

Given the problem of lack of specificity of the content of feedback,[10] we had developed support materials for tutors and students using GeCoS. This contained up to eight pre-formulated strategies for improvement for each GeCoS competency to help tutors scaffold learning through specific feedback. Our intention was to enhance the educational potential of the feedback by:

a) Addressing the 'specificity gap' in feedback by providing the tutor who has identified a need for improvement with a set of specific strategies (or scaffolding) for improvement from which the tutor can suggest those s/he feels will help the learner.[10]

b) Saving time as the tutor doesn't need to "re-invent the wheel".

c) Offering the potential to 'automate' the production of written feedback. After the discussion of how to improve, the selected strategies plus any other tutor comment can be recorded for the learner. To save the tutor time writing, the strategies for improvement have been included in an online version of the workplace-based assessment feedback form with text boxes for additional comment. This online workplace-based assessment tool generates an email to both the tutor and the learner containing a written summary of the discussion between them. We call this email the "educational prescription".

d) Enabling learners who have self-identified a need to improve a particular competency with suitable strategies to do so.

We developed the current set of strategies from the LAP strategies,[16] amendments and additions being informed by our course materials which were based on the Calgary-Cambridge model[14]. An example of a competency and its corresponding strategies for improvement is shown in figure 1.

Table 1 maps the number of strategies originally developed for each domain of consultation skills. Although GeCoS had been validated, the set of strategies for improvement had not and we considered that they needed improvement and validation by both teachers and students.

Operational definitions: We defined a valid strategy as one which if both understood and then adopted by students, should result in improvement in the relevant competence. We defined comprehensiveness as including all distinct strategies suggested by teachers and students during this study and during the workplace-based assessment of students in the first years of its use.

Basis of judgments: The judgment of whether adopting the competence should enhance performance would be made by hospital and general practice tutors (because the support materials would be used in both hospitals and general practices) and by students (who would be asked to adopt them). The judgment of whether strategies were understandable would be made by students who have to be able to understand them to adopt them. The judgment of whether strategies were distinct was made by the author group which included hospital and general practice tutors and a student.

Fig 1: An example of one GeCoS competency and its corresponding strategies for improvement (validated in this study)

Competency (from the History Process and Content domain of consultation skills)

Enable the patient to fully elaborate presenting problem

Strategies for improvement of this competency (one or more selected depending on what the student had done)

- Resist the temptation to interrupt at the start of the consultation, although this may be necessary later if the patient becomes repetitive.

- Use open questions to begin with e.g. 'How did it start?'; 'What happened next?'

- Use prompts as appropriate e.g. 'I see'; 'Tell me more about that'

- If the patient makes a significant statement and then stops, encourage the patient to continue, for example by repeating the last statement or word.

- List the symptoms so far and check in different ways, for example, 'Was there anything else you noticed?', 'Were there any other symptoms?'

Table 1:   The original strategies for improving competencies in GeCoS mapped to the nine domains of the consultation (categories of skills)

Category of consultation skills	No. of competencies	No. of strategies for improvement
Opening	3	12
History process + content	10	35
Examination	6	23
Management	12	47
Clinical reasoning	6	29
Building and maintaining the relationship	4	10
Organisation	8	22
Record keeping	6	13
Case presentation	4	14
Total	**59**	**205**

We therefore had to access the views of a broad range of experienced general practice and hospital tutors and of medical students. We judged that a single methodology would not suffice. We used a modified Delphi study[17] questionnaire administered by a commercial survey web site (SurveyMonkey™) to access the views of a broad cross section of teachers; methods which the authors[11;13;18] and others have used previously.[11;19-21] We considered that an electronic survey would not adequately access the views of students: we wished to access not only their candid views on the usefulness of the strategies but also why they thought those which were not useful were not and how they could be improved. We therefore used a modified nominal group technique[22] which allowed us to assess the discussion between students about the strategies but modified voting by maintaining anonymity using an audience response system which reflects aspects of Dephi methods.[17] We also considered that student participants would be more able to contribute if they had experience of workplace-based assessment and receiving written summaries containing a mix of the original and novel strategies. We therefore recruited students from years 4 and 5 of our undergraduate course as they had this experience.

Methods

Delphi study of tutors: Tutors were asked to express their opinions of the usefulness of each of the original strategies for formative feedback following workplace-based assessment of an undergraduate medical student. The questionnaire used a four point Likert response scale anchored by "Useful" (1) and "Not useful" (4). As well as alternative wordings, they were also asked to offer any additional strategies for improvement they had found helpful. The instructions to the respondents are reproduced in Figure 2.

Fig 2: Instructions for respondents to Tutor Delphi study

What we want is your opinion on:

1. How useful would each strategy be for your students?

    a. If you wish to suggest a rewording of any of the strategies to improve them, please use the text box.

    b. If you have additional strategies to suggest for students to improve any of the competencies in this domain, please use the text box. Put several items in one box if you wish.

Because of the number of strategies (205, see table 1), we developed nine online surveys, one for each GeCoS domain. Participants were asked to complete two surveys but were also sent the web addresses for the other seven so that they could respond to those in which they were interested. The panel of participants was drawn from hospital and general practice clinical tutors and examiners in order to include experts in all types of consultations. We continued to recruit participant tutors until we had 10 responses for each domain, aiming for half GP and half hospital tutors.

Data capture, analysis and synthesis: All categorical and free text data from the tutor Delphi were downloaded. To this we added rewordings and distinct new strategies from the written summaries of actual workplace-based assessments ("educational prescriptions") of the cohort of year 3 students from May 2010 through to the end of their 4th year in June 2011.

Analysis was by a round table group of the authors. Each meeting consisted of at least two hospital doctors and at least two GPs to bring both perspectives to the data.

All suggestions for rewording and new strategies were reviewed using previously described methods[11] to decide whether:

- Reworded strategies were an improvement on the original strategy.

- Suggested new strategies were indeed new rather than being rephrased existing strategies.

The usefulness ratings of each strategy provided the researchers with an indication of how important it was to seek an improved wording. The working protocol was that a rewording was sought for any strategy with a "Usefulness" rating of more than 2.0 in the Delphi study (2.5 is the mid-point of the "Useful" to "Not useful" 1 to 4 scale). The rewording was informed by the suggestions of the surveyed tutors and those identified from educational prescriptions as mentioned above.

We were also careful in preparing the list of strategies for validation by the student participants to ensure comprehensiveness. To this end we were mindful of the need to present a variety of unique suggestions for every GeCoS competence.

Modified nominal group study with students: The outcome of the tutor study was a new (expanded) list of strategies which entered the student study for validation. Both the original and reworded versions were presented to the student group.

Student participants were recruited by AT (a year 4 medical student). To accommodate the number of strategies, multiple groups were run. The methodology combined elements of Delphi consensus[17] and nominal group technique. The groups were facilitated by AT and the discussions were recorded, either by a 'scribe' or by digital tape recorder, and transcribed. Each group session lasted 90 minutes during which approximately 30 strategies were considered.

Group process: Each strategy was presented for validation to one student group only, but if any changes were made it was then presented to a second group after revision.

The strategy was presented using a PowerPoint presentation with the following wording:

> "If you need to improve on competency <<*Selected GeCoS competency>>*
>
> Would the following be a useful strategy for you?
>
> <<*Selected strategy for improvement of that competency>>*"

The student facilitator first checked "Do you understand the competency?" to enable clarification if necessary. The group then voted on the utility of the strategy for improvement using electronic keypad devices.

Each individual voted "Useful" or "Not useful" before discussion of each strategy. "Useful" was defined as the strategy being both clear (the student could understand what was suggested) and relevant (the student could envisage themselves or others improving by doing what was suggested).

The definitions of student consensus were:

- If 70% or more responded "Useful" then the item was "Validated".
- If more than 30% and less than 70% responded "Useful" the item was "Not validated".
- If fewer than 30% responded "Useful" the item was "Rejected".

Students who responded "Not useful" were asked whether it was because the strategy lacked relevance or clarity and why they considered this to be the case. These reasons were captured and transcribed to inform the rewording of the competence by the subsequent "round table" of the authors.

Following the vote on the usefulness of each original strategy, any possible rewordings from the tutor study were shown on a second slide. The student group then voted on which was the preferred wording.

The round table of the authors (including AT to ensure that the students' comments were represented) met to revise the strategies following the student validation exercise.

All "Rejected" strategies were discarded unless the round table felt they could improve the wording and all "Not validated" strategies were returned to a final student group whether or not they could be improved. All "Validated" strategies about which students made comments about improvement were also reconsidered by the round table of the authors and returned to the final student group if potential improvements were made.

All strategies reconsidered by the final student group were validated if they achieved more than a 70% "Useful" vote and rejected if not. Nevertheless, the "round table" (including AT) retained the final editorial control to ensure that there was choice of strategies available for each GeCoS competence.

Results

Figure 3 is a flow chart of the results which may help understanding of the process.

Fig 3: Flowchart of the study process

205 original strategies

Workplace-Based Assessments - GP tutors produced potential new strategies

| 24 revisions | 16 new |

Delphi exercise with hospital and GP tutors

| 85 revisions | Usefulness score for 204 of the 205 | 68 new |

Round table of researchers filtered all tutor suggestions

| 76 revisions | 64 new | 1 original omitted |

Student nominal groups voted on usefulness of 265 strategies with 76 revisions as alternatives, 55 of which were preferred. Advised on wording

| 190 validated | 42 not validated | 33 rejected |

Round table of researchers

| **154 accepted** | 35 revisions, 3 added | 1 omitted | 34 revisions | 8 omitted | 22 revisions | 11 omitted |

Student nominal group voted on usefulness of all but 5 of 94 revised /added strategies , plus 3 missed in round one = 92. Advised on wording

| 77 validated | 14 not validated | 1 rejected |

Round table of researchers

| **77 included**, 3 further revisions | **12 included**, 5 further revisions | 2 omitted | **1 revised and included** | **5 carried over** |

Final revised and validated set of 249 strategies

In 984 educational prescriptions for 128 students from May 2010 to June 2011, GP tutors had used 200 of the 205 strategies but also had offered alternative wordings and new strategies.

These were reviewed by one author (JL) and from these, 24 amendments of existing strategies and 16 new strategies were identified for this study.

In the tutor Delphi study 35 hospital tutors (22 male/13 female; 21 consultants/14 juniors all with some experience of teaching covering surgical, emergency and general medicine, paediatrics, psychiatry anaesthetics and radiology) and 21 experienced GP tutors (10 male/11 female) generated 250 comments containing 84 potential amendments and 68 potential additions to the original strategies.

These were all considered at three round table meetings attended by between five and eight members of the research team. This resulted in a new list of 268 strategies.

Seven student groups of between five and eight participants voted on and then discussed each strategy. A total of 41 students were involved at this stage; 16 male and 25 female; 33 from year 5 and 8 from year 4 with five students attending two groups. Although 190 of 265 strategies were validated on their first consideration by these groups, the student discussion suggested that 35 of the validated strategies could be further improved. 33 strategies were rejected (received <30% approval) and 42 were not validated (received 30-70% approval). 3 were inadvertently omitted at this stage and were therefore included in the second stage nominal group instead.

A further five round table meetings of the researchers including AT reviewed the outcome of the student groups. The round table omitted 11 of the rejected strategies together with 8 strategies which had not been validated and 1 validated strategy which was considered to be a duplicate. The remaining 22 strategies which had been rejected by the students plus the remaining 34 non-validated strategies and the validated strategies for which improvements had been suggested were revised by the round table of researchers being guided by the students' comments. Three strategies were added at this stage by suggested splitting of strategies. A total of 97 strategies (including the three missed from the first nominal group stage) were prepared for the final student group.

Ten Year 5 students attended the final group, five male and five female, of whom five had attended a previous group, bringing the total number of students involved to 46. 77 strategies were validated by this final group. 14 strategies were not validated by that group, and the researchers therefore had to determine what was to be done with them. Seven of these had been validated but refined by the previous group. For six the previously validated version was included. One amended version was included as it was preferred by the students. Five were included even though not validated by either group, but were further amended guided by student comment. Two strategies were considered not worth including or amending and were dropped. One strategy was rejected by the final student group but they suggested improvements. On discussion, the researchers determined to revise this according to student comment and included it in the final set of strategies. Five other revised strategies which had slipped the attention of the final student group were also considered worth including by the round table. Four of these had been previously validated.

The final 249 strategies modified and face-validated by the two-stage student group process can be seen in Appendix 2.

An example of the changes made to one strategy by this process is shown in Fig 4.

Fig 4:   Examples of the journeys of two strategies through this process

---

Example 1: To improve on the clinical reasoning competency: *Correctly interprets information obtained*

One original strategy: *Avoid over-reliance on features that may support a conclusion you reached prematurely*

Scored 1.67 on the usefulness scale in the tutor study (1 = useful; 4 = not useful) but was revised according to tutor suggestions to:

If there appears to be an obvious diagnosis, consider alternatives

Which was voted useful by 100% of the student group and preferred by 83%

Example 2: To improve on the history taking competency: *Skilled use of questioning including use of open and closed questions*

One original strategy: *Use facilitation to encourage the patient to tell their story*

Scored 2.0 on the usefulness scale in the tutor study (1 = useful; 4 = not useful)
Was rejected by the student group (0% useful) and revised according to student comment to:

Encourage the patient to tell their story by using expressions like "And then….?" Or "What happened next?"

Which was voted useful by 78% of the final group

Discussion

Two major challenges in providing useful feedback to trainee clinicians have been lack of specificity in the information and advice given and lack of time in which to give it. A step towards making workplace-based assessment and feedback more effective could be the provision of support materials for those giving feedback. To address the specificity gap and the time challenge, such materials would contain accepted suggestions for improving skills and should save time taken in constructing the advice given.

In this study, we have developed a valid and comprehensive set of strategies to inform the content of formative feedback on consultation skills and the strategies are considered useful by the medical student recipients of such feedback. We are only aware of one similar set of strategies for enhancement of generic consultation skills (which informed the development of our original strategies)[16] and one for the enhancement of clinical procedural skills[10] and neither has gone through a process of formal validation by tutors and students nor a review of 'usefulness' by students. Thus we believe that this is a unique resource for students and teachers which was developed using a novel combination of methods and participants.

The major strength of this study is that a total of 110 stakeholders with experience of workplace-based assessment contributed: 46 students who had had a minimum of six workplace-based assessments each, 56 experienced clinical teachers all of whom had conducted workplace-based assessments and the 8 strong research team, representing the major stakeholders in workplace-based assessment and formative feedback (students, clinical teachers and the core school staff). In order to enhance its utility for workplace-based assessment in both hospital and general practice and enhance its accessibility to students for self-assessment and feedback, it was desirable that clinical tutors from both hospital and general practice were involved in its development, that medical students should validate those strategies which were useful to them and that they should also suggest improvements. In addition, the methodologies were carefully chosen. The modified Delphi tutor study enabled a broad consensus to be achieved across the hospital and GP tutor base with important refinements on advice to give to medical students who need to improve specific consultation skills. The modified nominal group student study involving the final consumers, medical students, enabled both validation of the strategies as 'fit for purpose' and, by capturing the discussion, the final refinement of the strategies incorporating the student perspective. Furthermore, even if students had validated a strategy but had suggested ways in which it could be improved the 'round table' then attempted to do so. We consider that our refinement of the nominal group by anonymising voting through use of 'audience response' electronic key pads, so that students voted 'blind' to each other's opinions, was a particular strength.

A clear limitation is that this was a study conducted in a single school and with a particular consultation skill curriculum and assessment framework albeit that GeCoS has been validated in a multi-speciality, multi-school study. Furthermore the consensus on individual strategies was derived from small groups and a different student group might have reached a different consensus. Nevertheless few items were rejected as a result of a single student group without being reconsidered by the researchers and by the final group and we consider our methodology to have been a pragmatic solution to the problems this project presented.

This set of strategies for improving each skill required by tomorrow's doctor can be considered a 'primer' for consultation skills. We are using them throughout our curriculum from the skills lab in year 1 to final year workplace-based assessment and in all clinical settings from hospital to general practice in which our students learn. Our tutors now conduct approximately 1,400 formal

workplace based assessments incorporating these strategies on our students each year with very positive student evaluations. Having developed the tutor support materials, we will next improve and evaluate the electronic feedback interface to make them easier and quicker to use in any clinical setting. Studies of the effect of such feedback on student learning will be the final test of the utility of this teaching resource. One such study is planned in which the use made by students of the various sections of their web-based feedback portal will be evaluated. A study of sequential assessments and the feedback given would be another way of assessing the impact of this feedback although, because this is a complex system, this would be insufficient to attribute change exclusively to use of the strategies.

The wider application of this study is that any medical school with consultation skills curricula based on the Calgary-Cambridge framework for effective consultations can immediately take any or all of these strategies and map them to their own consultation skills assessments to inform the content of either verbal or written feedback. We offer them as a resource to others interested in supporting their tutors to improve the quality of the feedback they give their students.

Ethics: The study protocol was approved by the Keele University School of Medicine Research Ethics Committee.

Reference List

(1) Norcini J, Burch V (2007): 29(9):855-871. Workplace-based assessment as an educational tool: AMEE Guide No. 31. *Med Teach*.

(2) Overeem K (2010): 44(12):1159-1161. 'Paying it forward': performance improvement through feedforward interviews. *Med Educ*

(3) Van De Ridder JMM, Stokking KM, McGaghie WC, Ten Cate OTJ (2008): 42(2):189-197. What is feedback in clinical education? *Med Educ*

(4) Shute VJ (2008): 78(1):153-189. Focus on Formative Feedback. *Review of Educational Research*

(5) Archer JC (2010): 44(1):101-108. State of the science in health professional education: effective feedback. *Med Educ*.

(6) Hattie J, Timperley H (2007): 77(1):81-112. The Power of Feedback. *Review of Educational Research*.

(7) Harrison CJ, Konings KD, Molyneux A, Schuwirth LW, Wass V, Van der Vleuten CP (2013): 47(7):734-744. Web-based feedback after summative assessment: how do students engage? *Med Educ*

(8) Holmboe ES, Yepes M, Williams F, Huot SJ (2004): 19(5p2):558-561. Feedback and the Mini Clinical Evaluation Exercise. *Journal of General Internal Medicine*.

(9) Boulet JR, McKinley DW, Norcini JJ, Whelan GP (2002): 7(2):85-97. Assessing the Comparability of Standardized Patient and Physician Evaluations of Clinical Skills. *Adv Health Sci Educ Theory Pract* .

(10) McKinley RK, Williams V, Stephenson C (2010): 7:161-166. Improving the content of feedback. *The Clinical Teacher*.

(11) Lefroy J, Gay SP, Gibson S, Williams S, McKinley RK (2012): 5:115-125. Development and face validation of an instrument to assess and improve clinical consultation skills. *Int J Clinical Skills*.

(12) Fraser RC, McKinley RK, Mulholland H (1994): 44:293-296. Consultation competence in general practice: testing the reliability of the Leicester assessment package. *Br J Gen Pract*.

(13) Fraser RC, McKinley RK, Mulholland H (1994): 44:109-113. Consultation competence in general practice: establishing the face validity of prioritized criteria in the Leicester assessment package. *Br J Gen Pract* .

(14) Kurtz SM, Silverman J, Draper J (2005):Teaching and learning communication skills in medicine. Radcliffe.

(15) General Medical Council. Tomorrow's Doctors (2009): http://www gmc-uk org/education/undergraduate/tomorrows_doctors_2009 asp [2009 [cited 2009 Oct. 1]

(16) Fraser RC (1999): Clinical Method: a general practice approach. 3rd ed. Oxford: Butterworth-Heinemann.

(17) Goodman CM (1987): 12(6):729-734.The Delphi technique: a critique. *J Adv Nurs.*

(18) McKinley RK, Strand J, Gray T, Schuwirth L, Alun-Jones T, Miller H (2008): 42(6):619-627. Development of a tool to support holistic generic assessment of clinical procedure skills. *Med Educ* .

(19) Redsell SA, Lennon M, Hastings AM, Fraser RC (2004): 24(3):180-187. Devising and establishing the face and content validity of explicit criteria of consultation competence for UK secondary care nurses. *Nurse Educ Today* .

(20) Redsell SA, Hastings AM, Cheater FM, Fraser RC (2003): 23(4):299-306. Devising and establishing the face and content validity of explicit criteria of consultation competence in UK primary care nurses. *Nurse Educ Today.*

(21) Baldwin PJ, Paisley AM, Brown SP (1999): 86(8):1078-1082. Consultant surgeons' opinion of the skills required of basic surgical trainees. *Br J Surg*.

(22) Cantrill JA, Sibbald.B., Buetow.B (1996): 4:67-74. The Delphi and nominal group techniques in health services research. *International Journal of Pharmacy Practice.*

Authors' contributions:

The study was conceived by JL and RKM (both GPs) and the study protocol development was led by JL and RKM with input from all authors. Hospital tutors were recruited by FO'M (a gynaecologist and obstetrician), SW (a paediatrician) and RK (an emergency physician). General practitioners were recruited by JL and SPG (a GP). Medical student participants were recruited by AT (a medical student). All authors contributed to the round-table discussions. The first draft of the paper was prepared by JL and RKM with extensive input by SW and CH (a GP) and additional input from all authors to subsequent drafts. All authors have approved the final version. JL and RKM are guarantors for the study.

APPENDIX	
Category and competency	**Strategy**

Opening

OP1: Introduces self

1. If you are unknown to anyone in the consultation, introduce yourself professionally using your name and role.
2. Even if the patient launches in as soon as they are through the door it is worth saying who you are.

OP2: Establishes identities of patient and third parties and preferred forms of address

1. Check the identity of the patient against the name of the person you expect to see.
2. Ask accompanying people their names and relationship with the patient.
3. Ask those attending 'What would you like me to call you?'
4. Introduce any other people (health staff, students) in the room and check that it is acceptable for the patient for them to be there.
5. Check the pronunciation of unfamiliar names with the patient.
6. Use your judgement to decide what is appropriate. The default strategy is to start formal with an older person (than yourself) and to consider what will feel appropriate for a younger person.

OP3: Establishes agendas

1. Identify the patient's agenda. Develop a range of opening questions for different situations with which you are comfortable
2. Check that your understanding of the patient's agenda is complete: 'Is there anything else you would like me to do today'
3. Clarify your agenda for the patient: 'I understand that you have come because/for XX'
4. Consider all presenting complaints and then quickly prioritise them and pay attention to what is necessary. Involve the patient in prioritising 'What is the most important thing to deal with today?'
5. Explain your agenda if you are a learner, and seek consent for this 'I am learning how to consult with patients. Could I interview you before you see Dr X and I will then report to her and we will complete the consultation together?'
6. Although confidentiality may be assumed in a healthcare consultation, consider whether it would help to make it explicit in this consultation
7. Recognise that it may not be possible to sort out all the problems presented on that day

History Process

H1: Enable the patient to fully elaborate presenting problem

1. Resist the temptation to interrupt at the start of the consultation, although this may be necessary later if the patient becomes repetitive.
2. Use open questions to begin with e.g. 'How did it start?'; 'What happened next?'
3. Use prompts as appropriate e.g. 'I see'; 'Tell me more about that'.
4. If the patient makes a significant statement and then stops, encourage the patient to continue, for example by repeating the last statement or word
5. List the symptoms so far and check in different ways, for example "was there anything else you noticed?", "were there any other symptoms?"

H2: Listens attentively

1. Demonstrate to the patient that you are listening by using appropriate body language and maintaining eye contact.
2. In a patient-centred consultation you will receive information out of sequence. Remember key points. For example: 'You said earlier you are a smoker, how much do you smoke?' is preferable to asking the same patient 'Do you smoke'.
3. If you need to write information down, or record data on the computer, do so in a way that does not interfere with your communication with the patient.
4. Don't stop listening to the patient whilst you think about the next question to ask. Use other strategies if you need time to think eg mini summary
5. If you need time to think, tell the patient that you are gathering your thoughts. Make some brief notes if necessary.

H3: Skilled use of questioning including open and closed questions

1. Move from open to closed questions e.g. 'Why have you come today?' 'Can you tell me more about that?', 'Is it getting worse?'
3. Encourage the patient to tell their story by using expressions like 'And then….?' or 'What happened next?'
4. Avoid using 'leading' questions, i.e. those that imply a particular answer e.g. 'Your baby doesn't have diarrhoea does he?'
5. Don't use 'double' or 'nested' questions e.g. 'What is your pain like and how long have you had it?' 'Is your appetite normal and have you lost weight?'
6. Tailor the questions you ask to the level of the patient's ability to understand. Don't patronise or talk down to the patient.
7. Don't use technical jargon.
8. It may be that you have to ask the same question again or in a different way if the patient has misunderstood or evaded answering. Don't be afraid to do that or you will be left feeling unclear

H4: Clarifies words used and/or symptoms presented by patient as appropriate

1. If you don't understand what the patient means, ask them to explain.
2. If the patient uses a medical or technical term (e.g. constipation) make sure you understand exactly what they mean by it.

H5: Recognises and responds appropriately to verbal and non-verbal cues

1. Listen carefully for and follow up all cues that the patient gives you e.g. 'My husband's at home all day now'
2. Notice unusual words and/or surprising omissions and follow up on these.
3. If a symptom is shown during the consultation, consider acknowledging it and ask whether it is typical (eg if patient coughs or has a tremor).
4. Acknowledge patients' expressed feelings to give them a chance to explain them or feel that they have been shared. e.g. 'I can see that this is difficult for you to talk about . . .'
5. If the patient is having difficulty telling the story or is distressed, allow time for the patient to regain composure.
6. Try to tolerate the discomfort of appropriate silences. Resist the temptation to talk when the patient is thinking about their response.
7. If the patient seems particulary uncomfortable, pause to assure the patient of the confidentiality of the interview and check whether the patient is happy to continue with the topic
8. Be sensitive to behaviour that is incongruous e.g. the patient who laughs when stating something serious.

H6: History Content - Sequence of events

1. Ask the patient to describe and clarify when and in which order each event occurred
2. If a patient appears to have skipped a period of time and you are aiming to determine a chronology, ask what happened in the relevant period

3   If a patient gives a jumbled sequence of events, repeat the sequence in the order you understand for confirmation

4   Always check when they were last well or when their new symptom(s) FIRST started

5   If the patient has had symptoms for some time, find out why the patient has presented now?

6   If the patient has difficulty ask 'Can you tell me about it from the beginning?' and follow up with 'What happened next?' until the story is complete

H7: History content - Details of symptoms

1   Allow the patient to finish their opening statement and clarify their presenting complaint(s) before you seek relevant associated symptoms

2   Use a mental checklist such as SOCRATES (which is useful for many symptoms) to clarify the presenting complaint(s)

3   Ensure you have checked whether or not the patient is experiencing the 'cardinal' symptoms for relevant system(s)

4   Note taking can help you to keep track of disordered information.

H8: History content - Effect on the patient's life

1   Ask the patient how his/her ability to sleep, toilet, wash, dress, cook, eat, work, relax or socialise (as appropriate) have been affected.

2   In consultations with a third party, ask about effects on the patient's behaviour if appropriate

3   Ask the patient 'how is this affecting you? How is it affecting others?'

4   Ask the patient "what does it stop you doing?"

H9: History content - Patient's ideas, concerns and expectations

1   In every consultation you must be satisfied that you know: What does the patient believe is wrong? What are they concerned about? What do they hope can be done? Sometimes this may require gentle but persistent questioning.

2   If the patient has indicated their ideas, concerns or expectations avoid direct questions. It is better to reflect back a remark they have made. E.g. 'You said your mother had headaches like these, what was the cause of her headaches?'

H10: History content - Background information including: Past Medical, Family and Social History; Systems review; Factors influencing health

1   Remember that a problem will often have physical, psychological and social components ('Triple Diagnosis').

2   Patients with psychological illness may have unrecognised physical disease, and vice versa, so ensure you have thought about this possibility.

3   When satisfied that physical disease is present always consider its impact on the social and psychological well-being of the patient.

4   Consider the impact on the patient of other social and psychological factors in their life such as their work, housing, family and other relationships, personality, sexuality, cultural background, spiritual beliefs and practices.

Examination

E1: Examination - Obtains maintains consent

1   Ask the patient's permission to carry out the examination, especially 'intimate' examinations

2   Check whether the patient has understood and has any questions before you proceed

3   If the patient is unable to give consent (lacks competency eg. a young child or confused adult) you must act in their best interests. At all times try to achieve their cooperation, with the help of a familiar person if appropriate.

4   If the examination is uncomfortable at any point, apologise and ask for permission to continue

E2: Displays competent practice of infection prevention

1. You must always cleanse your hands before (for the patient) and after (for yourself)
2. You should wear non-sterile gloves in examinations which might involve contact with body fluids. You should wear sterile gloves where the patient could be at risk of transmitted infection from your skin.
3. Dispose appropriately of gloves, apron, tissues etc. according to your workplace policy

E3: Displays sensitivity to patient's needs and dignity; offers chaperone if appropriate

1. Ensure a chaperone is available for intimate examinations and explain the need for this to the patient
2. When a chaperone is required either by the patient or by yourself, the chaperone should be acceptable to the patient
3. Give the patient privacy to undress and dress where possible
4. If the patient has difficulty in positioning or undressing themselves, ask whether you or the chaperone can help
5. Expose the part(s) to be examined with due sensitivity to the patient's dignity and cover them as soon as possible

E4: Gives clear instructions and explanations of process

1. Explain clearly to the patient what you want them to do. Demonstrate the required action if appropriate.
2. Give an explanation of what you are doing to the patient, particularly if this might involve discomfort.
3. Explain in terms the patient can understand
4. Explain to the patient that you will wait until they are dressed, settled, and ready to discuss your findings

E5: Performs examination competently

1. Review the examination in the textbook and/or watch a competent practitioner perform the examination
2. Be familiar with the instruments you use, first practising under supervision
3. Set the situation up to maximise your chances of success e.g. light from the side (JVP), low light levels (fundoscopy), correct side for your examination (apex beat).
4. Ensure the comfort of the patient before proceeding with an examination.
5. Ask the patient to point to the pain (if they have any).
6. Aim to do the examination once, correctly, and as fluently as possible
7. Watch the patient for signs of discomfort and respond accordingly
8. Find a competent colleague who is willing to observe you performing the examination and to give you feedback

E6: Elicits normal and abnormal findings

1. Obtain repetitive practice with feedback from a competent colleague who can tell you what signs they are finding
2. Never be afraid to ask a colleague for their opinion about a sign of which you are uncertain
3. Keep practising examination skills so that the sequence is second nature, as this will free up your mind to assess the significance of findings

M1: Relates explanations to patient's perspective

1. Check what the patient already knows before beginning your explanation.
2. If appropriate, determine what they want to know and whether they want anyone else to be present
3. Establish what you can about the patient's lifestyle, beliefs, cultural background and abilities and take these into consideration.
4. Whenever possible, link back in your explanation to the patient's initial ideas, concerns and expectations
5. Explain what you are thinking and seek their views
6. Mentally rehearse good questions about dilemmas in patient management for example "People who are nearing the end of life sometimes like to state a preference about where they should die. Do you have any views on that?"

M2: Gives clear information in small chunks

1. Use clear language, avoiding technical jargon
2. Provide information in 'small packages' particularly if it is distressing or complex.
3. Use the patient's response as a guide to how to proceed
4. Give information in ways which promote recall and understanding (eg using diagrams)
5. If appropriate use leaflets and good quality internet information to reinforce your explanation and advice.

M3: Negotiates a mutually acceptable plan with patient and/or third parties

1. Think about how the patient can actively participate in decisions about their care and encourage them to do so
2. Determine whether they want to be involved in planning and whether they have any preferences
3. Offer suggestions and choices rather than instructions
4. Discuss with the patient the management options and your recommendations and ensure they have sufficient knowledge to make informed decisions.
5. When planning, focus on the patient's goals rather than the patient's problems, for example the elderly patient with heart failure who doesn't want to risk incontinence by taking their diuretics
6. Check whether they agree to your suggested plan

M4:Reassures appropriately

1. Where appropriate, aim to reassure and offer hope.
2. Get the full picture before offering reassurance

M5: Checks understanding

1. Ask the patient whether they have understood what you have said and give them sufficient opportunity to question you.
2. Explore the patient's reactions (beliefs and feelings) about the information you have given them and address them where necessary
3. Sometimes it may be appropriate to ask the patient to repeat back their understanding of the management plan and what they are to do.
4. Enquire of the patient 'Is there anything else you would like to ask about what we have said?' before ending the consultation.

M6: Gives key evidence-based information

1. Guidelines for management are often published with the strength of supporting evidence. Choose management strategies in line with current best evidence
2. Identify and use routinely a trustworthy clinical evidence website such as http://clinicalevidence.bmj.com or www.evidence.nhs.uk to evaluate the treatments you propose.
3. Identify the patient's needs and adapt the information you give accordingly

M7: Explores available options, risks and benefits

1. Start exploring options by acknowledging the patient's expectations eg. "I realise you were hoping for antibiotics but…"
2. Explain the likely impact of each management option
3. Explain risk and benefit in terms the patient is likely to understand
4. Make sure options are realistic and relevant

M8: Investigates appropriately

1. Remember to consider the need for investigation and consciously be aware of the reasons for and against any potential investigation.
2. Remember that any investigation should only be performed if the result will change management
3. Discuss the value of the investigation with the patient
4. Make sure the patient knows when and how they will hear about the investigation and its results.

M9: Prescribes rationally

1. Think about the reasons for and against prescribing a particular drug.
2. Always consider the major side effects and interactions.
3. If in doubt don't guess – consult the British National Formulary. Don't be afraid to do this infront of the patient
4. Ensure the patient understands how prescribed items should be taken, the expected impact and the principal side effects to be expected.

M10: Refers appropriately

1. Remember to consider the need for referral and consciously be aware of the reasons for and against any potential referral.
2. Become familiar with the potential options including interprofessional referrals
3. In some cases self-referral for example to support groups, a religious advisor or complementary therapist may be appropriate.

M11: Makes appropriate use of opportunities for health promotion

1. Remember to provide preventive advice relating to the presenting problem. For example the need to give up smoking for the patient with angina.
2. Consider whether to address any of the opportunities for promoting good health which are not directly related to the presenting problem eg smoking cessation.
3. Check the patient's readiness and motivation to change before giving advice.
4. Emphasise the positive benefits for making the change, as well as the harmful consequences of continuing.
5. Focus on areas of the patient's responsibility and what they can and should do
6. Where appropriate, ask the patient to commit to the behaviour change they are going to make.

M12: Agrees appropriate follow-up

1. Remember to always "Safety-Net". Explain to the patient what the expected course is and what to do if it differs.
2. Make clear if and when the patient should return.
3. Consider who is the most appropriate health-care professional to follow up your patient

Clinical Reasoning

CR1: Seeks relevant and specific information from patient's record or third parties

1. Prior to consultation review patient's record to elicit key information such as age, significant past medical history, current medication, and reason for recent consultation(s).
2. Consider whether 3rd parties could contribute information useful to the patient's assessment or management and, if so, approach them with the necessary consent
3. During the consultation re-examine the record where this is likely to contain information you require, particularly if the patient is unsure of factual details. Signpost that you are doing this.

CR2: Generates appropriate working diagnoses or problem list

1. Where possible try to construct specific pathological, physiological and/or psychosocial diagnoses. If this is not possible, try to identify specific problems.
2. Consider your pre-diagnostic interpretation when generating appropriate hypotheses.
3. Consider using pathological sieves to help you to generate appropriate hypotheses.
4. Appreciate the importance of the background factors influencing the health of your patient
5. Consider your diagnostic hypotheses in the light of your pre-diagnostic interpretation and challenge any inconsistencies.
6. In generating any single hypothesis deliberately test it with information for and against, and then try to identify and fill any gaps.
7. When considering your diagnosis, think about what is MOST likely, what is LESS likely and what needs to be EXCLUDED
8. Be prepared to reject diagnoses for which there is little or no support.

CR3: Seeks relevant and discriminating information from history, examination and investigations to help confirm or refute working diagnoses

1. Consciously identify the key clinical features of each of your working diagnoses.
2. Use focused questions to fill gaps in the information you are attempting to gather.
3. Always assess whether the patient looks well or ill, particularly children, and consider how this might influence your working diagnoses.
4. Actively seek clinical signs that are appropriate to your differential diagnosis and its severity
5. Consider whether specific tests/investigation are needed to confirm/exclude important diagnoses

CR4: Correctly interprets information obtained

1. Take sufficient time to consider what the information you have gathered means and how to apply it
2. To help your thinking summarise and reflect back to the patient what they told you. This will confirm to the patient you have understood the problem, and will clarify your thoughts.
3. If you recognise a pattern of symptoms and signs that nearly fits a diagnosis, consider carefully any feature that does not fit, and think again.
4. If there appears to be an obvious diagnosis, consider alternatives
5. If in doubt, consult reference ranges for limits of normal values – you are not expected to memorise all of these.
6. All tests are subject to error, and false positive and false negative results are common so consider this in interpreting results.
7. Make sure you consider all the information you have gathered before making your final diagnosis
8. Each history/examination does not necessarily yield a clear diagnosis, and patients may have more than one condition. Be careful not to dismiss symptoms or signs that could be significant, particularly if felt to be so by the patient

CR5: Applies basic, behavioural and clinical sciences to solution of patient's problem

1. If in doubt about the nature of the problem think how your knowledge of anatomy or physiology can help you reconsider it from a different angle.
2. Improve your awareness of the key features of particular diagnoses.
3. Be prepared to check with books, 'on-line' sources; colleagues, etc., particularly for single items of information.
4. Focus your learning on the discriminating features of diagnoses.
5. Practise translating findings into abstractions (semantic qualifiers). E.g. 'last night' becomes 'acute', food getting stuck becomes 'dysphagia'.

CR6: Recognises limits of competence and acts accordingly

1. Do not be afraid to tell the patient you do not know something. They will usually appreciate your honesty.
2. When you have reached the limits of your competence, do not guess – seek appropriate help by asking a colleague, or consulting information sources.

Building and Maintaining the Relationship

R1: Develops and maintains a professional relationship with patient

1. Adopt professional courteous behaviour relevant to the circumstances
2. If you have met the patient before, remind them who you are, check what has already happened, and ask what has happened since last meeting
3. When presenting a patient to a colleague, remember that you are talking about a person who is in the room with you. Think how you would want your story told. For example, use the patient's name: 'This is Mr John Smith…' in preference to the term 'This patient has…'

R2: Respects the patient's ideas, beliefs and autonomy

1. Acknowledge the patient's coping efforts and appropriate self-care
2. Respect the patient's right to decline investigation/treatment, explain the impact of their decision and make it clear that that you will continue to care for them

R3: Responds empathically

1. Try to consider what it would be like to be in the patient's shoes and respond appropriately within professional boundaries. Appropriate responses can include verbal (e.g. 'I can see you are angry'; 'I can understand that', 'I can see why you are distressed about it') and non-verbal acknowledgement of the patient's state.
2. Do not make assumptions about how a situation may affect a patient
3. Beware using your personal experience to align with a patient
4. When examining a child consider it from the childs perspective
5. Be aware of your reaction when the patient says something which shocks or surprises you

R4: Fosters collaboration

1. Be prepared to explain your thinking to help the patient to understand their condition and to engage them in its management
2. Acknowledge the patient's views about the problem and its management when you are sharing decision-making.
3. If the patient does not want to collaborate with your management plan, explore why and consider alternatives
4. Specifically consider what information (good or bad) you can share and consider who this is shared with (relatives etc).
5. Using the patient's own words will sometimes help collaboration
6. Allow the patient the opportunity to ask questions

O1: Considers and optimises the setting

1. Organise your consulting space (e.g. chairs, screens etc) and minimise potential distractions (e.g. bleeps, telephone calls) for the benefit of the patient and the consultation.
2. If a consultation is still on your mind take a moment to compose your thoughts before seeing the next patient
3. When you have done what you can to optimize the setting and it is still not ideal, acknowledge this and apologise if appropriate

O2: Uses third parties appropriately

1. Ensure you identify and acknowledge any third parties within the consultation.
2. Where appropriate, obtain patient's consent for disclosure of information to third parties.
3. Be aware of the effect a third party may have on the information you can obtain and give. You may need to ask the patient whether they would like the third party to stay; you may need to ask the third party to let you talk to the patient alone first.
4. Make good use of the contribution that third parties can make to the different areas of the consultation such as the history, examination or patient management.
5. Consider the ideas, concerns, expectations and other agendas of third parties in your thinking, and explore those in more detail where it may be relevant to the consultation.
6. Keep the focus on the patient. Always make sure you address the patient first even if they cannot respond

O3: Exhibit a well-organised approach to gathering and giving of information

1. Be systematic in gathering information , for example - finish one area before moving on
2. Before you examine the patient, consider whether you have gathered sufficient information from the history.
3. When managing the patient, first reach a shared understanding of the problem and then move on to give advice or explain the treatment you are recommending.

O4: Makes organisation of consultation overt to patient

1. If appropriate, clarify the time both you and the patient have available for the consultation
2. Indicate to the patient what is going to happen next (Signposting).
3. At appropriate stages, summarise back to the patient the key elements of the consultation (for example the history) to demonstrate that you have understood each other
4. If you need time to think, tell the patient that you are gathering your thoughts. Make some brief notes if necessary.

O5: Prioritises agendas appropriately

1. Be sure you understand the patient's agenda by allowing them to complete what they wish to say, checking whether there is anything else.
2. Where there is more than one agenda (including your own), agree to deal first with the most urgent (medical priority) unless the patient cannot focus on that one before another is discussed (patient's priority).
3. Take note of the other agendas to be addressed later and indicate/negotiate how they are to be covered.

O6: Summarises appropriately

1. Summarise to enhance the consultation (e.g. to clarify, before signposting or to emphasise important points)
2. At appropriate stages, summarise back to the patient the key elements of the consultation (for example the history) to demonstrate that you have understood each other.

O7: Uses time appropriately

1. Be aware of the time. It may be helpful to keep a clock in view.
2. Having identified your priorities, allocate time appropriately to the tasks of the consultation.
3. Be efficient (have your tools to hand; good pace; concise choice of words and examination)
4. Aim to be successful with your examination at your first attempt so that you avoid repetition

O8: Closes consultation appropriately

1. Indicate that you are about to close and ask whether there is anything else the patient would like to say or ask
2. Summarise the consultation briefly and clarify the plan
3. Remember safety netting - tell the patient what you expect to happen, things to be concerned about and what to do if it doesn't happen as predicted
4. Medical students should thank the patient for what they have gained from the consultation. This may sometimes be appropriate for doctors too

Record Keeping

RK1: Makes concise and accurate notes without interfering with dialogue or rapport

1. Do not write during the patient's opening statement, as you will miss important information and may appear not to be listening
2. Your notes during a consultation should be minimal – train yourself to remember, and write only what you will forget
3. Particularly important to jot down are: people present; key words in information gathered from and given to the patient; examinations and procedures carried out
4. If you are taking notes, explain why and gain the patient's consent

RK2: Record - Diagnoses/ problems

1. After every consultation record the problems or diagnoses in specific terms

RK3: Record - Relevant history and examination

1. As a minimum, record the features of history and examination which support or refute possible diagnoses
2. (For computerised records) If there is a read code the general rule is use it rather than writing free text.
3. Record assessment of capacity to consent if this might be in question
4. Record your impression at that time (diagnosis and differential)

RK4: Record - Outline of management plan; therapy, investigations, referral and follow up

1. Document what tests will be done and, if appropriate, how these might affect management choices.
2. Record in the notes to whom a referral has been made, and how (by telephone, fax, Choose and Book, Post etc) Indicate whether the referral was routine or urgent.
3. Keep a copy of the referral in the patient records.
4. Document plan for unexpected deterioration for example who should be contacted and how
5. Record management options discussed with the patient and the patient's choices.
6. Ensure referring professional and others involved in patient's care are copied into correspondence, as appropriate.

RK5: Record - Information, instructions and special precautions given to the patient

1. This information should appear on the prescription and also in the patient's records.

RK6: Record - Identification of the author and date of record

1. When recording in the patient's record ensure that you document clearly: Date, time, your name and role (and when available GMC number)
2. Sign all entries you make in the notes

CP1: Engages and orientates colleague

1. State purpose of communication if not implicit in situation e.g.: 'I would like to practice case presentation'
2. Orientate listener with basic patient details and key background information. E.g. 'This 24 year old man with diabetes has been admitted with a vomiting and since admission he has become drowsy'
3. Consider what the function of your presentation is and frame your presentation in this light for example a teaching presentation will be long, requests for emergency assistance will be very brief

CP2: Delivers relevant detail with clarity and logical order

1. Paragraph grouped data appropriately with headings and their relevant content
2. Invite the listeners to ask questions at appropriate points in your presentation
3. Signpost the hypotheses you are considering or have considered.
4. Present relevant data. This will depend on the context e.g.social factors may be less important on acute admission than when planning the patient's subsequent discharge
5. Identify and present data that allow determination of the patho-physiology, the aetiology and the functional effect of the health problem.
6. Use SBAR (Situation, Background, Assessment, Recommendation) to organise your presention.
7. Allow, promote and manage dialogue during the presentation to ensure that all important aspects are adequately explored. This may require that you point out that there is more data which you consider relevant e.g. 'There are social factors which I feel need to be considered'

CP3: Communicates Interpretation of data transparently

1. With your interpretation offer the evidence on which it is based. E.g. 'This patient has rapidly progressive dysphagia. He has gone from difficulty swallowing meat to only swallowing water in 4 weeks.' 'I have a patient who is in shock with a BP of 90/50 and pulse of 120. '
2. Distinguishes clearly between historical report, examination findings and interpretation / opinion.
3. Be open about omissions in your assessment, for example 'I forgot to percuss the chest'

CP4: Draws purposeful conclusion

1. Consider the triple diagnosis (at the level of physical, psychological and social patho-physiology) and present what is relevant
2. Summarise succinctly with backing evidence. Be honest about uncertainty.
3. Invite comment on specific request, suggested management plan or need for clarification in a way that relates to the purpose for the communication declared previously.

Chapter 7

Grades in formative workplace-based assessment: a study of what works for whom and why

Janet Lefroy[1]

Ashley Hawarden[1]

Simon P Gay[1]

Robert K McKinley[1]

Jennifer Cleland[2]

[1]School of Medicine, Keele University, Keele, UK

[2] School of Medicine, University of Aberdeen, Aberdeen, UK

Abstract

Context Grades are commonly used in formative workplace-based assessment (WPBA) in medical education and training but may draw attention away from feedback about the task. The dilemma is that the self-regulatory focus of a trainee must include self-awareness relative to agreed standards, which implies grading.

In this study we aimed to understand the meaning which medical students construct from WPBA feedback with and without grades, and what influences this.

Methods Year 3 students were invited to take part in a randomised crossover study in which each student was their own control. Each student had one WPBA with and one without grades, and then chose whether or not to have grades with their third WPBA. These preferences were explored via semi-structured interviews. A realist approach to analysis was used to gain understanding of student preferences and the impact of feedback with and without grades.

Results and discussion Of students who had feedback with and without grades, 65 (78%) then chose to have feedback with grades and 18 (22%) without grades their third WPBA. 24 students were interviewed. Students described how grades locate their performance and calibrate their self-assessment. For some, low grades focused attention and effort. Satisfactory and high grades enhanced self-efficacy. Grades are also concrete, powerful and blunt, can be harmful and need explanation to help students create helpful meaning from them. Low grades risk reducing self-efficacy in some and may encourage others to focus on proving their ability rather than on improvement. A metaphor of the semi-permeable membrane is introduced to understand how students reduced potential negative effects and enhanced the positive effects of feedback with grades by selective filtering and pumping.

Conclusion This study illuminates the complexity of the processing of feedback by its recipients, and informs the use of grading in provision of more effective, tailored feedback.

Background

Feedback is an important teaching and learning tool (1–3). Indeed, workplace-based assessment (WPBA) of the competence and attitudes of trainees with feedback from their assessor is considered one of the most powerful interventions in medical education (2,4,5). For example, Veloski *et al*., found that feedback had a positive impact on physician performance in 74% of 41 studies reviewed (5). However, there are sometimes powerful negative effects and problems with feedback in higher education (6). Students commonly report feedback as difficult to understand (7); non-specific (providing only vague and sweeping generalisations) (8); or difficult to act on

116

(6,9). Overly critical feedback does not encourage learning, can be perceived as an attack and have a destructive effect (1,10–12).

Grades (e.g. below/meets/above expectations) have long been a component of WPBA in medical training (for example in the mini-CEX (13) and the Leicester Assessment Package (14)). This seems logical in a profession where both competence in clinical practice and knowing your own limits of competence are essential, and in which enabling learners to self-monitor in relation to competency requirements is an important goal (15,16). Yet there is a paradox: giving grades as part of a student's feedback in school settings and elsewhere in undergraduate education can reduce the effectiveness of feedback and may reduce student performance (1).

In research into what makes feedback effective, ineffective or harmful, the relationship between the recipient and giver, the state of mind and maturity of the recipient, the feedback process and the feedback content (11,12,17–19) are key factors. Feedback should clarify the student's position and progress relative to required goals, and suggest how to attain the goals (1); feedback is most useful if it is from a trusted assessor who knows the student (5) who is in a learning frame of mind (20) and there is a diagnostic and supportive dialogue between assessor and student which enhances motivation (21) because the suggestions in the feedback align with the student's goals and therefore seem relevant (11).

Factors which affect the impact of grades in formative feedback have been researched (8,10,22). While grading can give an unambiguous answer to a student's questions "Where am I relative to where I am going?" and "How well am I progressing?" which are the basis of future learning, and may increase involvement and effort, it may not improve performance because it mixes feedback about the person with feedback about the task thus distracting attention from feedback about the task (1,10). Grading may affect the tutor-student relationship because the tutor is both helping the individual to learn and passing judgement, and is thus explicitly in a position of power (5,8,23–25). Grades may also lack value in WPBA because differences in assessor calibration make standardisation of assessment almost impossible (26,27). This would be of particular concern if students were more interested in their grades than in the qualitative information which may accompany them.

So why do we use grades at all in formative WPBA? Should we stop? To date, most research on grading in formative assessment has been in school children and non-medical higher education. While useful for hypothesis generation (10), the findings may not be directly applicable to today's outcomes-based and outcomes-focused medical education. While there is a move towards portfolio-based self-directed learning without grading in some medical schools (28), medical students learning skills value grades (29). Does grading encourage or discourage feedback-seeking and self-directed learning? Or does this depend on something within the student, as suggested by research into the differing receptiveness to formative feedback depending on performance in summative assessment (30,31), when students who paid least attention to their feedback were those who had narrowly passed the assessment. Self-regulation theories suggest that within each of our students are two basic self-regulation systems which co-exist but may conflict (22,32). These two systems - the promotion (seeking good) and prevention (avoiding

harm) approaches - may both be active in response to feedback. How does grading influence this balance? There is a clear need for research comparing the impact of feedback with and without grades within medical education, as a contribution to understanding how assessment influences learning (33,34).

In this study we aimed to understand first the meaning which medical students construct from workplace-based assessment and feedback with and without grades and, second, how this is influenced by the students' internal and external environment. A further aim was to use this information to develop more effective, individually-tailored feedback processes. We did so using a crossover study in which each student was their own control. We compared student experience of, and preference for, grades or no grades. A realist logic of analysis was used to gain understanding of student preferences and the impact of feedback with and without grades (35,36).

Context

Keele University School of Medicine has a five-year programme with repeated formative WPBAs of students' consultation skills using the Generic Consultation Skills (GeCoS) assessment tool (37,38). Each WPBA includes a face-to-face discussion between student and tutor about the student's strengths, areas which need to be improved and strategies for achieving this. Students receive a written summary of the discussion including their grades for each assessed domain of their consultations. The grading scale (Must Improve - Borderline – Proficient – Very Good) is referenced to the standard required of a graduating doctor (15). The first formal WPBAs take place during a four-week general practice (family medicine) placement at the end of Year 3 after the summative knowledge assessment and Objective Structured Clinical Examination (OSCE) when each student has three WPBAs - done in weeks one, two and four of this placement, and receives feedback from their GP tutor.

Methods

Recruitment

During the academic year of 2011-2012, all year three students and their GP tutors were invited by email to take part in the study. Participation was voluntary and non-participation had no influence on any aspect of teaching or progress.

Participation for students involved:

- Opting to receive grades for all three WPBAs (our practice at that time) or consenting to randomisation into one of two study groups:

  - Group 1: First WPBA with grades, second without, third student's choice.

  - Group 2: First WPBA without grades, second with, third student's choice.

- Consent (or not) for the research team to access their recent summative OSCE results (stations passed out of 12) and the three written WPBA summaries from the general practice placement.
- Consent (or not) to be interviewed in the four weeks following the placement. .

Participation for GP tutors involved:

- Consent for the research team to access their WPBA written feedback.

Students who consented to randomisation therefore experienced WPBA with and without grades before being asked their preference (grades or no grades) for the third and final assessment. This preference was obtained via an online questionnaire in the third week of their placement.

In order to sample the widest spectrum of student approaches to grading, maximal variation sampling was used when selecting interviewees with respect to gender, attainment in the recent OSCE and preference for each of the three options: grades with their third WPBA; no grades with their third WPBA; or grades with all three WPBAs. Semi-structured interviews were conducted at the end of the placement exploring what students felt about: their placement; their progress with consultation skills; their views on WPBA; what they could recall of each of the three WPBAs; why they chose grades or no grades and the impact of having or not having grades on their feedback; whether they agreed with their feedback and would wish to have such a choice in future WPBAs. One researcher (JL) conducted the interviews. Interviews were audio-recorded with consent, and transcribed verbatim.

Analysis

We chose a realist approach to data analysis, in which the question "What works for whom under what circumstances and why?" is asked. The aim of realist evaluation is to understand relationships between context and outcomes by discovering some of the workings or 'mechanisms' of the 'black box' of complex interventions (such as feedback with and without grades from a WPBA). These are termed CMO (context-mechanism-outcome) configurations (36,39). We considered realist evaluation to be closely aligned with the study aims: subjects were involved in the process of working out why (mechanism) they had reacted in the way they did (outcome) in different feedback situations (context). Furthermore, realist approaches can raise suggestions for adjustment of the complex intervention under study once the relationships between context, mechanisms and outcomes are better understood (36).

The realist evaluation approach involves testing an initial programme theory about the working of the 'black box'. A programme theory is a statement of what we think it is about the intervention which generates change. Our initial programme theory was based on two assumptions from the literature. Firstly that feedback should answer the student's questions "Where am I going?", "How am I progressing?" and "How can I make progress?" which should result in discovering

strategies for improvement (1), and that adding grades to the context would enhance this outcome by adding clarity. Secondly, based on theories of self-regulated learning, in which goal level, persistence, effort, and self-efficacy are the self-regulation constructs with the strongest effects on learning (21,40), that the motivational effects of grading would also have an important influence on the outcome. We were expecting these motivational effects to be either positive or negative depending on both the external context (feedback being positive or negative) and the student's multi-faceted internal context, and that mechanisms would include confirmation or conflict with self-assessment; desire to improve and desire to avoid failure (10).

The initial programme theory to be tested against our data and refined was therefore that if the student has a trusted assessor (external context) and a learning goal approach (internal context), they will find that grades (external context) clarify (mechanism) and energise (mechanism) their efforts to find strategies to improve (outcome). If students were more performance oriented (internal context) they would find that satisfactory grades (external context) reassured (mechanism) and therefore reduced efforts to improve (outcome). We were less sure what to expect to find as the outcome of lower than expected formative grades, and were looking for outcomes and their explanations.

The transcripts were therefore examined for context-mechanism-outcome (CMO) configurations – what effect (outcome) did the feedback with and without grades have (contexts), what caused these effects (mechanisms) and in what internal and external learning environment (contexts) did these occur.

The first two transcripts were analysed by all authors to develop our joint understanding of what constitutes a context, mechanism and outcome in formative WPBA. A table was produced for each transcript listing the CMO configurations identified, with columns for student comments about feedback with and without grades (the manipulated variable in the context). Subsequent transcripts were coded separately by JL and AH who compared their analyses. Where interpretation was difficult, one or two of the other researchers also analysed the transcript to reach a consensus.

The authors then compared CMO configurations containing cognitive, self-regulatory and other explanations of the effects of feedback with and without grades, seeking evidence to corroborate and refine our initial programme theory. Where it was not corroborated, alternative explanations were sought where different mechanisms might have been operating in different contexts.

Reflexivity

The research team comprised stakeholders in the programme under study (JL SPG RKM AH) in collaboration with JC from another UK medical school who was invited to provide an external perspective. JC is a clinical psychologist who has worked in medical education for 14 years. Her interests in feedback stem from working as a medical educator in undergraduate medicine, and from researching "failure to fail"(25). JL SPG and RKM are clinicians who were involved in the development of the instruments for, and the programme of WPBA and feedback. They therefore conducted analysis from that stakeholder viewpoint but with a genuine lack of certainty about

whether feedback should or should not include grades, AH was a Keele medical student who at the time of the study had recently completed Year 3 and therefore brought the student perspective to analysis. As JL was a clinical tutor known to the interviewees, we were aware that this may influence participant responses so efforts were made to diminish this effect in the preamble to interviews.

Ethical Approval

The study received approval from the Keele University School of Medicine Ethics Committee.

Results

There were 144 students in the year cohort. 110 (76%) volunteered to participate. 24 of these declined randomisation, opting to have all WPBAs with grades but five of the 24 changed their minds in week 4 and had their final WPBA without grades. 86 students were randomised and 83 of these chose whether or not to have grades for their third and final WPBA; three did not complete the choice questionnaire. Of the 83, 65 (78%) chose to have feedback with grades and 18 (22%) without (Fig 1).

Fig 1 Flow diagram of the participants' passage through the study.

Key: G=WPBA with grades; N=WPBA with no grades, GGG= all three WPBA with grades
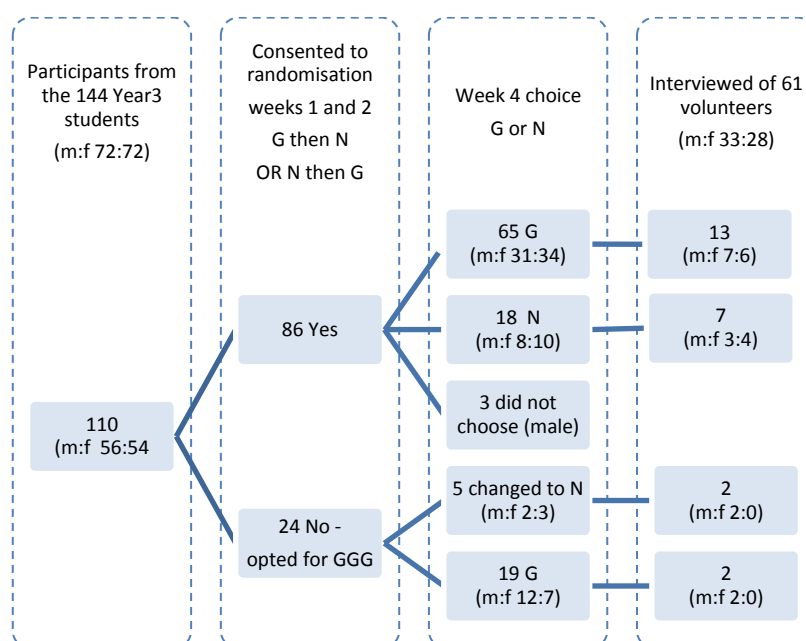
The student characteristics of gender, order of randomisation and what grades they got in their initial graded WPBA were facets of the internal context at the point of choosing grades or no grades in the final WPBA. However, these did not appear to influence the choice significantly (Table 1). Recent OSCE results did show an effect on choice but only around the pass/fail divide -

all the students who failed the OSCE chose grades in week four, whereas those who barely passed were a little more likely to choose no grades. Numbers in these two sub-groups were too small for analysis of significance but mirrored studies of feedback-seeking behaviour following summative assessment (30,31).

Table 1: Characteristics of students in the randomised cohort making a choice of grades or no grades their 3rd WPBA.

Characteristic		Choice	
		Grades	**No Grades**
Cohort	**All**	65 (78%)	18 (22%)
Gender	**Male**	31 (79%)	8 (21%)
	Female	34 (77%)	10 (23%)
Order of randomisation of WPBA	**1) Grades** **2) No Grades**	30 (73%)	11 (27%)
	1) No Grades **2) Grades**	35 (83%)	7 (17%)
Received Borderline (B) or Must Improve (MI) grades in any domains in the graded WPBA in week 1 or 2	**No B or MI grades**	15 (79%)	4 (21%)
	Did get B or MI	49 (77%)	15 (23%)
Recent OSCE results (if disclosed)	**Clear pass**	50 (79%)	13 (21%)
	Just passed	4 (57%)	3 (43%)
	Failed	10 (100%)	0 (0%)

Sixty-one students volunteered for interview. Of these, 24 were purposively sampled for interview (15 who chose grades and 9 who chose no grades; 14 male and 10 female; 19 passed the recent OSCE, three failed and were preparing to re-sit the OSCE, and two did not disclose their OSCE results for this study).

54 (83%) of 65 students who chose grades and 18 (78%) of 23 who chose no grades answered the question "Why did you choose grades/no grades?"on the choice questionnaire. These responses were included verbatim in the text for analysis and linked to the student's interview text.

Overview of how our initial programme theory changed as a result of the realist evaluation of interviews and questionnaire text responses

Our respondents validated the expected two types of mechanisms resulting from feedback: cognitive mechanisms which were generally associated with desired gains as described by Hattie and Timperley (1); and motivating and demotivating mechanisms affecting outcomes of effort and self-efficacy, which could conflict as students appraised their feedback. Students differed in which

of these mechanisms were activated, and to what extent, by the inclusion of grades in their feedback. In seeking to understand these differences in outcomes such as effort and self-efficacy we found evidence of both selective recall of feedback and selective focusing on feedback. In our resulting programme theory (fig 2), we have likened this selective recall and selective enhancement to the "filtering" and "pumping" mechanisms of a semi-permeable cell membrane which can process feedback as it allows it in. The metaphor of the semi-permeable membrane has been adopted because it is a good representation of what we have found, and will be understood by clinical educators and by medical students. We will postulate that this is how we all receive feedback, the settings of this semi-permeable barrier being attributable to various aspects of the internal and external context including the influence of being graded.

These findings are expanded and discussed below. Quotes from transcripts are identified by M=male F=female; participant ID in the order interviewed, and annotated with their choice of grades or no grades. Unless annotated as a questionnaire comment, the quotes are from interviews.

Findings: Cognitive and motivational mechanisms (representative samples provided in table 2)

Cognitive mechanisms triggered by receiving grades were locating self, clarifying goals and focusing attention. Students who chose grades perceived them as clear additional information regarding their position and progress (quotes 1&3) but might need the context of the same assessor for this to be worth the risk of harm (quote 2) and needed an explanation for the grade if it were low (quote 4). Some students described the importance of an external opinion on position and areas needing attention (e.g. quote 5). The outcome of this clarity was focused identification of strategies for improvement of consultation skills and also training in self-assessment by exit standard (e.g. quote 3). The counter argument that grades create a focus on the grade rather than on the content of feedback was made by only one student (F23 in quote 6).

Motivational effects of grades were more mixed. Grades could galvanize or reduce effort and this depended on the internal and external context as well as on the grades (quotes 7&8). Outcomes such as enhanced self-efficacy for consultation tasks and acquisition of an identity as a proto-professional were desired by many, and grades had the potential to promote or detract from these. M1 (in quotes 5, 8 &12) describes two motivational effects: grades galvanised effort but diminished self-efficacy at least initially. He was one of the (few) students who had an uncomfortable feedback relationship, describing his GP as "brutally honest" and in that context experienced some loss in self-confidence from the verbal feedback as well as from the low grades. Many students however felt they were in a very supportive relationship. Those who felt they were improving in the context of a supportive relationship described choosing to have grades in the confidence that they would get external validation which would boost their self-efficacy (e.g. quote 11).

To others, grades represented a clear risk of harm. This risk/benefit dilemma triggered varying mechanisms in these Year 3 students when faced with grading by Year 5 exit standards. Many described how being graded to exit standard orientated learning and motivated working to goals,

but risked triggering demotivation. The risks of grades causing demotivation were much more apparent for students who chose not to have grades with few describing any compensating cognitive or motivational mechanisms. In particular, grades did not galvanise effort for them, a response which some may have developed because of previous experiences. M22 describes this but was aware of his internal risk/benefit conflict and decided: "I guess grades are a necessary evil" (quote 9). F21 (quote 10) also felt that although grades could be beneficial (if explained) they were too risky: because she trusts her self-assessment, grading could be harmful if dissonant.

Some might have preferred grades in another context but chose no grades in the context of a lack of trust in the assessor or in the assessment system (e.g. quote 13). Others described grades as merely useless and irritating rather than potentially harmful.

Table 2: Inside the 'black box' - what works for whom and why? Context-Mechanism-Outcome (CMO) relationships described by students talking about the effects of grades

CMO configurations	Quote
Giving grades (C): an intervention here seen to be helping students to improve consultation skills (O) by cognitive mechanisms (locating self, clarifying progress, focusing attention) (M)	
Grades (C) help students to locate themselves clearly(M) so they know what needs improving (M), in the context of a trusted and consistent assessor (C)	**Quote 1:** "Although I have found the comments… very helpful, as it provides me with specific points to learn from, I feel the grades are useful as they give a more concrete and measurable representation of my performance." **M12** questionnaire response (chose grades) **Quote 2:** "I've had a change of tutor since the first and second weeks and I feel like there would be a lack of consistency which would be reflected in my grades so I would rather in this case not have grades than have a grade which isn't necessarily reflective of how I may have improved. But if I had had the same tutor, I would have chosen to have grades." **F19** questionnaire response (chose no grades)
Grades (C) help to clarify progress towards goals (M) but need explanation (C) to be useful	**Quote 3:** "whilst I feel feedback is all well and good and it's something that you need as well to sort of take on board and improve, I feel the only way you can see if you <u>have</u> actually taken that feedback on is if you have a quantitative benchmark to work from. ….' So now I have goals for the next two years to work towards." **M5** (chose grades) **Quote 4:** "Everyone has to reach proficiency so I wouldn't mind having grades, but I don't find grades useful unless there is detailed feedback. I can see how they are useful in terms of knowing where

	you are with respect to the exit… without feedback I can't use the grade. It merely demotivates me. I don't expect to be perfect but I need the feedback to explain the bad grade." **F21** (chose no grades)
Grades (C) focusing attention (M) resulting in increased effort in those target areas (O)	**Quote 5: Interviewer**: "In what way did receiving those grades affect your progress?" "Well then I knew… I had opening, history taking where I had specific areas of the consultation… I'd have to build upon rather than just getting generalised feedback … It sort of highlighted areas I thought were fine. Like I thought my history taking was fine… but the GP said no, you're missing all these things and then I tried to build on those things and make mine a lot more detailed." **M1** (chose grades)
Grades (C) focus attention (M) unhelpfully on the grade in the internal context of previous experience (C)	**Quote 6:** "I just don't like grades at all. I dunno. It's all the competition with other people as well – what did you get? How did you get on? I just find it really annoying. I just find myself more anxious and worried if I've got grades rather than just the no grades. …Because if you put grades on something I have a tendency to focus on the grade rather than… anything that was said… So like you could have said ten things that I improved on or did well in, but if you said to me 'But despite that your grade was this' and the grade wasn't necessarily very good I would go 'Oh'. And I would think of it as not very good rather than thinking about the fact that I did a lot of things well." **F23** (chose no grades)

Grades(C): an intervention increasing or decreasing effort (O) by motivational mechanisms (M)	
Galvanizing effort (O) by challenging complacency (M) in the context of competitiveness (internal C) or a valued student-assessor relationship (external C)	**Quote 7.** "there were quite a few that were slightly lower than I expected .., it gave me a bit of an idea of where I particularly needed to improve and have a think about which was quite useful and also it gave me more incentive to improve in those areas because I'm horribly competitive." **M4 (chose grades)** **Quote 8.** "I think you're trying to sort of get approval from the person that's assessing you and whatever and make them, you want to know that they think you're sort of improving." **M1** (chose grades)
Grades diminishing effort (O) by complacency (M) in the context of a satisfactory grade (C) or by demoralization (M) with	**Quote 9.** "A grade can be interpreted in an unhelpful way (for me at least). If I do well I won't have the drive to improve my skills further because I'm already up to standard. Likewise If I'm graded badly rather than seeing it as a reason to try harder, I'll be demoralised and unwilling to try the skill at all. This has happened twice already, it took me a while to get over it and start trying the skill again." **M22**

a low grade (C), in a student damaged by previous assessments (internal C).	questionnaire response (chose no grades)
But with potential to galvanise effort (O) in a different context (C)	**Quote 10.** "I get very nervous about grades and work myself up about them. Bad grades would knock my confidence… Grades can help you rest on your laurels. On the other hand, grades can motivate you to move into the next grade… Maybe the third WPBA in year 4 we should have grades so that if there is a significant need to improve we would have time to do that.." **F21** (chose no grades)

Grades (C) an intervention influencing self-efficacy (O) by demonstrating the presence or absence of progress or competence (M)

Grades (C) showing progress (M) therefore causing gains in self-efficacy (O)	**Quote 11.** "I feel that my consultation skills have been improving over the last 4 weeks and I have benefited from the feedback… provided… I think that having grades in the final week will show the progress I have made since week one and give me the confidence to carry this forward to my future medical practice." **M6** questionnaire response (chose grades)
(Risk of) Diminishing self-efficacy (O) in the context of –comparison with exit standard (C);	**Quote 12.** "I've recognised now that being graded against the standard of everyone else in the year, although it may sort of help my confidence a bit more than being marked at a higher level, it (being marked to exit standard) means that I know where I need to improve across sort of more long term goals than just short term comparisons to everyone else." **M1** (chose grades)
Or in the context of social norms and expectations about grades (C)	**Quote 13.** "It sounds… like, you know, you're <u>Borderline</u>, not good enough to be at medical school… going back to my family and saying "Oh I got a Borderline", that doesn't sound very encouraging." **F19** (chose no grades)

How our Programme Model was changed by Realist Analysis - Findings of selective recall and processing which we have termed "Filtering" and "Pumping", and the influence of grading on this.

A related set of mechanisms triggered by factors in the internal and external context (including getting grades) resulted in feedback being unequally assimilated. We were able to compare students' recall of their feedback with the written feedback summary they were sent a few days after each assessment. This helped us to explore what interviewees did with their feedback and why. Some did not recall much of their feedback or grades and were not clear why, but others explained mechanisms which involved their grades. Others paid special attention to parts of their feedback, both verbal and written, because of the grades they received.

Contrast these two reactions to Borderline grades:

➢ "Obviously Borderline is not something that you want but I know they use the system to assess doctors so I thought well I'm not gonna be at that stage yet so I didn't pay that much attention to it." F15 (chose no grades)

➢ "I feel that if you get a Borderline grade it makes you think "I really need to improve on this area" and for me gives me a push to do it! If you don't get grades, for me it doesn't give me the same push to think I really have to improve in this area." F25 Questionnaire comment (chose grades)

There seemed to be a 'semi-permeable barrier' to assimilation of feedback which we have likened to the cell membrane with open or closed channels set to pump or selectively transmit feedback. In this analogy it is a 'thick membrane' with embedded processing capacity. The 'membrane' effects were driven by self-protection, dealing with dissonance and focusing attention. Examples of these are given in Table 3.

<u>Filtering for self-protection</u>

Because of the risk of harm from feedback, especially from grades, it would not be surprising if students selectively filtered their feedback. This was demonstrated by F15 (quoted above in the text) who described ignoring her Borderline grades. She then chose not to have grades in the final assessment, rationalising that there was no concern or the tutor would have told her so. She and several other students indicated that they would in future be likely to choose grades now that they understood the exit standard grading system and could therefore make meaning out of it. Others found less than top marks too unpalatable, for example, F10 (quote14) chose grades to gauge where she was, how she could improve and by how much she had improved. She recalled some of her written feedback but could not recall her grades except she remembered not liking them. She described how her tutor had tried to scaffold her understanding of grading by exit standard but she was unable to make use of the grades in the way she had intended.

Table 3: How do students handle their grades?

CMO configurations	Example
Filtering (M) for self-protection from risk (O) in the context of low grades by exit standard (C); grades could be shocking	**Quote 14.** "My GP … kept saying don't be disheartened by this cus like although you're only proficient at it you will be very good very soon cus you're going along the right tracks but you won't be very good until you're a F1… I think when you sit down and think oh actually yes at this level I should only be proficient so it's o.k. that I was put as proficient but I think when you think you've done quite well and then someone says you're two or three out of five and you were expecting to be five out of five sort of thing I think it's quite a shock (laughing) so I probably wouldn't have chosen to have grades if I'd known, if I'd seen it before I probably would have preferred just to have the writing cus I think the writing reflects much better on what I was doing rather than the grading." **F10** (chose grades before seeing her week 2 grades)
And in the context of low self-assessment (C)	**Quote 15. Interviewer**: "Would there be any scenario you can think of where you'd say I won't have grades thank you?" "Where I had done rubbish and… I didn't want to know." **F24** (chose no grades)
Processing of the meaning of grades (M) to resolve dissonance (M) in the context of low grades and a trusted assessor (C). Once processed, the grades then triggered the next mechanism – "pumping" of feedback (M) resulting in seeking strategies for improvement (O).	**Quote 16. Interviewer**: "Did the grades have any impact on you?" "Yes. I felt a bit bad cus I thought especially in third year, I shouldn't be achieving like this since we've just done OSCEs (laughing), like two or three weeks ago. So I thought well I should improve and then erm… My friend was really, really good erm so I just started like looking, picking up things that he picked up and then looking at the GP's feedback a lot. And then over the next two weeks, I improved… And I had to like obviously respect their views because they've been doing it for a long time so… I thought - Well the criticism is not really like bad criticism, it's good criticism to help me sort of help me learn and move forward and improve my skills… I think with the grades it made me feel like I should <u>do</u> something." **M6** (chose grades)
Grades (C) induced a temporary performance approach (O) but after some processing (M) they enhanced learning	**Quote 17.** "I prefer to have grades so I can see what my abilities are at the moment and what I need to improve on. Grades give me reassurance if they are good or motivate me to work harder if they need improvement." **F14** questionnaire response, then in the same student afterwards at interview: "I was concerned that erm I'd been given Borderline for Examination and Management but the GP explained that that's the level I was at - at

effort (O)	the moment - in terms of erm a doctor… so I'm not sure if he meant that I was Borderline for a student or what. But I would obviously like to improve them.
	… In examination I think I tried to erm volunteer to examine more patients … I just wanted to sort of prove that I did know how to examine.
	… I think Examination is one of those things that takes a lot of practice anyway… I wouldn't expect it to be much above Borderline at the moment anyway because you know it's something to improve on." **F14** (chose grades)
Grades (C) triggering "Pumping" (M) to enhance learning in areas needing effort (O)	**Quote 18.** "If I just had verbal feedback erm… I just think having that Must Improve will keep me sort of in mind that it is something I will need to continue working on. But… if see a grade it sort of sticks in my mind quite a bit more than someone just talking to me." **M1** (chose grades)
	Quote 19. "She didn't want to give me erm for example an excellent rating because otherwise I would be too complacent with that so she gave me Can Be Improved (sic) for each of those categories."
	Interviewer: "And is that true that you would have been complacent?"
	"Erm perhaps so, perhaps so. …It was good that at the end of each week we'd get, erm an e-mail of all our results so I could look back on them and improve on them and the fact that she's ticked those boxes, the box that erm I could improve on to get associated symptoms so the next time I went to surgery I acted on those recommendations." **M3** (opted for all grades, was due to resit OSCE)

Filtering out the grades but making meaning from the feedback seems harmless. What is of more concern is when the filter blocks the entire set of feedback because of risk of harm from the grades. F24 (quote 15) could recall feedback from week 1 and week 4 when she had no grades but all she could recall of week 2 was one Borderline grade. We were able to examine the written feedback which was equivalent on all three weeks excepting the presence or absence of grades. F24 seems to have reacted to this grade by filtering out the entire set of feedback. F23 (quote 6) also indicated that she only saw the grade and not the feedback.

Processing to resolve dissonance

Several students described how they had struggled to make sense of their grades when they did not initially agree with them. M6 and F14 illustrate this (quotes 16&17). M6 in the context of the

GP being a trusted advisor had to resolve the dissonance. He described an initial drop in self-efficacy followed by a galvanizing effect of grades with Borderline grades forcing attention on why he was not as good as his peer. F14 wished to prove that she was better than the Borderline grade she had been awarded. She later returned to her initial learning goals as stated in her questionnaire response.
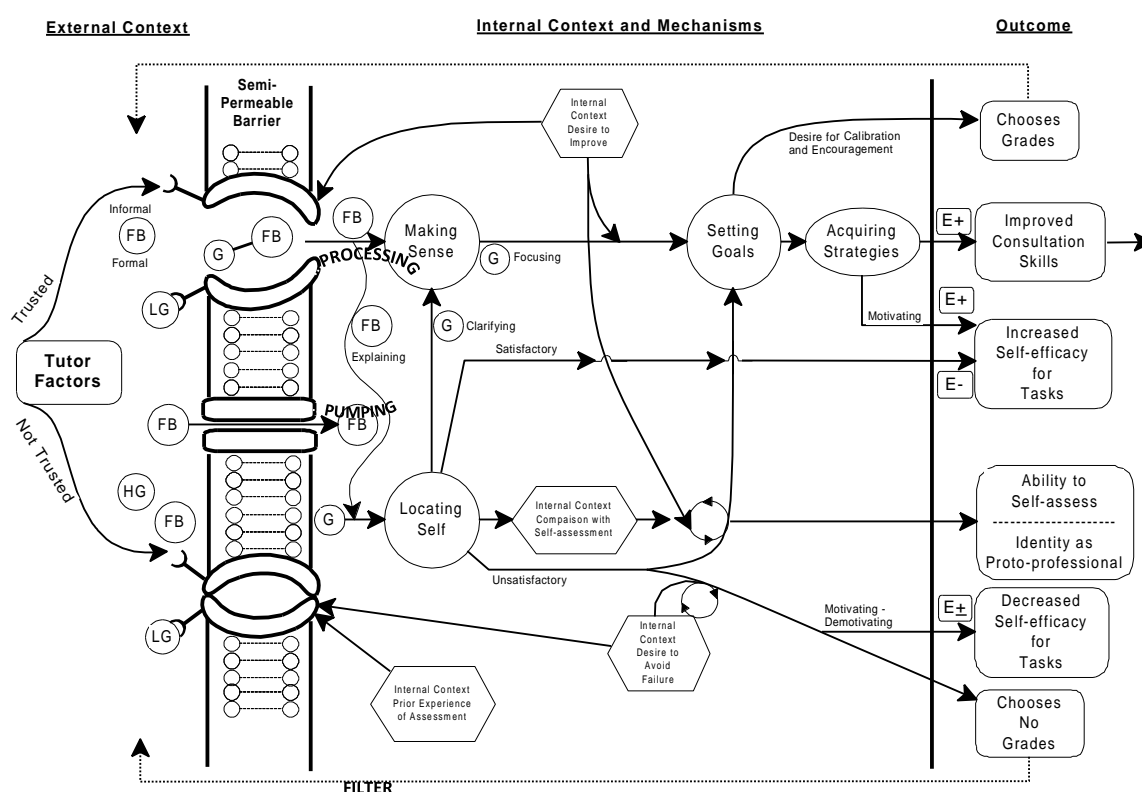
<u>Pumping to enhance learning</u>

There was also evidence of selective "pumping" of the feedback, and students described using their grades to trigger the pump, by focusing attention and effort as illustrated by quotes 18&19. This is contrary to the expectation that grades would prevent students from taking notice of their other feedback (1).

<u>Programme model</u>

So what works for whom and why? The mechanisms induced by grades are complex. Here we have gained a few glimpses into the "black box". During our analysis we refined our programme model (Fig 2). We considered that the act of choosing grades or no grades was an interim outcome which fed back into the context as shown.Not shown in Fig 2 are the other elements of the external context independent of WPBA which were also mentioned as contributing to the learning outcomes such as the types of consultations students were conducting and their feedback from peers at the weekly video feedback session.

Fig 2 Programme model depicting some of the contexts, mechanisms and outcomes found in this study

(FB) = feedback (informal and formal descriptive)

(G) = grades

(HG) = high grades

(LG) = low grades

[E+] = increased effort to learn / to perform

[E-] = decreased effort to learn / to perform

Contexts are shown as hexagonal shapes ⬡

Mechanisms are arrows ⟶

Mechanisms which are also internal outcomes are in circles ◯

Context - Mechanism - Outcome configurations illustrated in Fig 2

Initial Programme Theories confirmed:

1. Feedback has sense made of it leading to goal-setting for areas needing improving. This leads to acquiring strategies for improvement and the final outcome is improved consultation skills.
2. Acquiring strategies for improvement and noticing improvement also increases self-efficacy.
3. Grades enable the student to locate him/herself against the goals thus clarifying the meaning made of the descriptive feedback.
4. Satisfactory grades increase self-efficacy and can cause bypassing of the goal-setting route.
5. Unsatisfactory grades can also have negative outcomes by demotivating.

Additional CMO configurations found:

6. Feedback channels in the semi-permeable barrier may be opened or closed by the degree of trust in the tutor.
7. Feedback channels may also be opened or closed in response to low grades, as set by the internal context – desire to improve or desire to avoid failure, and prior experience of assessment.

8. Grades need explanation so draw attention to relevant feedback i.e. cause "pumping" of feedback.
9. Grades focus goal-setting.
10. Unsatisfactory grades trigger different mechanisms depending on the internal context – in the context of a desire to improve, low grades enhance goal –setting; in the context of desire to avoid failure, low grades decrease self-efficacy.
11. The initial motivating effect of low grades in the performance goal approach turns to demotivation if the additional effort to perform doesn't result in higher grades. This contrasts with the learning goal approach where low grades motivate additional effort to learn.
12. The interim outcome of choice of grades results from desire for calibration and/or encouragement triggered by previous grades.
13. The interim outcome of choice not to have grades results from either the desire to avoid harm, or inability to make use of the grades for a number of reasons (not shown to avoid over-complicating the diagram).

Discussion

This is the first study directly comparing feedback from WPBA with and without grades in undergraduate medical education. We found that 78% of our mid-stage medical students want grades and that grades can be helpful when linked to formative assessment. Grades locate the student and calibrate their self-assessment, promoting self-awareness of competence and limits and helping them track progress, especially if referenced to an understood standard such as that expected of the graduating doctor (15). In students who have well-developed self-regulation, low grades can focus attention and effort; satisfactory and high grades can enhance self-efficacy; and in the context of constructive feedback from a trusted assessor, students will usually create constructive meaning out of low grades.

However, we also found that 22% of students preferred not to have grades for a number of reasons including avoiding harm. This suggests that for such students, if grades are given, care must be taken to adjust the external context to diminish the risk of harm and to help the student orient their internal context (self-regulate). In other words, grades are concrete, powerful and blunt and need explanation to help students create helpful meaning from them. Grades can also create complacency: it is easy to focus on the grade and to ignore carefully prepared advice on how to improve. When the student has a prevention approach to learning (22), low grades risk reducing self-efficacy in some; and low grades may encourage others to focus on proving what they can do rather than on looking for ways to improve what they do.

The self-protective filtering, processing and "pumping" we found are comparable with the feedback-seeking behavior found in Veterinary students by Bok *et al* (41). Bok *et al* found that students adjusted their feedback-seeking behavior to avoid harm and promote gain, and the

contextual factors promoting or deterring feedback-seeking were the learning climate and relationship with assessors. So we suggest that the "settings" of the protective filters depend on prior experiences and promotion/prevention focus but are influenced by the feedback relationship. The clarity and simplicity of grades make them more difficult to ignore than undesired narrative feedback. Once allowed through the filter, sense must be made of both narrative feedback and grades. This meaning-making is also influenced by gain-seeking and loss-avoidance; hence the students' insistence that a grade must be explained and advice given on how to improve, so that they can turn a potential loss into an educational gain. Students then seem to switch the "pump" on and suck in the constructive criticism attached to the low grade. This again aligns with self-regulation theory - the student's motivational regulatory focus affects the way they handle feedback. This could spiral in either direction as the feedback affects the motivational focus (40).

Comparing these findings with our initial expectations from the literature, we did find as expected that the most prevalent contextual factors in the positive or negative influence of grading were the perception of tutor support and the student's prevention or promotion approach to and past experiences of assessment (5,22). The finding that potential negative effects were diminished and positive effects were enhanced by selective filtering and pumping modifies the expectation from the literature that grades will neutralise feedback about the task (1,10). While we did find that neutralising of feedback was triggered by grades in a few, and other students mentioned it as a potential adverse mechanism, these were capable of making adjustments to acquire the feedback they felt they needed. This is also a form of self-regulation.

While we acknowledge that other mechanisms could be operational, we postulate that the metaphor of the semi-permeable membrane for the way we handle feedback may be valid and may have resonance for many in medical education.

Strengths and weaknesses of this study

The crossover design has enabled us to examine a complex real-life phenomenon – the effect of including grades in formative feedback, and explore how students both make choices about receiving grades and make sense of feedback which does and does not include grades. The study design enabled sophisticated purposive sampling of students who received feedback in different contexts but could not control for all complexities in the context. There are undoubtedly other influences, but this design has reduced them. While we do not have direct access to the feedback discussions between students and tutors, we did have the written summaries and thus insight as to what was discussed to compare with students' recollections. The number of participants has been sufficient to enable an understanding of the process in the various learner contexts. The approach was obviously acceptable as so many participated.

The study is limited by being in a specific context: a post assessment end of year clinical placement with assessment by exit standard two years before exit with students from one medical school in (usually) supportive longitudinal relationships with tutors. Nevertheless these are also strengths: students should have been motivated by learning as they were not

approaching summative assessment, a large proportion received low grades in at least one assessed domain and we were able to analyse their response to this and found that these protective mechanisms are evident even in supportive relationships. Finally, the outcomes reported are self-reported by students but such self-report is a feature of qualitative inquiry.

Implications of the study

Given the prevalence of grades in medical education, this study has wide implications, assuming the findings are replicated in other settings, for example when students are approaching summative assessment.

The conclusions of this study for "best practice" formative WPBA are as follows: Firstly, grades are important to many medical students for a number of reasons. These reasons can be explored by a supportive tutor who can encourage a learning approach to WPBA, aiming for self-awareness of competency and prioritisation of areas for improvement. If it seems that receiving grades will enhance the seeking of strategies for improvement, they should be offered as an element of formative feedback. Secondly, the criteria for allocation of grades must be understood by students who receive them. Thirdly, not all students will accept the offer of grades where choice is available, and this decision should be respected but perhaps explored by tutors. Finally, feedback is processed in the light of previous feedback so all tutors must understand that today's feedback affects the response to tomorrow's feedback. This study provides further evidence of the complexity going on inside the heads of feedback recipients, and helps the push towards a more personalised approach to feedback. We now need to determine whether attention to these factors does indeed help students to make more use of their feedback and enhances learning.

References

1.   Hattie J, Timperley H. The Power of Feedback. Rev Educ Res. 2007 Mar 1;77(1):81–112.

2.   Norcini J. The power of feedback. Med Educ. 2010;44:16–7.

3.   Carless D, Salter D, Yang M, Lam J. Developing sustainable feedback practices. Stud High Educ. 2011 Jun;36(4):395–407.

4.   Norcini J, Burch V. Workplace-based assessment as an educational tool: AMEE Guide No. 31. Med Teach. 2007;29(9-10):855.

5.   Veloski J, Boex JR, Grasberger J, Evans A, Wolfson DB. Systematic review of the literature on assessment, feedback and physicians' clinical performance: BEME Guide No. 7. Med Teach. 2006;28(2):117.

6.   Gibbs G. How assessment frames student learning. In: Bryan, C and Clegg K, editor. Innovative assessment in higher education. London: Routledge; 2006. p. 23–36.

7.   Weaver MR. Do students value feedback? Student perceptions of tutors' written responses. Assess Eval High Educ. 2006 Jun;31(3):379–94.

8.   Higgins R, Hartley P, Skelton A. Getting the Message Across: The problem of communicating assessment feedback. Teach High Educ. 2001;6(2):269.

9.   Poulos A, Mahony MJ. Effectiveness of feedback: the students' perspective. Assess Eval High Educ. 2008 Apr;33(2):143–54.

10.   Kluger A, DeNisi A. The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. . Psychol Bull. 1996;119(2):254.

11.   Shute VJ. Focus on Formative Feedback. Rev Educ Res. 2008 Mar 1;78(1):153–89.

12.   Urquhart LM, Rees CE, Ker JS. Making sense of feedback experiences: a multi-school study of medical students' narratives. Med Educ. 2014 Feb;48(2):189–203.

13.   Norcini JJ, Blank LL, Duffy FD, Fortna GS. The Mini-CEX : A Method for Assessing Clinical Skills. Ann Intern Med. 2003;138(6):476–81.

14.   McKinley RK, Fraser RC, Van V der, Hastings AM. Formative assessment of the consultation performance of medical students in the setting of general practice using a modified version of the leicester assessment package. Med Educ. Department of General Practice and Primary Health Care, University of Leicester, Leicester General Hospital, Gwendolen Road, Leicester LE5 4PW, UK; PM:10886641; 2000;34(7):573–9.

15.   GMC. Tomorrow ' s Doctors The duties of a doctor registered with the General Medical Council. 2009.

16.   Eva K, Regehr G. Self-assessment in the health professions: a reformulation and research agenda. Acad Med. 2005;80:46.

17.   Eva KW, Armson H, Holmboe E, Lockyer J, Loney E, Mann K, et al. Factors influencing responsiveness to feedback: on the interplay between fear, confidence, and reasoning processes. Adv Health Sci Educ Theory Pract. 2012 Mar;17(1):15–26.

18.   Murdoch-Eaton D, Sargeant J. Maturational differences in undergraduate medical students' perceptions about feedback. Med Educ. 2012 Jul;46(7):711–21.

19.   Watling C, Driessen E, van der Vleuten CPM, Vanstone M, Lingard L. Understanding responses to feedback: the potential and limitations of regulatory focus theory. Med Educ. 2012 Jun;46(6):593–603.

20.   Molloy E, Boud D. Seeking a different angle on feedback in clinical education: the learner as seeker, judge and user of performance information. Med Educ. 2013 Mar;47(3):227–9.

21.   Sitzmann T, Ely K. A Meta-Analysis of Self-Regulated Learning in Work-Related Training and Educational Attainment: What We Know and Where We Need to Go. Psychol Bull. 2011;137(3):421.

22.   Kluger A, Van Dijk A. Feedback, the various tasks of the doctor, and the feedforward alternative. Med Educ. 2010;44:1166.

23.   Govaerts M, Van der Vleuten C, Schuwirth L, Muijtjens A. Broadening Perspectives on Clinical Performance Assessment: Rethinking the Nature of In-training Assessment. Adv Heal Sci Educ. 2007;12:239.

24.   Monrouxe L, Rees C, Lewis N, Cleland J. Medical educators' social acts of explaining passing underperformance in students: a qualitative study. Adv Heal Sci Educ. 2011;16(2):239.

25.   Cleland JA, Knight L V, Rees CE, Tracey S, Bond CM. Is it me or is it them? Factors that influence the passing of underperforming students. Med Educ. 2008 Aug;42(8):800–9.

26.   Holmboe E, Yepes M, Williams F, Huot S. Feedback and the mini clinical evaluation exercise. J Gen Inter Med. 2004;19:558.

27.   Yeates P, O'Neill P, Mann K, Eva K. Seeing the same thing differently: mechanisms that contribute to assessor differences in directly-observed performance assessments. Adv Health Sci Educ Theory Pract. 2013 Aug;18(3):325–41.

28.   Altahawi F, Sisk B, Poloskey S, Hicks C, Dannefer EF. Student perspectives on assessment : Experience in a competency-based portfolio system. Med Teach. 2012;34:221–5.

29.   Manzone J, Tremblay L, You-Ten KE, Desai D, Brydges R. Task- versus ego-oriented feedback delivered as numbers or comments during intubation training. Med Educ. 2014 Apr 9;48(4):430–40.

30.   Harrison CJ, Könings KD, Molyneux A, Schuwirth LWT, Wass V, van der Vleuten CPM. Web-based feedback after summative assessment: how do students engage? Med Educ. 2013 Jul;47(7):734–44.

31.   Sinclair HK, Cleland JA. Undergraduate medical students: who seeks formative feedback? Med Educ. 2007 Jun;41(6):580–2.

32.   Higgins E. Making a good decision: Value from fit. Am Psychol. 2000;55(11):1215.

33.   Schuwirth LWT, Van der Vleuten CPM. Programmatic assessment: From assessment of learning to assessment for learning. Med Teach. 2011 Jan;33(6):478–85.

34.   Galbraith RM. Got feedback? Med Educ. 2013 Mar;47(3):224–5.

35.   Pawson R, Tilley N. Realistic Evaluation. London: Sage; 1997.

36.   Wong G, Greenhalgh T, Westhorp G, Pawson R. Realist methods in medical education research: what are they and what can they contribute? Med Educ. 2012 Jan;46(1):89–96.

37.   Lefroy J, Gay SP, Gibson S, Williams S, McKinley RK. Development and face validation of an instrument to assess and improve clinical consultation skills. Int J Clin Ski. 2011;5(2):115–25.

38.   Lefroy J, Thomas A, Harrison C, Williams S, O'Mahony F, Gay S, et al. Development and face validation of strategies for improving consultation skills. Adv Health Sci Educ Theory Pract. 2014 Jan 22;

39.   Pawson R. Evidence-based policy: a realist perspective. London: Sage; 2006.

40.   Nicol DJ, Macfarlane-Dick D. Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. Stud High Educ. 2006 Apr;31(2):199–218.

41.   Bok HGJ, Teunissen PW, Spruijt A, Fokkema JPI, van Beukelen P, Jaarsma D a DC, et al. Clarifying students' feedback-seeking behaviour in clinical clerkships. Med Educ. 2013 Mar;47(3):282–91.

42.   Lefroy J, Hawarden A, Gay SP, McKinley RK, Cleland J. Grades in formative workplace-based assessment: a study of what works for whom and why. Med Educ. 2015;49:307–20.

Chapter 8

Positive impacts of written feedback:

Does formal workplace based assessment add value to informal feedback?

Janet Lefroy

Ashley Hawarden

Simon Gay

Robert McKinley

Keele University School of Medicine

Abstract

Introduction Feedback is a key component of learning but effective feedback is a complex process with many aspects. One aspect may be a written summary which is passed to the learner but this may not be valued by learners.

Methods   We examined the role of written feedback in the feedback process to determine whether it does more than provide a simple summary of the interaction. We conducted a secondary analysis of data gathered for a study of formative workplace based assessment. Interview data from 24 interviews with students and written summaries of workplace based assessments for 23 of them were reanalysed by two researchers who were already immersed in the data and examined all references to verbal, informal feedback and written, formal feedback or the assessment tool used.

Results   We found that students valued the verbal feedback discussion highly and that they often considered the written summaries superfluous. We also found that the act of preparing written feedback augmented the feedback discussion and tutors had adopted the language of the formal instrument in the verbal feedback and free text written feedback.

Discussion   What this study adds to existing research is evidence that there may be a secondary faculty development effect of requiring the preparation of written feedback which has served to enhance the educational content of feedback. Although this is not proof of causality (the requirement to provide written feedback alone producing the positive effects), we consider that the likelihood is sufficiently strong to continue the practice.

Background

Feedback to students in the clinical setting comes in a variety of forms including conversations with a clinical tutor following a student-patient encounter or case presentation; written feedback using instruments such as mini-CEX; formal clinical supervisor discussions which may involve a 360 degree appraisal; self- and peer-assessment. Many of these feedback opportunities are embedded within educational systems in which workplace-based assessment (WBA) is a required component and the formative potential can become lost in the ticking of boxes (1). Feedback at its best is an ongoing dialogue between a motivated learner and a supportive and trusted advisor through which goals are identified and strategies for improvement are agreed and reinforced (2–6). Social aspects of this interaction are key to its success and the particular educational culture influences how both learners and tutors approach feedback (7). Medical students have been found to value informal verbal feedback more than formal WBA with written feedback (3,8). One explanation for this may be that feedback works best soon after the event, especially for a complex task such as consulting with a patient (9,10).
Clinical tutors might prefer to dispense with the formal written feedback if it is true that their immediate specific snippets of spoken advice are more likely to be heeded. The destiny of most

Workplace Based Assessment (WBA) feedback forms is to be filed in a portfolio of evidence. In their darker moments, tutors sometimes question the likelihood of that painstakingly crafted piece of written feedback ever again crossing the consciousness of the learner (11). Conversely, there might be some benefits of written feedback other than ticking the box as evidence for progression. Writing is known to promote information processing in learners (12,13). Committing feedback to the written word might improve its quality by a similar cognitive process in the assessor. Medical student and trainee doctor satisfaction with their workplace feedback increases when a written card system is used to prompt feedback (14–16). It is however not clear what lies behind such enhanced satisfaction, and in one study faculty giving the feedback perceived no difference in quantity, quality and timeliness of feedback (16). The observation of a bias towards leniency in written WBA feedback with a lack of recommendations for improvement warns against equating student satisfaction with enhanced learning (15,17,18) Such findings beg further qualitative research to be understood.

Before we conclude that we are wasting everybody's time by requiring the completion of feedback forms which have little chance of being read and even less chance of improving performance, let us check for the presence of a baby in the bathwater we are considering throwing out. It is worth taking a closer look at exactly what is happening more widely in WBA systems which involve the generation of written feedback.

Context:

At Keele, we believe that the learning culture should support the giving of useful feedback. We have therefore invested heavily in serial formative WBA of students' consultation skills including the provision of formal written feedback using a validated assessment and feedback instrument 'generic consultation skills' (GeCoS), an iterative quality enhancement process (19) and a purpose-designed app and web resource. These resources enable clinical tutors to select the competencies assessed and to add their own free-text comments to relevant student-validated strategies for improvement (20,21). When we found that routine end-of year student evaluation consistently showed greater satisfaction with verbal feedback (table 1) despite this impressively detailed written feedback they were getting we decided to explore other aspects of this phenomenon.

Table 1: Evaluation data student satisfaction Y5 2012-13

How useful was the feedback?	Informal verbal		More formal (verbal and written)	
Location	Primary Care	Secondary Care	Primary Care	Secondary Care
Students selecting "very useful" and "useful"	98%	87%	80%	63%

Our research question was: Does the process of completing a written summary of WBA feedback add value for the learner?

We thought that we perhaps had access to data that might address this question in interviews collected for a qualitative study about WBA that we had conducted with our 3rd year student cohort in 2012 (22).

Methods

A secondary analysis was made of the transcripts from semi-structured interviews with students who had recently had three WBAs with verbal and written feedback during a four week clinical placement. The data had been collected for a study on the impact of grades in formative WBA (22). The methodology of the original study can be seen in that publication. In brief, all year 3 medical students and their GP tutors for the placements of 2012 were invited to participate. Maximal variation sampling was used when selecting interviewees with respect to gender, attainment in the recent OSCE and preference for feedback with and without grades. Semi-structured interviews were conducted at the end of the placement exploring what students felt about questions including their views on WBA. One researcher (JL) conducted the interviews. Interviews were audio-recorded with consent, and transcribed verbatim. The interview schedule can be seen in appendix 1.

Participants in the original study had been asked to consent for their data to be stored for future research. Ethical approval for this secondary analysis was granted by the Keele School of Medicine Ethics Committee.

In the secondary analysis, transcripts from the interviews held in 2012 with 24 Year 3 medical students were examined. The written feedback to 23 of these students was also examined (one student had not received written feedback).

We used a coding process to generate two subsets of data. Two researchers (JL and AH) independently re-read the interview transcripts to identify all references to *verbal/informal/feedback* and to *written/formal/GeCoS feedback* (GeCoS = Generic Consultation Skills – the assessment and feedback instrument used (20,21)). In a grounded secondary analysis, cross-comparison of the data within these two codes was made. Each of the researchers analysed student experiences and views about their feedback discussions with tutors and their written feedback summaries looking for direct comparisons made by students or interpreting the differences in their talk about each. The second dataset was the written feedback provided in the WBAs of consenting tutors. This was also examined to compare and contrast it with the students' descriptions of their informal verbal feedback. When the student recalled their feedback we checked to see whether the same terms were used in the written feedback. The free text comments in the written feedback were also checked for similarity with the phrases embedded in the GeCoS instrument.

Researchers were already familiar with the data following the previous study and were thus immersed within it. The recoding and the new lens of comparative analysis were used to make the familiar unfamiliar again. Following their independent analyses the researchers then met to discuss their findings. Any disparities in their findings were discussed to reach consensus.

Results

Four main themes relating to the research question were identified: the value of verbal feedback, written feedback being felt to be relatively superfluous, the augmentation effect on verbal feedback of having a written summary to produce, and the adoption of the language of the formal instrument in both verbal feedback and free text written feedback. These themes are expanded below.

Verbal feedback was valued

Verbal feedback on their consultations were seen by students as supportive, rich in content, immediate, specific and desirable.

F10: *My doctor was very good at (feedback) if he was sitting in on a consultation as soon as the patient had gone out we went through what was good, what was bad, what could be improved sort of thing and I think having that immediate feedback on it while it's fresh in your mind still erm I think that was really good. I think yes definitely the immediate feedback really helped with me cus then the next person who came in two minutes later, I was thinking about what he'd just said and I was bringing that into the next consultation straightaway.*

The feedback dialogue also gave more opportunity for the student to probe the tutor about what they meant by their feedback. For example this student was initially unhappy with his critical feedback but came to appreciate it after some discussion:

M1 *he seemed to notice at times that I was quite sort of annoyed maybe or a bit put down by the criticism I got and … he said things like oh …. I hope you don't think I'm being too critical of you or whatever. And we had a chat about it so I don't know if that influenced him to be, to give a bit more praise…. cus I was saying that he was comparing my consultation skills to sort of his very high standards that he has. Erm and he said oh who do you want to be compared to? And I said well I suppose, I think about comparing myself to the other members of my peer group and he just said, you shouldn't compare yourself to people in your group. You should be comparing yourself against like the highest standards you can be compared to.*

The friendliness of the verbal feedback was appreciated:

M7: *I felt he was being more almost as if he is telling from experience, as a friend – one person to another; almost like that information you give on the side kind of thing (laughing) so he was telling you the crucial details. He was like "really do this cus this is important"*

Written feedback was less valued

Most felt that the written feedback was superfluous except as a record of what had been spoken, and this record was not always welcomed.

Interviewer: *And how do you view workplace-based assessment? So those GeCoS assessments that you had?*

F23: *I thought they were a bit silly. (laughs) I just found it a bit odd. Because the doctor's given me feedback on my consultation skills anyway, I didn't feel I needed the GeCoS on top of it. It just put added stress into a situation which really didn't need any more stress.*

This student went on to explain her preference for informal feedback:
F23: *It was informal, there wasn't anything to fill in or any grades put on it or any… points to improve. It was just sort of said "Oh here's an idea, you should just look at this or go away and read this" rather than "this is what you have to do"…*

And of the stress:
F23: *just knowing that someone is … assessing you. And that it's going to be written down somewhere on record, I think makes it more stressful than just… being an informal thing that... only really you and the person in the room knows.*

They also felt that the written feedback was constrained by the assessment and feedback instrument used for the formal assessment.
F13: *I think it was mainly cus I had just presented these cases and the questions weren't… really relevant to the history that I'd just presented… I think we both found it quite difficult to tick boxes.*

The inability to get an immediate query about written feedback answered was a problem for a few students:
M1: *I thought oh what am I doing wrong. I must be doing quite a lot wrong. Erm so yes, I didn't really know what to think after that cus it was more, it was quite rushed going through it and... it wasn't discussed with him at the time. I received it via e-mail*

Several students had not engaged with their written feedback:
F13: *I can't remember off the top of my head but I did get given a form, summary form thing I think somewhere.*
There were exceptions to the opinion that verbal feedback was better than written, especially from students who wanted specific, well-thought-out feedback or those who wanted grades:
M22: *There was a lot more focus on what you can do next time, in a very clean bullet -pointed note. It wasn't a sort of wishy washy you know "Yeah you did OK but you should…" It was like "Do this next time" or "don't do that"… It's very constructive to know exactly what to do differently next time.*
M6 *I mean with feedback it's more of a like subjective thing so you'll just, you'll just hear you know a few words, you might forget them in a few weeks but with the grades like it's written down.*

Having to produce a written summary has the effect of augmenting feedback
Students reported that their clinical supervisors undertook multiple informal assessments with verbal feedback before completing each formal written assessment. There was often a three-stage feedback process of immediate post-consultation feedback, summary discussion and written advice. While they preferred verbal feedback to written, they acknowledged that their

assessors took great trouble over the formal assessment and often discussed it with them at length. This additional discussion and the written feedback felt qualitatively different from the immediate feedback – more constructive.

M12: *I got feedback straight after. Sort of just face to face, him telling me about the consultation… Then a bit more specific feedback - more detailed - particularly after the assessment. Just about history taking skills and questions I should ask or ways I could phrase things.*

And then I got more sort of written feedback quite soon after. It was always quite prompt.

And:

F21 *Before we went on the computer she sort of said what did you do well? What might you want to change in the future? And then she sort of gave me a brief thing as well and then we went on the computer and looked at the more specific answers we could have – the strategies.*

And:

M18 *I think the main thing was that she talked me through it, as opposed to me just reading it. So even though it was pretty thorough what she said, and it was really good on GeCoS but the fact that she talked me through was really helpful, and allowed me to – sort of - cos sometimes when you read something you have your own perceptions and you think that: Maybe she's saying this; Maybe she's saying that. But it sort of clarified to me what she actually meant and what she wanted me to say.*

Even this student who preferred the informal feedback could see that his tutors were adding value to it when they wrote it down:

M12 *I think perhaps because we'd always discussed it face to face in some form before getting the written feedback, you sort of could appreciate what was written there more because you've already discussed it.*

Personally I prefer face to face feedback. But I like that you've got a written record so you remember what's come back. You might not remember everything that's said to you when you have face to face feedback.

And I think when they're writing feedback more formal and written, they've had a bit more time to think about it. And any sort of constructive elements to it were always more in the written.

So when we were discussing face to face it was more sort of just noticing what had gone on and just talking a little bit about it. But when it was written it was always how to take that forward a bit more.

The language of the formal instrument was being adopted in informal verbal and free text feedback

What was notable in students' descriptions of their informal feedback was that it did correlate closely to their written feedback. This may indicate that the assessors were adapting to using the same terms in informal feedback discussions which they had been required to use in the formal assessment instrument. This is certainly the assumption made by student F11, and she approved, considering the feedback as more valuable because it aligned with their curriculum:

F11: *I think it was really good in the GP's because obviously they were kind of told what to do and they were told, they were given the guidelines on how to do it. So it was quite helpful… It was nice having someone watching me do it so that I could, you know, get good feedback.*

We were able to corroborate this as illustrated by these three examples from comparative analysis of the student interviews with their written WBA feedback:

Student F21 described the informal verbal feedback received:

F21: *After each patient left he told me what he thought of the history – the good bits and bits missed – and my examination.*

Interviewer: *What feedback do you recall?*

F21: *I asked multiple questions of the patient.*

In her written feedback summary the tutor selected the same concept of avoiding multiple questioning from the pre-formulated strategies for improvement (see figure 1):

Fig 1. Extract from written feedback to student F21

Opportunity for improvement:	
HISTORY PROCESS: Skilled use of questioning including open and closed questions	
GeCoS Strategy selected:	Don't use 'double' or 'nested' questions e.g. 'What is your pain like and how long have you had it?' 'Is your appetite normal and have you lost weight?'

Student F24's tutor used the language of the assessment instrument in free-text feedback reminding her of the previous week's selection from the GeCoS strategies for improvement in her feedback (See figure 2&3).

Fig 2: Week 3 WBA extract for student F24

Opportunities for improvement:	
OPENING: Establishes agendas	
GeCoS Strategies selected:	Identify the patient's agenda. Develop a range of opening questions for different situations with which you are comfortable
	Check that your understanding of the patient's agenda is complete: 'Is there anything else you would like me to do today'

Fig 3: Free text comment in WBA for student F24 week 4

> *"(You) reflected on last week's feedback and started to clearly establish patient's agenda at start of each consultation"*

In free text feedback to student M5 his tutor used the language of the assessment instrument to reinforce the relevance of the selected strategy to the student's need to improve (figure 4):

Fig 4. Week 3 WBA extract for student M5 and relevant GeCoS strategy

Opportunities for improvement:	
HISTORY: History content - Details of symptoms	
Free text comment: *"You missed a few key elements to the history such as weight loss or gain. Using SOCRATES would have helped"*	
GeCoS Strategy selected:	Use a mental checklist such as SOCRATES (which is useful for many symptoms) to clarify the presenting complaint(s)

Discussion

The secondary analysis of this data set has confirmed the high value placed on verbal feedback from their GP tutors by our students. This mirrors previous research in which verbal feedback was viewed more positively than written (3,8).

This preference for verbal over written feedback appears to relate to its immediacy and also to feedback as a social interaction,(9,10) which both work better face to face than in writing. For example, verbal feedback enables reciprocity of dialogue in which students can act quickly to resolve questions and conflict. By comparison one-way written feedback seems less satisfactory. Another socially important aspect described by some students is the intimacy of verbal feedback: it appears to be less threatening than the written (and potentially public) "assessment" even in this entirely formative situation. These data, however, suggest a paradox: written feedback may contain more useful advice than verbal feedback (M12 above) yet is still not preferred.

Understanding feedback as a social interaction and the value students place on the interaction may help us to resolve this paradox. We have demonstrated further social attributes of verbal feedback which make it preferable to written feedback such as keeping criticism private and allowing dialogue about it. In summary, verbal feedback was valued more highly than written

because of its richness, immediacy, intimacy and interactional effects which written feedback cannot match.

Notwithstanding this very clear preference for verbal feedback, these data suggest that the requirement to provide 'formal' written feedback may be augmenting 'informal' verbal feedback. The description given by students of the WBA process suggests three stages – immediate feedback, then a discussion in which feedback is summarised and strategies suggested, and finally receipt of the written feedback. The feedback discussion was linked by students to the requirement for formal assessment and although this qualitative data cannot prove that students in a similar clinical placement without the requirement for written feedback would get less feedback in total, the large quantity of feedback described here contrasts with the usual student and graduate complaint of insufficient feedback (23,24). In this study, the verbal feedback also appears to be aligned with the written feedback which in turn is aligned with the formal consultation skills curriculum (20,21) which we consider to increase its educational value. We consider that production of the formal written WBA using the online instrument to facilitate corrective feedback enabled tutors to give feedback aligned with the curriculum and in the language of medical education. It may also help to familiarise tutors in the language of consultation skills assessment which they then use in their feedback. Tutor engagement in feedback is key to its success (10). Requiring formal WBA will have ensured tutor engagement. Giving tutors the feedback tools and language may have encouraged this wealth of feedback. Without this realisation, the medical education intervention of mandatory written WBA might be seen as of limited value. We acknowledge that these two findings are tentative but, if they are confirmed, they extend knowledge about complexity of the feedback and the culture and environment in which it is embedded.

Elements of a learning culture which support the exchange of meaningful feedback have been identified as: a systems approach; supporting the development of trusting supervisor-trainee relationships; using video review with feedback; promoting communities of practice in which feedback is routine, regular and valued; making sure that the trainer's role and the learners' educational objectives are understood, and ensuring that competency in providing feedback is maintained and improved by reflection and refresher courses (25). The system of regular formal formative WBAs with training and a feedback instrument which we have introduced appears to have made an important contribution to the development of a such a learning culture across our network of 100 teaching practices(26). Sometimes you have to wait to see and then look for the benefits of an educational intervention.

Strengths and weaknesses
This is a secondary analysis of data. We consider in this case this is a strength. Data was collected to explore the impact on students of having grades with their written feedback. The interview schedule asked about recall of informal feedback to understand whether there was any difference in the discussions with tutors when grading was part of the process. This provided a good source of data for this subsequent research question.

The main weaknesses in this study are first that it is located firmly within one school which has invested heavily in its workplace assessment and feedback culture and second because it is a secondary analysis, we have not been able to examine causality. There were no parallel tutor interviews in that study to corroborate our inferences about staff development impacts of the WBA system and app. Whether or not these virtuous impacts would be driven by simply requiring the production of written feedback without a considerable parallel development of tutors is unknown (26).

Conclusions

Although the WBA was entirely formative for these students, they wanted to perform well and they reported that their tutors were engaged with helping them. Students recalled expansive and rich feedback and this was mirrored in their written feedback. The feedback process may have been enhanced by the requirement for written WBA. Furthermore, assessors have used the school's language for assessing and giving feedback on consultation skills. Using the school's app for producing written feedback summaries may have had a staff development effect of internalising the medical education language. Consequently students also had verbal feedback which was closely aligned to the formal curriculum. This at least in part explains the disparity in student satisfaction with verbal and written feedback. However, we consider that this is likely to be a positive unintended consequence of the requirement to complete three formal workplace-based assessments. Justification for medical education interventions needs to look beyond the obvious or immediate. Here is one example of a possible secondary beneficial effect of an intervention.

What we found:
- Tutors appeared to be engaged based on student report
- Language in the verbal and written feedback was aligned with the curriculum
- Students recalled expansive and rich feedback and this was mirrored in the written feedback
- The feedback process may have been enhanced by the requirement for written WBA
- The curriculum-aligned language may have been promoted by use of a feedback app

---

Feedback has a culture (7), a culture that it seems is open to influence in unexpected ways. The introduction of this system supporting mandatory serial formative WBAs seems to have changed the feedback culture in a beneficial way for our students.

We conclude that even if students never read their written feedback, the process of generating it is worthwhile and should not be abandoned. The practice of reflecting on the feedback (using the written summary as an aide-memoire) is likely to add to the formative outcomes but this remains to be proven.

149

Appendix 1

Interview guide for the study "Grades in formative workplace-based assessment - a study of what works for whom and why":

This interview is about your consultation skills. How did you get on in the Consolidation of Clinical Skills block?

How good are your consultation skills? How are your consultation skills progressing?

What is your self-reference? (What do you compare yourself to?)

What has affected your progress? (Helped? Hindered?)

How do you view Workplace-Based Assessment?

What happened in each of your 3 WBAs? (What feedback do you recall?)

What did you do with the WBA feedback? (What effect did the WBA feedback have on you?) (Where did you focus your attention after the feedback?)

Were there actions you could have taken arising out of the feedback you were given?

Why did you choose all grades/grades for your week 4 WBA/no grades for your week 4 WBA? How important was it to you?

Was there any difference in how you responded to feedback with and without grades? Why?

Is it helpful or unhelpful to be graded by the standard you should be on graduating from medical school?

Did the feedback resonate with you? Fair? Unfair? Too kind?

(What were the differences between your self-assessment and the grades, verbal and written feedback you got?)

How did the feedback and grades compare with previous assessments you have had?

Would it help to be allowed to choose whether to have grades or not with your future workplace-based assessments?

What difference would it make?

Would you want grades? Would they be good for you? What are the advantages and disadvantages?

Did you discuss the choice with your tutor? (What effect did that have?)

References

1. Norcini J, Burch V. Workplace-based assessment as an educational tool: AMEE Guide No. 31. Med Teach. 2007;29(9):855–71.
2. Hewson MG, Little ML. Giving feedback in medical education. Verification of Recommended Techniques. J Gen Inter Med. 1998;13:111–7.
3. Urquhart LM, Rees CE, Ker JS. Making sense of feedback experiences: a multi-school study of medical students' narratives. Med Educ. 2014;48(2):189–203.
4. Eva KW, Armson H, Holmboe E, Lockyer J, Loney E, Mann K, et al. Factors influencing responsiveness to feedback: on the interplay between fear, confidence, and reasoning processes. Adv Health Sci Educ Theory Pract. 2012;17(1):15–26.
5. Archer JC. State of the science in health professional education: effective feedback. Med Educ. 2010;44(1):101–8.
6. van de Ridder JM, Stokking KM, McGaghie WC, Ten Cate OTJ. What is feedback in clinical education? Med Educ. Centre for Research and Development, University Medical Centre Utrecht, Utrecht, The Netherlands. j.m.m.vanderidder@umcutrecht.nl; 2008 Feb;42:189–97.
7. Watling C, Driessen E, van der Vleuten CPM, Lingard L. Learning culture and feedback: An international study of medical athletes and musicians. Med Educ. 2014;48:713–23.
8. Bates J, Konkin J, Suddards C, Dobson S, Pratt D. Student perceptions of assessment and feedback in longitudinal integrated clerkships. Med Educ. 2013;47:362–74.
9. Hattie J, Timperley H. The Power of Feedback. Rev Educ Res. 2007;77(1):81–112.
10. Watling C, Kenyon C, Zibrowski E, Schulz V, Goldszmidt M, Singh I, et al. Rules of engagement: residents' perceptions of the in-training evaluation process. Acad Med. 2008;83(10 Suppl):S97-100.
11. Higgins R, Hartley P, Skelton A. Getting the Message Across: The problem of communicating assessment feedback. Teach High Educ. 2001;6(2):269.
12. Quitadamo IJ, Kurtz MJ. Learning to Improve: Using Writing to Increase Critical Thinking Performance in General Education Biology. CBE - Life Sci Educ. 2007;6:140–54.
13. Klein P. Reopening inquiry into cognitive processes in writing-to-learn. Ed Psychol Rev. 1999;11(3):203–70.
14. Ozuah PO, Reznik M, Greenberg L. Improving Medical Student Feedback With a Clinical Encounter Card. Ambul Pediatr. 2007;7(6):449–52.
15. Richards ML, Paukert JL, Downing SM, Bordage G. Reliability and Usefulness of Clinical Encounter Cards for a Third-Year Surgical Clerkship. J Surg Res. 2007;140(1):139–48.
16. Yarris LM, Fu R, LaMantia J, Linden JA, Hern HG, Lefebvre C, et al. Effect of an Educational Intervention on Faculty and Resident Satisfaction with Real-time Feedback in the Emergency Department. Acad Emerg Med. 2011;18(5):504–512.
17. Prystowsky JB, DaRosa DA. A learning prescription permits feedback on feedback. Am J Surg. 2003;185(3):264–7.
18. Bandiera G, Lendrum D. Daily encounter cards facilitate competency-based feedback while leniency bias persists. Can J Emerg Med. 2008;10(1):44–50.

19.   Bartlett M, Crossley J, McKinley R. Improving the quality of written feedback using written feedback. Educ Prim care. 2017;1:16–22.

20.   Lefroy J, Gay SP, Gibson S, Williams S, McKinley RK. Development and face validation of an instrument to assess and improve clinical consultation skills. Int J Clin Ski. 2011;5(2):77–87.

21.   Lefroy J, Thomas A, Harrison C, Williams S, O'Mahony F, Gay SP, et al. Development and face validation of strategies for improving consultation skills. Adv Heal Sci Educ. 2014;19(5):661–85.

22.   Lefroy J, Hawarden A, Gay SP, McKinley RK, Cleland J. Grades in formative workplace-based assessment: a study of what works for whom and why. Med Educ. 2015;49:307–20.

23.   van de Ridder JMM, Stokking KM, McGaghie WC, ten Cate OTJ. What is feedback in clinical education? Med Educ. 2008;42(2):189–97.

24.   Reddy ST, Zegarek MH, Fromme HB, Ryan MS, Schumann S-A, Harris IB. Barriers and Facilitators to Effective Feedback: A Qualitative Analysis of Data From Multispecialty Resident Focus Groups. J Grad Med Educ. 2015;7(2):214–9.

25.   Lefroy J, Watling C, Teunissen P, Brand P. Guidelines: the do's, don'ts and don't knows of feedback for clinical education. Perspect Med Educ. 2015;4(6):284–99.

26.   Bartlett M, Webb M, McKinley RK. The investment of a new medical school in its local primary care community. Educ Prim Care. Radcliffe Publishing Ltd.; 2015;26(2):89–94.

Chapter 9

Educational impact of electronic assessment processes: Utility of an app-based system to improve feedback following workplace-based assessment

Janet Lefroy

Nicola Roberts

Adrian Molyneux

Maggie Bartlett

Simon Gay

Robert McKinley

Keele University School of Medicine UK

Abstract

Objectives To determine whether an app-based software system to support production and storage of assessment feedback summaries makes workplace-based assessment easier for clinical tutors and enhances the educational impact on medical students.

Methods We monitored our workplace assessor app's usage by Year 3 to 5 medical students in 2014-15 and conducted focus groups with Year 4 medical students and interviews with clinical tutors who had used the apps. Analysis was by constant comparison using a framework based on elements of van der Vleuten's utility index.

Results The app may enhance the content of feedback for students. Using a screen may be distracting if the app is used during feedback discussions. Educational impact was reduced by students' perceptions that an easy-to-produce feedback summary is less valuable than one requiring more tutor time and effort. Tutors' typing, dictation skills and their familiarity with mobile devices varied. This influenced their willingness to use the assessment and feedback mobile app rather than the equivalent web app. Electronic feedback summaries had more real and perceived uses than anticipated both for tutors and students including perceptions that they were for the school rather than the student.

Conclusion Electronic workplace-based assessment systems can be acceptable to tutors and can make giving detailed written feedback more practical but can interrupt the social interaction required for the feedback conversation. Tutor training and flexible systems will be required to minimise unwanted consequences. The educational impact on both tutors and students of providing pre-formulated advice within the app is worth further study.

Introduction

Clinical tutors (or preceptors) often have difficulty in providing feedback to their trainees because of time constraints. The design of workplace-based assessment (WBA) tools can result in a greater focus on assessment than on feedback even when the intention is formative.[1–3] One way to manage time pressure and to promote constructive feedback is to assist clinical tutors to scaffold learning for their trainees by providing pre-written strategies for use in the event of various possible predicted competency gaps.[4] The challenge is to present this bank of strategies in a format which is accessible to busy clinicians, enhances the feedback conversation and enables the clinical tutor to add their own ideas. Ideally, a trainee, having had a feedback conversation with their clinical tutor, would be able to reflect on the feedback given, adopt strategies to improve their competence and monitor their own progress with validation and further advice from their clinical tutor.[5] A written summary of each feedback discussion is not essential but could enhance

the value of the assessment by providing an aide memoire for reflection and subsequent discussion, as well as documenting that the process has taken place.

The advent of mobile devices and supporting software systems has made it possible to create "apps" for reference and data-capture uses in almost any setting, including educational assessments. Electronic data capture of mini-CEX assessments via hand-held devices has already been successfully substituted for paper-based assessment of both doctors in training (called "junior doctors" in the UK or "residents" in North America) and medical students, improving timeliness and efficiency.[6–8] However, changing the format of an assessment and feedback tool is likely to alter the utility of the assessment in both expected and unexpected ways. For example, the amount of written feedback given decreased when using a hand-held electronic device vs paper-based in a comparative study of mini-CEX assessment of medical students.[9]

The utility of a system of assessment depends upon its reliability, validity, feasibility (or perhaps more appropriately, practicality), acceptability and educational impact.[10] Assessment in the workplace inherently has low reliability but high validity because it is situated in a real and variable clinical context and within the complexities of a social relationship between trainee and assessor.[11,12] It has been argued that the purpose of WBA should therefore be to understand how, why and what trainees are learning rather than attempting to "objectively" or "accurately" measure learning outcomes.[13] Educational impact should be an additional stated purpose, given the natural suitability of an assessment situated within the learning environment to improve performance.[14] In studying a system of assessment where the assessment and feedback tool and process are embedded within an app, important questions are whether or not its use is practicable and acceptable for WBA, enhances the feedback conversation and enables learning. For example, mobile devices in the hands of students are not always welcomed by staff and patients in the clinical workplace[15,16] yet a small-scale study with medical students in a remedial placement has suggested that student-held devices containing an app for the mini-CEX acted as an ice-breaker in their request for feedback from clinical tutors.[17]

Our study used these "later elements" of van der Vleuten's Utility Index[10] as a framework to investigate the utility of an app-based system for WBA and feedback across three clinical years of a medical school programme.

Research questions were:
How practical is it to use an app (mobile or web app) during WBA? (Feasibility)

How acceptable to tutors and students is using an app for WBA and what influence does it have on the feedback process?

What do tutors and students consider is the educational impact of this app-based system of WBA and feedback?

Methods

Approach to the study
This study is based in an ongoing action research project to embed and enhance WBA of our students while on clinical placements (or clerkships) at Keele University School of Medicine in the UK. Action research uses mixed methods to triangulate monitoring and evaluation data during programmes of change. This is done rigorously enough to be able to answer research questions and extend general knowledge, as well as solve problems for the local programme.[18–20]

Context – WBA developed by action research
At Keele University School of Medicine our WBA programme is entirely formative and does not inform progression decisions. Nevertheless, engagement with the process is mandatory. In primary care (general practice) placements in years 3, 4 and 5 of their undergraduate medical course students consult with patients under supervision and have three WBAs with feedback on their observed consultation skills during each placement. The assessor is the GP tutor who has observed them in practice that week. In secondary care, WBA by observation and feedback is currently optional and is generally done in speciality teaching clinics. Each student therefore has a minimum of 3 WBAs of patient encounters per year.

We have developed our WBA through a series of action research projects.
In one series, we developed an assessment tool (GeCoS – Generic Consultation Skills) which contains the 59 clinical encounter competencies expected of a graduating doctor in nine domains (Opening, History, Examination, Management, Clinical Reasoning, Building and maintaining the relationship, Organisation, Record Keeping and Case Presentation) which now underpin our consultation skills curriculum. The face-validated assessment tool and feedback suggestions are published for others to use. [4,21] We have been using the GeCoS competencies for both formative and summative assessment of consultation skills since 2010. The competencies have not been changed but we have reformatted them in various paper and electronic versions. The set of accompanying strategies for improvement was modified considerably and validated by medical student panels in 2012.[4]

We have also used action research to develop the WBA support systems which contain the GeCoS consultation competencies, the suggestions for how medical students can improve each competency and free text boxes for assessors to remind students about what they did and give additional advice on how to improve. The early iterations of the online WBA system required networked Internet access and generated a utilitarian feedback summary in an unattractive format, which students struggled (or neglected) to read. Despite this, we decided to continue generating written summaries because this seemed to be enhancing the quantity and quality of verbal feedback.[22] Wishing to improve the feedback system, we postulated that a handheld assessment device which supports audio recording (such as the student's or assessor's mobile phone or tablet) should facilitate the dialogue between tutor and student during WBAs. In addition it might be used to capture parts of that dialogue and save time. We considered that a

WBA platform that generates a feedback summary as a downloadable PDF should also improve the acceptability of the feedback summary to the student, thereby increasing its utility.

In Cycle 1 (2013-14) of the current Reflect-Plan-Act-Observe action research cycle, we developed, piloted and refined the consultation skills WBA and feedback system. This system comprised a web app and mobile apps for the two predominant mobile communication platforms, along with server-based infrastructure for collecting, processing, analysing and storing the completed assessments. The web app and two mobile apps perform the same function and each contain the GeCoS tool for assessment and feedback but present it differently to suit the format of the device being used. The apps are freely available online and in app stores, but require a Keele log-in to use. We are happy to provide a test student log-in on request.

Cycle 2 (2014-15) involved roll-out of the apps to all year 3, 4 and 5 medical students and their clinical tutors for all WBAs that year. Students received written summaries of the assessments via the School's online feedback portal.

Participants

Student participants in this part of the action research project were drawn from year 4 medical students who had experienced formative WBAs over the previous two years in both general practices and hospitals and had thus had experienced the "old" WBA system in their third year and the "new" app-based system in their fourth year. All students on two successive women's health blocks were invited to participate. These students had all had three WBAs during a four week block in general practice in year 3 and another three WBAs in a four week block in general practice in year 4. Some of them had also had WBAs in hospital teaching clinics in their women's health block. The students were assessed by numerous GPs, whereas the hospital assessments for these students were conducted by one of the authors (NR). Of the 32 year 4 students invited to two focus groups, 21 participated. Participants of focus group A had their GP block at the start of year 4 while those in focus group B had their GP block half way through the year and shortly before both focus groups were held.

In addition, tutors who had used the app four times or more during the study period were invited by email to consent to a telephone interview. Of 40 clinical tutors invited to interview, 11 volunteered and 10 (23%) were interviewed (7 GPs and 3 hospital doctors). One GP and one hospital doctor among the ten interviewed are authors of this study so their data is not quoted in this paper but their feedback about the app was valuable to the problem-solving side of the action research.

Monitoring data about the usage of both the web and mobile apps, together with the students' use of the feedback portal are securely stored in a relational database system housed within a Keele-based server. Usage data from all WBAs for year 3, 4 and 5 students was anonymised and monitored for this study.

Data collection and analysis

We used mixed methods: amalgamating quantitative utilisation data with the qualitative experience data from interview and focus group data.

The medical school's database was queried to show:

- o Numbers of WBAs carried out in GP practices and in hospitals.
- o Numbers of WBA feedback summaries created using the web app and the two mobile apps.
- o Per-click usage monitoring of the student feedback portal - how many times individual students had accessed their feedback summaries following each WBA.

We conducted two focus group meetings with year 4 medical students. 32 students were invited to attend the focus groups, which were facilitated by two research assistants who were not involved in the students' education. An agreed discussion guide was used in each focus group (Appendix 1).

Telephone interviews with clinical tutors were initially conducted by an independent research assistant. The final four interviews took place after the research assistant's contract had ended, and were therefore conducted by one of the researchers (JL) using the same interview schedule (Appendix 2).

Focus groups and interviews were audio-recorded and transcribed, and any identifying information removed. Analytic rigor was ensured by two researchers (JL and NR) independently coding the data before testing and achieving consensus. We then used an analysis framework based on the selected elements of van der Vleuten's utility index[10] and arranged our codes by constant comparative analysis of all the data that could inform each element.[23,24] Using open coding and memo-writing we searched for both expected and unexpected emergent themes in order to develop explanatory theories about the effect of the app-based system on the feedback process and content and, for the purposes of the next cycle of action research, how to improve the process and the technical features of the app.

Results

In the academic year 2014-15 the Keele Workplace Assessor app was used a total of 1581 times for conducting WBAs of 405 year 3, 4 and 5 students by 261 clinical tutors (248 in GP and 13 piloting it in a hospital setting). Of these, the web app was used 1339 times (85%) and the combined mobile apps were used 242 times (15%). Each tutor used the app between one and twelve times. Students accessed their WBA feedback portal to read their feedback summaries 992 times in the academic year 2014-15 accessing 63% of summaries produced (if each click was to a different summary). These monitoring data are represented in Fig. 1.

Fig 1: Monitoring data for usage of the new WBA and feedback system 2014-15
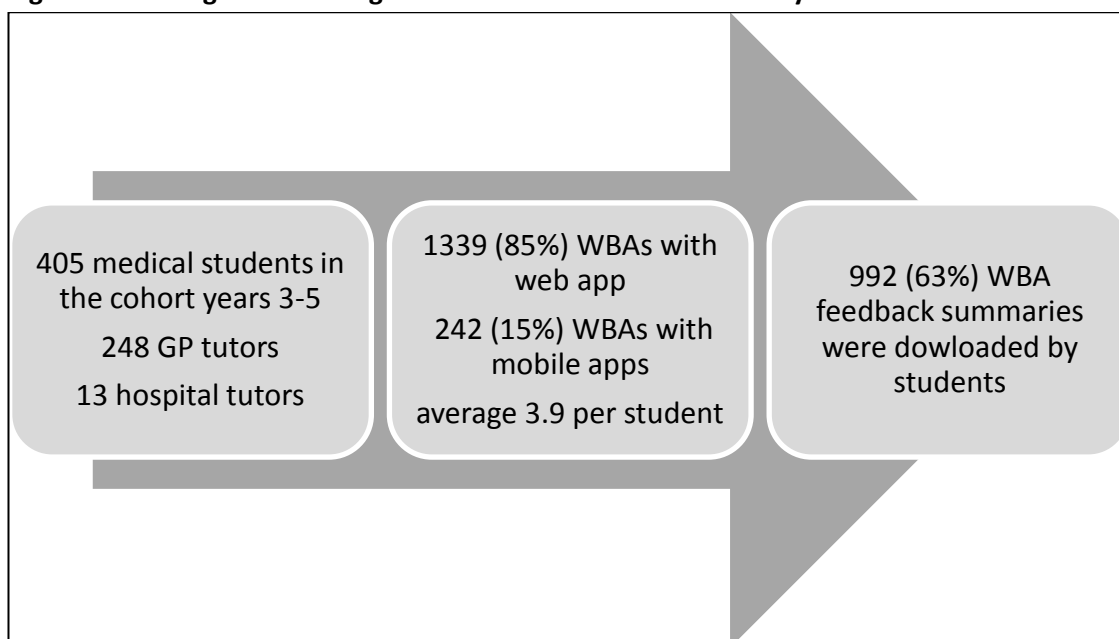
From the focus groups and tutor interviews we present evidence of influences of the app on the practicability and acceptability of WBA, on the feedback conversation, and perceived educational impact of this app-based system for formative WBA. Under each of these headings there were expected (or intended) and contrasting findings.

Feasibility of using an app (mobile or web app) during WBA

Table 1 shows examples of student and tutor comments about the accessibility of the mobile or web app. Students and those tutors who used it considered that the mobile app was easy to download and could be used anywhere; a particular benefit in hospital settings with competition for the available computers, also obviating the need to find the web app. Tutor comments suggested that if not using the same computer each time then the search for the email with the link to the web app made it a struggle to find. Using the mobile app was also found by some to be a faster process than using the web app (Comment 1). However, the vast majority of feedback summaries (85%) were generated using the web app rather than the mobile app. A few (described by students as 'younger') GPs were however using the mobile app and were dictating the text of their feedback. Students commented that the mobile phone screen size and typeface are too small. As typing on a mobile phone was not considered practical by students or tutors, it worked well only if the phone had speech recognition. Some tutors used their own tablet devices.

Students commented on the range of tutors' ability to use mobile devices and navigate the app. They felt that some tutors were looking to them for help implying that they needed further training (Comment 2).
Tutor preference for the web app over the mobile app was unexpected. GP tutors explained that they preferred to type, partly through habit but also because they found that they were more

able to think when typing which was helpful when they were trying to craft good feedback (Comment 3). We also infer from the interviews that this is because GPs may not have Wi-Fi or good cellular coverage but all have computers with Internet access. The inaccuracy of voice recognition was also a problem for some (Comment 4).

By contrast, those who expressed a preference for the mobile app found it faster to access and appreciated being able to dictate their free text comments (Comment 5). The mobile app was compared favourably to the web-based mini-CEX and similar WBAs for trainees by a hospital tutor, largely because of the ease of access and use of speech recognition.

Table 1: Examples of comments about the feasibility/accessibility of an app for WBA

Comment 1: *In the hospital it was easier having it on a phone than trying to find an empty monitor. You have to tailor it to the situation if you are with a younger GP or in the hospital where you haven't got much time you use the app on the phone. If you have got time you use the online version.* **Student 3(f) Focus Group A**

Comment 2: *My doctor is an older guy who couldn't use a phone. He gave it to me and said just tick that. So it wasn't an assessment.* **Student 1(m) Focus Group B**

Comment 3: *I'm used to writing reports …I prefer to reflect a little bit myself and just formulate what it is I want to write down because when I'm giving feedback there's a lot more communication going on than just verbal feedback… So when you're writing it in black and white, you have to be a little bit more careful of how you might phrase something.* **Tutor 5 (GP)**

Comment 4: *I don't dictate, because you're constantly checking the words, I find I actually type faster than I can speak, dictate and check.* **Tutor 4 (GP)**

Comment 5: *There's always that 'what am I going to say?' first, you think about it and then you say it. I'm so used to dictating in clinic so I can think what I want to say and it doesn't faze me.* **Tutor 10 (H)**

Comment 6: *You choose your three [strengths or priorities for improvement] which you want to make comment on and it takes you to those and so you're not having to screen scroll down pages of areas you don't want to comment upon.* **Tutor 6 (GP)**

Comment 7: *There's still the automatic kind of assumption - I can only really give a short feedback like I normally would. But once you kind of get comfortable to the actual dictation then I think it would help to give more detailed feedback. We give detailed feedback all the time verbally. We just don't give them any record of it.* **Tutor 10 (H)**

Comment 8: *Mine was really good in GP. She wrote everything down in each consultation and then summarised in the GeCoS.* **Student 2(f) Focus Group A**

The app (both web and mobile) was generally regarded by tutors as a time-saving and effort-

saving way of producing feedback (Comment 6) although habits formed on other feedback tools may shape the way this one is used (Comment 7).

Although few tutors completed the assessment using the app during student consultations, the majority used it to give an overview feedback summary afterwards (Comment 8).

Acceptability of the app for WBA and effect on the feedback conversation

Table 2 shows examples of student and tutor comments about the acceptability of the app for WBA, and how it was used in feedback. The mobile app was designed to be downloaded to the student's mobile device and handed over by the student when seeking feedback from a tutor. There were students and tutors who voiced discomfort about the student's mobile phone being used to capture feedback. They indicated that the mobile phone was private and not for others to handle (Comment 9). This had not been anticipated.

Most tutors used the app as a summarising tool for discussion that had already taken place and not as a teaching tool during the discussion. A barrier preventing it from enabling discussion was "screen distraction" - the need to look at a screen to search and type (Comment 10). The analogy of computers in patient consultations was used by tutor 2 who implied that tutors might learn to use the app during the feedback discussion (Comment 11).

Table 2:   Examples of comments about the acceptability of using an app for workplace assessment and its impact on the feedback conversation

Comment 9:   *Handing over your phone is awkward. Why do the med school expect you to use your (own) phone?* **Student 4(m) Focus Group A**

Comment 10: *After looking at them consult I would have given them quite a lot of feedback verbally and then I would use GeCoS to sort of back that up. I might have done it once with them sitting beside me but I found that I'm more concentrating on the computer. I'd rather be concentrating on the student in front of me so I wouldn't do it that way. It's an opportunity for them to ask questions.* **Tutor 5 (GP)**

Comment 11: *I think you spend more time trying to make the IT work than you do try to make the conversation work, you get distracted by it. It's a bit like a GP consultation and the role of a computer, the computer is there as an aid and it's not there to guide the consultation. Patients complain if you spend all the time looking at the screen, so students get frustrated if you're kind of there texting or whatever it is … the principle of it was brilliant and I am sure it can be made to work a bit better. You know, students live on their phones so that's great.* **Tutor 2 (GP)**

Comment 12:   *If we got feedback more often, like in that clinic, like if every placement had something like that, if doctors were more aware of the app it would be so useful for getting feedback, 'cause it'd just be like right we're going to do a quick GeCoS now on what you just did, and that'd be great.* **Student 3(f) Focus Group B**

Students commented that it was sometimes embarrassing to be present when the feedback was being given, and it might cause the written feedback to be less honest. There was some debate in

focus group B about whether they would prefer their feedback summary to be generated by their tutor when they were not present; some suspected the tutors would forget what had been discussed.

Although in its first year the app has been used mainly for mandatory WBAs in GP, it has been piloted in hospital teaching clinics. Student feedback about the impact that had on the feedback obtained was very positive (Comment 12).

Educational impact of the app-based GeCoS WBA system

<table>
<tr><td colspan="2"><u>Table 3: Examples of comments about the educational impact of the app-based GeCoS workplace assessment system</u></td></tr>
<tr><td>Comment 13:</td><td>Within each section there were various descriptors of what you would hope to see in good consulting and those are quite helpful to be able to read through with the student to sort of say these are things that I'm talking about Tutor 5 (GP)</td></tr>
<tr><td>Comment 14:</td><td>Did you find that your GP people would just tell you anyway? That you didn't really need the whole formality of (using the app)..?
Yeah but they'd forget … whereas if you'd got those bullet points in GeCoS they'd think oh actually you've done x, y and z really well, but I forgot to tell you about that. Students 2(f) and 3(f) Focus Group B</td></tr>
<tr><td>Comment 15:</td><td>We have lots of discussions erm and then I suppose one of the difficulties is you're having discussion and then trying to put that in a structured format within GeCoS.
Interviewer: So is GeCoS a constraint?
Yes I suppose in some ways it is but if you didn't have that structure then other areas of feedback might be missed. I don't know really. I can't see it working if you didn't have some formal assessment tool…. I think one of the downsides of it is that it's so comprehensive that actually sometimes it's quite hard to find the slot to put your feedback. Then it gives you a structure because otherwise you may end up just with a bit of waffle in a box which could easily turn into very sort of limited value for the student. Tutor 7 (GP)</td></tr>
<tr><td>Comment 16:</td><td>I would really like a truthful feedback rather than a tick box thing, nearly a letter from my GP just saying this is what I thought you did really well but this is what you've improved on.
It's finding the middle ground, isn't it? Something that will prompt them but doesn't then limit them
It's so easy for them to tick the box whereas if they have to physically write something, they have to think about what it is and if they can't think of anything they don't write anything, it's as simple as that. Students 4(m), 6(m) and 2(f) Focus Group B</td></tr>
</table>

Table 3 shows examples of student and tutor comments about the impact of this WBA system on the content of feedback. Students and tutors appreciated the structure of the GeCoS tool and the

162

guidance provided within the app because it prompted the giving of detailed feedback about both strengths and areas for improvement in a framework that was aligned with their curriculum (Comments 13 and 14).

The GeCoS app incorporates suggested text for giving students specific advice on each competency identified as needing improvement. This caused considerable concern to students particularly because they perceived the feedback tool as being too big and too difficult to navigate. Tutors interviewed also mentioned the size as being a problem and some saw the tool as prescriptive but could still see a benefit in terms of the structure and specificity it could give to their feedback which would make it more helpful for the students (Comment 15).

Students suggested that it might be too easy for tutors to tick boxes and that might result in them not thinking enough about what they needed to say, but could see that the app provided helpful guidance (Comment 16).

The value (educational and otherwise) of the app in capturing a written record

Table 4 shows examples of student and tutor comments about the value of creating a written feedback summary. Capturing feedback was seen as worthwhile by tutors (Comment 17), but to most students the written feedback summary was not seen as being of much value compared to the informal feedback discussions and some were unsure of its purpose, believing it to be feedback for the medical school (Comments 18 and 19). This is borne out by triangulation with the usage data from the student feedback portal – one third of feedback summaries were not accessed by students, meaning that they could not have read those written summaries. There was recognition, too, by tutors that the feedback discussion was more important than the written summary (Comment 20). The students who did value the written summary were those who got something more in writing than they had been given face to face, and those who liked to compare one with the next (Comments 21 and 22).

One unexpected value of the written feedback mentioned by both students and tutors was that it could be used to say things that were awkward to say face to face (Comments 23 and 24). Also unanticipated and less desirable was that some tutors prepared their feedback summary (because it was submitted electronically) with two recipients in mind – the student and the medical school, which seemed to influence the feedback they gave (Comment 25).

Some tutors also mentioned the value of having their written summaries of feedback to use as evidence of teaching or to remind them about the student if they were later asked by them for a reference.

Table 4: Examples of comments about the value of the app in capturing a written record

Comment 17: *I think it's a good idea because when I give feedback to students it's not recorded, I just say what I think and they nod away and they say thank you very much, that was useful, but it's not recorded in any way…They can't remember everything I say because it's quite a lot.* **Tutor 3(H)**

Comment 18: *I haven't even opened the GeCoS feedback emails because I sat down and did it with him so I knew exactly what he said.* **Student 4 (m) Focus Group B**

Comment 19: *If you do well in GeCoS it doesn't seem to matter? And if you do badly what does it matter? What is the point of it then?* **Student 6 (m) Focus Group B**

Comment 20: *At the end of the day the tool provides a record and it's great if the students have a record afterwards so they can go back and look at the points but that is a very small part of the interaction, the main part of the interaction is what goes on as you have that meeting.* **Tutor 2 (GP)**

Comment 21: *My GP in 3rd year didn't tell me when he was doing them and I only found out afterwards and that could have been really bad but I found it really useful because he was really honest. At the time he had told me things, but (the written feedback) was really detailed because he had gone and done it in his own time.* **Student 5 (f) Focus Group A**

Comment 22: *Mine was really good. He did a GeCoS with my first consultation and then every GeCoS after that he referred back to it.* **Student 3 (f) Focus Group A**

Comment 23: *I think you always get better feedback if you're not there. They can't say the negative stuff when you are sat there.* **Student 2 (f) Focus Group B**

Comment 24 : *There was one that I had erm, maybe it was a bit of a cop-out but he was quite tricky, quite difficult to talk with so I used the GeCoS feedback as a means of being a bit more direct in terms of some advice and feedback for him and then he came back to me on it and we had a much more open discussion about it. So in that sense it actually worked quite well 'cause once it's written down on paper he took a bit more notice of it.* **Tutor 7 (GP)**

Comment 25: *You feel that somewhere at the top there's people collating data and they want you to fit into that box.* **Tutor 6 (GP)**

Discussion

We studied aspects of the feasibility, acceptability and educational impact of an app-based system to support the production of written summaries of formative WBA for undergraduate medical students. This framework for analysis was used to find both expected and unexpected themes. We expected that the system would enhance the feasibility and acceptability of

formative WBA and produce more useful feedback. Unexpected emergent themes were deliberately sought in order to develop our understanding of the impact of new technologies on existing medical education practice.

As expected, tutors perceived both web and mobile versions of the app to be time-efficient and helpful when crafting useful feedback. The requirement to use WBA feedback apps three times in a placement did appear to promote the importance of formative assessment and empower tutors to give detailed and specific feedback. Tutors were confident that they would find appropriate supporting text in the GeCoS tool to help them to create the summaries.

Students were less impressed by the embedded strategies, viewing them as too easily selected and therefore of doubtful personal relevance while recognising the alternative might be to receive less feedback. Students valued highly the free-text feedback provided by their tutors, especially when they perceived that the tutor had put a lot of effort into it. This suggests an unexpected negative feature of an app developed specifically for time-efficiency, if the value placed by students on their feedback is a function of their perception of the effort required to produce it.

Contrary to expectation, we found a preference for the web app amongst GP tutors. Incorporation of the dictation facility in the mobile app was not as appealing as we had expected: those who tried dictating using speech recognition facilities on a mobile device liked the speed but some were concerned about its accuracy. Preference for the web app was explained in three ways: either GP tutors were accustomed to using computers, *or* they wanted to consider what they committed to writing *or* they were reluctant to use the student's device. Tutors who are used to touch screens and speech recognition are starting to use the mobile app in hospital and GP. It may become more acceptable as it becomes more familiar.

As to acceptability, the face-to-face feedback discussions were generally highly valued by students, but some reported that being present when their feedback summary was generated was uncomfortable. This was unexpected and seemed to relate to witnessing their tutor struggle to use the app or waiting passively while they were typing. Some students disliked their mobile phone being used as an educational tool by their tutor, a barrier we had not anticipated.

The app does facilitate learning, but not as envisaged. Although it was designed to enhance the feedback conversation, in the first year of using the app, the majority of tutors have used it at the point of generating feedback summaries rather than using it as a teaching aid during the feedback dialogue. Our previous research suggests that the requirement to create a feedback summary may enhance the feedback dialogue. [22] In this study, tutors appreciated that the app-based assessment and feedback tool provides structure and curriculum-aligned advice. We suggest that this could influence the feedback dialogue as tutors become familiar with the students' curriculum through repeated use of the app. Students were less sure than tutors of the utility of the app and only a minority felt that the written feedback summary added value. This was more likely if the student was not present when the written summary was generated.

This study contributes to the literature about the utility of apps in clinical WBA in two ways. Firstly, plurality of platform is important: our provision of both mobile and web apps for the same assessment and feedback tool showed that, given the choice, tutors tend to work on a platform

with which they are familiar. This echoes previous comparisons of paper-based and electronic marking, when for example, few used the new electronic assessment system with their trainees.[25] We have found that it is important that an app-based system is flexible enough to accommodate different experience and skills in typing, dictating and incorporating technology into work and life. This implies that future clinical tutors who have grown up with mobile devices and are "digital natives"[26] will use technology differently than the current generation of tutors who are mostly "digital immigrants".[27]

Secondly, the study gives insight into the complexity of introducing an electronic agent into a social interaction, the feedback conversation. Electronic devices in patient consultations such as a mobile device or computer screen become a third and intrusive party.[28–31] In the same way, the assessor's focus on the screen can subvert the feedback discussion, even though it can enhance the content. Training in the use of electronic devices may improve the feedback process, as it has done for the doctor-patient-computer consultation.[32]

Limitations of this study

Action research is real and messy. It facilitates improvement of a system but findings relevant in one setting must be applied with caution in different contexts. Nevertheless we consider that the lessons we have learned are generalizable to the implementation of other app-based systems in other settings.

While both tutor and student stakeholders evaluated the system, the number of participant tutors was smaller than students. The students in focus group A all had their workplace assessments at the start of the academic year and group B later in the year: this provided some insight into the tutor's learning curve. Tutors who volunteered to be interviewed may have had stronger positive or negative feelings about the app than those who did not.

Though a change of interviewer, particularly to one who might be perceived to have a vested interest in the outcome of the evaluation (JL), might be viewed as a limitation, there were no obvious differences between the resultant themes of the interviews by the independent researcher and those interviewed later by JL, although the interviewees did ask questions of JL about the app.

The study was conducted in a single school with a 'bespoke' solution to its assessment support.

Implications for practice and next steps

Multiple interfaces are needed when setting up a system of electronic WBA. Mobile and web apps suit different environments (with varied computer and Wi-Fi availability) and people (with varied competencies). While the dictation facilitation was useful, some tutors prefer to type their feedback. In setting up such a system, it is worth profiling the users for their familiarity not only with mobile technology, but also being a "native" at typing or dictation in their daily work, as such preferences are strong.

A written summary has multiple roles and students are not the only beneficiary of their feedback discussion being captured in writing: it also benefits tutors by educating them about the students' curriculum, evidencing their teaching and aiding their recall for the future.

Co-completion of the summary of a feedback discussion has advantages and disadvantages. Although feedback should be a conversation between student and tutor, the incentive to read a summary of that conversation may be removed by the feeling that it contains nothing unknown.

Provision of preformed strategies for improving clinical skills has a downside as well as benefits. The ease of selection, which is an attractive feature to tutors, makes students suspicious of their personal relevance because convenience may have bypassed thought. This broadens the canvas for staff training to include information-sharing skills for formative assessors and how to use standardised materials in a customised fashion.

In conclusion, medical educators are adapting to the digital era but interactions with learning and assessment systems risk interrupting the inherent social interactions in education. Continued successful integration of technology in medical education will require carefully planned training and mentoring and systems sufficiently flexible to cope with the subtle demands placed upon them.

Conflicts of interest
The authors declare that they have no conflict of interests.

Ethics
These monitoring and evaluation activities were approved by the Keele School of Medicine Ethics Committee on 26.1.14.

Contributions (not included in the published paper)
JL and AM convened the action research group to plan this cycle of development and evaluation of workplace assessment. AM designed the workplace assessor app. NR coordinated student focus groups. JL coordinated tutor interviews and conducted the final few interviews. JL and NR conducted the initial analysis of focus groups and interviews. AM collected and analysed the monitoring data about usage of the system. All authors contributed to interpretation of the data. JL wrote the first draft of the paper. All authors contributed to revisions.

References

1. Holmboe E, Yepes M, Williams F, Huot S. Feedback and the mini clinical evaluation exercise. J Gen Intern Med. 2004;19:558-561.

2. Saedon H, Salleh S, Balakrishnan A, Imray CH, Saedon M. The role of feedback in improving the effectiveness of workplace based assessments: a systematic review. BMC Med Educ. 2012;12(25):1-8.

3. Setna Z, Jha V, Boursicot KAM, Roberts TE. Evaluating the utility of workplace-based assessment tools for speciality training. Best Pract Res Clin Obstet Gynaecol. 2010;24(6):767-782.

4. Lefroy J, Thomas A, Harrison C, Williams S, O'Mahony F, Gay SP et al. Development and face validation of strategies for improving consultation skills. Adv Health Sci Educ Theory Pract. 2014;19(5):661-685.

5. Lefroy J, Watling C, Teunissen P, Brand P. Guidelines: the do's, don'ts and don't knows of feedback for clinical education. Perspect Med Educ. 2015;4(6):284-299.

6. Torre DM, Simpson DE, Elnicki DM, Sebastian JL, Holmboe ES. Feasibility, reliability and user satisfaction with a PDA-based mini-CEX to evaluate the clinical skills of third-year medical students. Teach Learn Med. 2007;19(3):271-277.

7. Ferenchick GS, Solomon D, Foreback J, et al. Mobile technology for the facilitation of direct observation and assessment of student performance. Teach Learn Med. 2013;25(4):292-299.

8. Mooney JS, Cappelli T, Byrne-Davis L, Lumsden CJ. How we developed eForms: An electronic form and data capture tool to support assessment in mobile medical education. Med Teach. 2014;36(12):1032-1037.

9. Torre DM, Treat R, Durning S, Elnicki DM. Comparing PDA- and paper-based evaluation of the clinical skills of third-year students. Wisconsin Medical Journal. 2011;110(1):9-13.

10. van der Vleuten CPM, Schuwirth LWT. Assessing professional competence: from methods to programmes. Med Educ. 2005;39(3):309-317.

11. Holmboe ES, Huot S, Chung J, Norcini J, Hawkins RE. Construct validity of the miniclinical evaluation exercise (miniCEX). Acad Med. 2003;78(8):826-830.

12. Norcini JJ, Blank LL, Duffy FD, Fortna GS. The Mini-CEX : A Method for Assessing Clinical Skills. Ann Intern Med. 2003;138(6):476-481.

13. Govaerts M, van der Vleuten CP. Validity in work-based assessment: Expanding our horizons. Med Educ. 2013;47(12):1164-1174.

14. Norcini J, Burch V. Workplace-based assessment as an educational tool: AMEE Guide No. 31. Med Teach. 2007;29(9):855-871.

15. Robinson T, Cronin T, Ibrahim H, et al. Smartphone use and acceptability among clinical medical students: a questionnaire-based study. J Med Syst. 2013;37(3):9936.

16. Ellaway R. The informal and hidden curricula of mobile device use in medical education. Med Teach. 2014;36(1):89-91.

17. Coulby C, Hennessey S, Davies N, Fuller R. The use of mobile technology for work-based assessment: The student experience. Br J Educ Technol. 2011;42(2):251-265.

18.   Meyer J. Action research. In: Pope C, editor. Qualitative Research in Health Care. London: BMJ books; 1996.

19.   Zuber-Skerritt O. New Directions in Action Research. London: Falmer; 1996.

20.   McKay J, Marshall P. The dual imperatives of action research. Information Technology & People. 2001;14(1):46-59.

21.   Lefroy J, Gay SP, Gibson S, Williams S, McKinley RK. Development and face validation of an instrument to assess and improve clinical consultation skills. International Journal of Clinical Skills. 2011;5(2):77-87.

22.   Lefroy J, Hawarden A, Gay SP, McKinley R. Positive impacts of written feedback: Does formal workplace based assessment add value to informal feedback? 2017 (cited 9.2.17) available from: https://www.mededpublish.org/manuscripts/832/v1

23.   Glaser B. The Constant Comparative Method of Qualitative Analysis. Social Problems.1965;12(4):436-445

24.   Kennedy TJ, Lingard LA. Making sense of grounded theory in medical education. Med Educ. 2006;40(2):101-108.

25.   Dudek NL, Papp S, Gofton WT. Going Paperless? Issues in Converting a Surgical Assessment Tool to an Electronic Version. Teach Learn Med. 2015;27(3):274-279.

26.   Prensky M. Digital Natives, Digital Immigrants Part 1. 2001 (cited 1.4.17) available from: http://marcprensky.com/marcs-essays/#students

27.   McAlearney AS, Schweikhart SB, Medow MA. Doctors' experience with handheld computers in clinical practice: qualitative study. BMJ. 2004;328(7449):1162.

28.   Pearce C, Trumble S, Arnold M, Dwan K, Phillips C. Computers in the new consultation: Within the first minute. Fam Pract. 2008;25(3):202-208.

29.   Kazmi Z. Effects of exam room EHR use on doctor-patient communication: A systematic literature review. Inform Prim Care. 2013;21(1):30-39.

30.   Alkureishi MA, Lee WW, Lyons M, et al. Impact of Electronic Medical Record Use on the Patient-Doctor Relationship and Communication: A Systematic Review. J Gen Intern Med. 2016;31(5):548-560.

31.   Alsos OA, Dabelow B, Faxvaag A. Doctors' concerns of PDAs in the ward round situation: Lessons from a formative simulation study. Methods Inf Med. 2011;50(2):190-200.

32.   Booth N, Robinson P, Kohannejad J. Identification of high-quality consultation practice in primary care : the effects of computer use on doctor – patient rapport. Inform Prim Care. 2004;12:75-83.

Appendix 1

<u>Focus Group Facilitator Guide</u>

INTRODUCTORY QUESTIONS

- What do you understand by WBA?
- What are your thoughts about WBA?

TOPICS FOR EXPLORATION

- General exploration of the app
    - Have you made use of the WBA app?
    - Thoughts on the app?
        - Ease of use?
        - Ease of access?
        - Quality of the app?
    - How is the app being used?
        - When is it accessed?
        - How often is the app accessed?
        - Is it used as a resource or repository?
        - Is text dictated or typed?
- Usefulness of the app to students
    - Is the app more useful than the previous online method?
        - If so, how and why is the app more useful?
    - Does the app improve the quality of feedback
    - How could the app be made more useful?
- Outcomes of the WBA
    - What do students do with the feedback they receive?
        - Ignored or acted upon?
        - Value placed upon the feedback?
        - Does the app influence how the feedback is used?

CONCLUSION

- Are there any additional comments anyone would like to make regarding our discussion around WBA?

Appendix 2

Guide for telephone interviews of clinical tutors

QUESTIONS

1. What do you think of your workplace based assessments using the new GeCoS or LCAT app (ask for any other positive and negative views after they have said what they want to) (ask how many GeCoS assessments they have had last year and this year so far)
   o Thoughts on the app
     ▪ User-friendly?
     ▪ Ease of access?
   o How is the app being used?
     ▪ When is it accessed?
     ▪ Whose device is being used (the student's or the assessor's? mobile device or computer)
     ▪ Is it used as a resource for feedback content or as a repository after the feedback discussion?
     ▪ Is text dictated or typed?
   • Usefulness of the app to students
     o How does feedback with the GeCoS (and LCAT) app compare with the previous online process?
     o Does the app improve the quality of feedback?
     o How could the app be made more useful?
2. Do you think using the app altered the content of the feedback you give compared to informal feedback discussions? (quantity? Detail? Alignment with curriculum?)
3. Any other comments?

Chapter 10

Guidelines on feedback for clinical education
The Do's, Don'ts and Don't Knows of feedback
for clinical education

Janet Lefroy[1]

Chris Watling[2]

Pim W Teunissen[3]

Paul Brand[4]

[1]Keele University School of Medicine, Staffordshire, UK
[2]Schulich School of Medicine and Dentistry, Western University, Ontario, Canada
[3]Maastricht University and VU University Medical Center, Amsterdam, The Netherlands
[4]Isala Klinieken, Zwolle, The Netherlands

Abstract

Introduction   The guidelines offered in this paper aim to amalgamate the literature on formative feedback into practical Do's, Don'ts and Don't Knows for individual clinical supervisors and for the institutions that support clinical learning.

Methods   The authors built consensus by an iterative process. Do's and Don'ts were proposed based on authors' individual teaching experience and awareness of the literature, and the amalgamated set of guidelines were then refined by all authors and the evidence was summarized for each guideline. Don't Knows were identified as being important questions to this international group of educators which if answered would change practice. The criteria for inclusion of evidence for these guidelines were not those of a systematic review, so indicators of strength of these recommendations were developed which combine the evidence with the authors' consensus.

Results   A set of 32 Do and Don't guidelines with the important Don't Knows was compiled along with a summary of the evidence for each. These are divided into guidelines for the individual clinical supervisor giving feedback to their trainee (recommendations about both the process and the content of feedback) and guidelines for the learning culture (what elements of learning culture support the exchange of meaningful feedback, and what elements constrain it?)

Conclusion   Feedback is not easy to get right, but it is essential to learning in medicine, and there is a wealth of evidence supporting the Do's and warning against the Don'ts. Further research into the critical Don't Knows of feedback is required. A new definition is offered: Helpful feedback is a supportive conversation that clarifies the trainee's awareness of their developing competencies, enhances their self-efficacy for making progress, challenges them to set objectives for improvement, and facilitates their development of strategies to enable that improvement to occur.

> **Do's -** educational activity for which there is evidence of efficacy
>
> **Don'ts** – educational activity for which there is evidence of no efficacy or of harms (negative effects)
>
> **Don't Knows** – educational activity for which there is no evidence of efficacy

Introduction

Feedback is considered of utmost importance for learning. Despite the importance of feedback and the attention it has received in scholarly literature, effective feedback remains difficult to achieve within the context of clinical education. The guidelines offered in this paper aim to

amalgamate the literature on formative feedback into practical Do's, Don'ts and Don't Knows. The guidelines relate to formative feedback (i.e. exchange of information with the intent to support development) in clinical education (medical students and doctors learning in the workplace), but are also relevant to formative feedback associated with a summative assessment.

We have not attempted a systematic review of the considerable and growing body of literature on feedback in medical education. Rather, we offer recommendations based on published evidence from scientific exploration of the feedback process, and on our combined experience and study in this area. Below, we list the *Do's, Don'ts,* and *Don't Knows*. In the supporting paper that follows we briefly articulate what we regard as the key evidence for each Do and Don't we have listed. In the summary (Table 2) we indicate the strength of this evidence and therefore of our recommendation using the criteria outlined in Table 1.

It is not easy to know what feedback will be useful to a trainee. There is a recognized feedback gap (between feedback given and what is received by the trainee (1)). What this means for supervisors is that delivering feedback without first diagnosing our trainee's need and receptiveness risks wasting effort. The impact of formative feedback will depend on the strength of the trainee's desire to improve and their confidence in their ability to do so (2). To some extent these are personality traits (innate or learned earlier in life) but they can change with the trainee's situation and we need to know how to promote both.

The guidance we have compiled is intended not only for clinical supervisors, but also for learners and for the institutions that support clinical learning. We suspect that one of the reasons that the quality and quantity of feedback has not improved greatly despite all the years of scrutiny and the libraries of words written about it is that the focus has been largely on how supervisors as individuals should construct and deliver feedback, with considerably less attention directed to how learners receive and respond to feedback, and to how institutions can create a culture in which feedback works. Clinical tutors may not be averse to giving useful feedback, but they may operate in an environment that limits their opportunity to do so. Learners want feedback, but they may be motivated more by competition for status or fear of failure than by the desire to improve as a clinician. Overcoming barriers to meaningful feedback demands both individual and institutional efforts. We therefore include a set of Do's and Don'ts regarding the learning culture which are directed primarily at institutions wishing to promote feedback, in addition to our guidelines for the individual supervisor. We hope that by setting out the known Do's and Don'ts and by encouraging study of the many Don't Knows about feedback within our complex systems of clinical coaching, we can provide direction for these important efforts.

Terminology

The term 'trainees' is used for both undergraduate and postgraduate learners, but where the stage of training is thought to influence the giving or receiving of feedback this is specified.

Table 1. Criteria for strength of recommendation

Strong	A large and consistent body of evidence
Moderate	Solid empiric evidence from one or more papers plus the consensus of the authors
Tentative	Limited empiric evidence plus the consensus of the authors

Table 2: Summary of guidelines with strength of recommendation

For the individual clinical supervisor giving feedback

Do's for the process of feedback	Strength
1. Do realize that feedback is not just one person providing information to another to help them improve. Feedback is part of a social interaction influenced by culture, values, expectations, personal histories, relationships, and power. Do treat feedback as a conversation rather than as a commodity	Strong
2. Do recognize that trainees must perceive feedback as credible in order for it to be influential. Credible feedback is well-informed, typically by direct observation of the task or event, and it comes from a trustworthy source. Make sure that you as supervisor set a good example as a credible role model	Moderate
3. Decide the timing of feedback depending on the competence level of the trainee and on the complexity of the task	Moderate
4. Do encourage trainees to look for feedback and use it to enhance their performance	Moderate
Do's for the content of feedback	
5. Do tailor bespoke feedback to the individual trainee. The trainee might benefit from: –Reinforcement of key points done well –Identification of key points which might have been done better or omissions –Working out strategies for improving the quality of their work –An increased self-awareness	Strong
6. Do give specific feedback, focused on how the task was done and how that type of task should/might be done	Strong
7. Do make sure to indicate whether feedback is about necessary improvement for minimally acceptable performance or whether it is a reflection on possible variations to build upon adequate performance - Consider offering grades as an element of formative feedback if it seems that receiving grades will enhance the seeking of strategies for improvement. Conversely, avoid giving grades to trainees who you suspect will stop trying to learn if they get a good enough grade and to those who will give up if they get a poor grade	Tentative
8. Do ensure that feedback is actionable, enabling the trainee to construct strategies for improvement. After discussing the trainee's performance of a task, provide some guidance or 'scaffolding' to enable them to step beyond their current competence	Strong
9. Do attend to trainee motivation when discussing strategies for improvement	Moderate
10. Regardless of the specific approach to feedback that is used, do engage the trainee	Tentative

in a reflective conversation that marries their self-assessment with your observations and elaborations Several approaches have been described in the literature (sandwich, Pendleton, reflective feedback conversation, agenda-led outcome-based analysis, feedforward), but no single approach has been established to be the most effective. Rather, the likely best approach varies according to the learner, the teacher-learner relationship, and the context	
Don'ts	
11. Don't assume that a single approach to feedback will be effective with all trainees or in all circumstances. As the players and the contexts change, so too does the most useful approach to feedback. Don't assume:–You know what a trainee wants to learn–You know why a trainee is struggling –You know if or why a trainee wants feedback – You know what information a trainee takes out of a situation or feedback conversation	**Moderate**
12. Don't provide feedback without follow-up. Trainees are unlikely to be influenced by feedback that is not followed by an opportunity for them to demonstrate improving performance	**Moderate**
13. Don't provide feedback that is poorly informed (or is based on hearsay); doing so diminishes the value that trainees assign to feedback in general	**Moderate**
14. Don't underestimate the emotional impact of feedback that is perceived as negative. Emotional distress may be a barrier to acceptance and use of feedback	**Moderate**
15. Don't give grades without explaining the criteria for allocation of grades and providing strategies for improvement	**Moderate**
Don't knows	
16. What determines the credibility of feedback?	
17. How much is the right amount of content when giving feedback?	
18. What determines the 'open and safe interaction' in the feedback conversation?	
19. What influences the trainee's response? (constructive or destructive outcomes)	
20. Is overt comparison with peers—when made by the supervisor—helpful to the trainee? Indeed, is overt comparison with required performance standards helpful?	
21. Does a written summary of the feedback discussion enhance learning?	

Guidelines for the learning culture (what elements of learning culture support the exchange of meaningful feedback, and what elements constrain it?)

Do's	Strength
22. Do have a systems approach, building feedback into the learning processes	**Moderate**
23. Do support the development of longitudinal, trusting supervisor-trainee relationships in medical training; influential feedback thrives in the context of trusting relationships	**Moderate**
24. Do use video review with feedback as a component of training	**Tentative**
25. Do promote communities of practice in clinical workplaces in which feedback is routine, regular and valued	**Moderate**
26. Make sure that those who have a formal role in a workplace's educational system are aware of that role and understand what learners' educational objectives should be	**Moderate**
27. Make sure that the team give feedback regularly, reflect on the practice of giving feedback, and follow refresher courses to maintain and improve competency in providing feedback	**Moderate**
Don'ts	
28. Don't rely exclusively on faculty development to improve the effectiveness of feedback.	**Moderate**
29. Don't allow formal assessments of clinical skills, such as the mini-CEX, to be completed without observation and feedback	**Moderate**
Don't knows	
30. What are the vital components that ensure a constructive system of workplace learning that caters to trainees, workers, and the educational system? How can the institution nourish a climate which encourages the provision and seeking of feedback?	
31. Is it most effective to give feedback to individuals alone or in a group setting?	
32. Does the use of formative assessment outcomes for summative purposes (such as having supervisors provide formative feedback that at the end of a rotation is also used for a summative assessment) corrupt a well-intentioned educational system?	

Foundation paper

Methods and 'way of working'

The authors built consensus by an iterative process. Following an initial discussion to agree on the scope of the guidance and the criteria for selection of guidelines (see below), each author independently listed their Do's, Don'ts, and Don't Knows. These were amalgamated by JL and discussed for clarification where there was obvious conflict of Do's and Don'ts. A lead author was identified for each of the compiled list who would provide an initial outline of the evidence. The compiled table of Do's, Don'ts, and Don't Knows with supporting evidence was then circulated for all authors to add evidence and comments. Where we considered that evidence was still conflicting or there was not a clear consensus following consideration by all authors, items which had been thought clear Do's or Don'ts were moved into the Don't Know section. This process was repeated once more for final agreement and the strength of each recommendation was determined by consensus.

The criterion for identifying a Do or a Don't was that it was considered important to us as medical educators with our individual teaching experience and awareness of the literature. We did not set out to perform a systematic review of the large and growing body of literature on feedback in medical education. The range of undergraduate to postgraduate education teaching and research experience we had across three countries' health systems led us to believe that we could compile useful guidelines. The Don't Knows were identified as being important questions to this international group of educators which if answered would change practice.

Since the criteria for inclusion of evidence for these guidelines were not those of a systematic review, we avoided using the ABC indicators of strength and devised our own indicators which combine the evidence with the authors' consensus (see table 1).

Results

In the initial discussion of the scope of the guidelines, it became clear that while we could provide guidance to individual clinical supervisors wishing to give better feedback, the impact of that guidance would be limited if they were working within a system that didn't actively promote feedback as a way of improving. We therefore determined to divide our guidance into that for the individual clinical supervisor giving feedback and for those in positions of influence over the feedback culture in training systems and workplaces of medical students and doctors.

Our initial list of Do's, Don'ts, and Don't Knows numbered 65. We reduced this to the 32 listed largely by amalgamation with only two being dropped as unimportant on group reflection.

Items which had been thought clear Do's or Don'ts but after examining the conflicting evidence were moved into the Don't Know section were: Is comparison with peers helpful? Is comparison with required performance standards helpful? Can the same people give summative and formative feedback? (item no.32).

Conflicts of individual authors' Do's and Don'ts arose over whether trainees benefit from receiving grades with formative feedback. The consensus was that this may be useful to some trainees and a tentative recommendation was included.

The background evidence to each guideline is described and referenced in the following paragraphs. Table 2 is annotated with our judgement on the strength of our recommendation based on that evidence.

Background evidence to guidelines for the individual clinical supervisor giving feedback

The process of feedback

1. **Do realize that feedback is not just one person providing information to another to help them improve. Feedback is part of a social interaction influenced by culture, values, expectations, personal histories, relationships, and power. Do treat feedback as a conversation rather than as a commodity.**
   In a review paper on the role of feedback in self-assessment, Sargeant et al. (3) described how feedback from medical colleagues is part of a social process in which information is used to construct an understanding of one's own performance. Reconciling and assimilating negative feedback with views held by the individual was described to be influenced by social context. Watling et al. explored how different professions, i.e. music, teacher training, and medicine, deal with feedback. The differences between professions described in that study highlight the influence of social and cultural values on the role and impact of feedback (4). Viewing feedback only as 'specific information about the comparison between a trainee's observed performance and a standard, given with the intent to improve the trainee's performance' (5) ignores the complex ways in which culture, values, expectations, personal histories, relationships, and power manifest themselves through feedback (6).

2. **Do recognize that trainees must perceive feedback as credible in order for it to be influential. Credible feedback is well-informed, typically by direct observation of the task or event, and it comes from a trustworthy source. Make sure that you as supervisor set a good example as a credible role model.**
   A number of qualitative studies have shown that learners value feedback that they deem to be credible, but may dismiss feedback that they perceive to lack credibility (7–10). Feedback that is negative or corrective is especially likely to be subjected to an appraisal of its credibility before learners will accept or act upon it. The credibility of feedback is influenced by the credibility of the source, by the process by which the feedback was informed and created, and by the content and characteristics of the feedback itself (9).

3. **Decide the timing of feedback depending on the competence level of the trainee and on the complexity of the task.**

Studies of learners' perceptions of effective feedback have highlighted the importance of timeliness to learners' acceptance and use of feedback (7,11), confirming that the all-too-frequent practice, within medical training, of providing performance feedback long after the event is rarely perceived by learners as useful. Although there is general agreement that feedback should be 'timely', the concept of optimal timeliness appears to be a nuanced one. For example, for simulation training of procedural skills, terminal feedback (at the end of the task performance) may be superior to concurrent feedback (during the task performance) for enhancing learning (12).

Hattie and Timperley (13) provide evidence that different levels of feedback deserve different timing. Thus immediate error correction during task acquisition is more effective than delayed, whereas immediate correction when trying to build fluency will detract from the learning of automaticity which is a process and therefore better discussed after the event. Feedback after an audit showing comparatively poor clinical performance was most effective if given more than once and in writing as well as verbally (14).

4. **Do encourage trainees to look for feedback and use it to enhance their performance.**

Our trainees may approach feedback with trepidation about the harm it might do to their self-esteem (15); they may desire to make a good impression on their trainer among others; they may also desire the information which feedback gives them about how to improve (16). These are the complex and largely unconscious psychological influences on feedback-seeking (17). Trainees may hesitate to seek feedback on the very occasions when they might benefit from it most: situations where their performance has fallen below the required standard. In light of evidence for a heightened impact of feedback in these circumstances, the need to support trainees to seek and use feedback is especially pressing (14).

Research in non-clinical higher education shows that learners ask for feedback more frequently and see more benefits than costs in it as it is perceived to contain more valuable information. This assessment made by the learner of the potential value of feedback information is influenced by goal orientation (18,19). Teunissen et al. showed that this relationship between goal orientation and increased frequency of feedback seeking also holds in a population of postgraduate medical trainees (20). There are experimental studies showing that although goal orientation is a fairly stable concept, a learning goal orientation can be fostered (21). Supervisors should therefore encourage a learning frame of mind – this makes trainees more likely to accept formative feedback (17,21). In practical terms, this will involve welcoming discussions of the need to improve, encouraging goal-setting and planning of learning (22).

The content of feedback

5. **Do tailor bespoke feedback to the individual trainee. The trainee might benefit from:**
   - **Reinforcement of key points done well**
   - **Identification of key points which might have been done better or omissions**
   - **Working out strategies for improving the quality of their work**
   - **An increased self-awareness**

   Feedback needs to be tailored to the trainee's perceptions (2). It is most effective if directed at unsatisfactory elements of performance and linked to specific learning aims (23). The content of feedback should therefore arise from a diagnostic and supportive dialogue between supervisor and trainee (24).

   Learners actively process some (but not all) of the information they get in feedback (25,26). Relevance and credibility are important parameters for learners to decide how to act on feedback (10). Both appear to increase when feedback is tailored to an individual's needs.

6. **Do give specific feedback, focused on how the task was done and how that type of task should/might be done.**

   That feedback should be specific seems self-evident, and advice to teachers on giving feedback almost universally endorses the provision of specific feedback. General information unrelated to the performance, comments about a good or poor performance or compliments are less effective than specific comments (27,28). Lack of specificity has repeatedly been identified as an all-too-common weakness of the feedback that is typically exchanged in medical training (29). When, however, one looks for evidence that increasing feedback specificity leads to more effective learning, the waters become murkier. Goodman et al. (30), for example, showed that increasing the specificity of feedback benefits initial performance, but discourages exploration, potentially undermining the deeper learning required for independent performance.

   Kluger and DeNisi's feedback intervention theory, derived from their meta-analysis of over 130 studies of feedback interventions in various settings, also posits that feedback becomes less effective as attention shifts away from the task and toward the individual; in short, feedback that is threatening to self-esteem is unlikely to be effective (28). Sargeant invoked this theory to explain the difficulty practising physicians experienced in accepting and using negative or critical multisource feedback (31).

   To sum up the advice from Hattie and Timperley (13) and Kluger and DeNisi (28), which is echoed by Archer (32) in his overview on the topic, feedback is most effective when directed at the task level and may assist in 'deep processing and mastery of tasks' when it is about processing of tasks or self-regulation. A 'Don't' is providing feedback that focuses on the person level. According to Hattie and Timperley, person-oriented feedback 'usually contains little task-related information and is rarely converted into more engagement, commitment to the learning goals, enhanced self-efficacy, or understanding about the task'(13, page 96).

7. **Do make sure to indicate whether feedback is about necessary improvement for minimally acceptable performance or whether it is a reflection on possible variations to build upon adequate performance.**
   Consider offering grades as an element of formative feedback if it seems that receiving grades will enhance the seeking of strategies for improvement. Conversely, avoid giving grades to trainees who you suspect will stop trying to learn if they get a good enough grade and to those who will give up if they get a poor grade.

   Self-regulation theories suggest that within each of our trainees are two basic self-regulation systems which co-exist but may conflict (33). These two systems - the promotion (doing things because you want to) and prevention (doing things because you have to in order to avoid harm) approaches - may both be active in response to feedback (34). It is important that the supervisor recognizes that his/her trainee is predominantly in promotion or prevention focus with respect to the focus of feedback, because positive feedback is more effective in motivating performance improvement for learners in promotion focus, while negative feedback is more useful in motivating performance improvement for learners in prevention focus (28). Linking this with the evidence about goal orientation in Guideline No. 4, the promotion system generates goals which are experienced as desire for gratification, so learning goals when achieved will excite an increased desire to learn. The prevention self-regulatory system may encourage learning for fear of failure but this will feel like a necessity and achievement will cause relaxation rather than a desire for further learning (34). The prevention system is active in individuals with performance goals – aiming to prove that one is already adequately competent and avoiding criticism. Feedback works best for learning when the trainee has learning goals rather than performance goals (17,35) so it is important that the feedback itself should not push the trainee towards performance goals.
   Grades are a clear and non-nuanced form of feedback which can trigger both promotion and prevention responses in trainees (28). If a trainee is keen to know where they are in the opinion of the supervisor, their reasons can be explored by a supportive supervisor who can encourage a learning approach, aiming for self-awareness of competency and prioritization of areas for improvement. Receiving grades in this frame of mind was found to enhance the seeking of strategies for improvement, especially if criteria for allocation of grades are understood (26,36). Harmful effects of grades have also been noted in some participants in school, higher education and medical education (13,26,34), suggesting that making grades optional in formative feedback may be wise, with trainee choice being respected but perhaps explored by supervisors.

8. **Do ensure that feedback is actionable, enabling the trainee to construct strategies for improvement. After discussing the trainee's performance of a task, provide some guidance or 'scaffolding' to enable them to step beyond their current competence.**
   Sadler suggests that for information to become feedback, it must enable the learner to take action to remedy the gap between actual and desired performance (37). Information about 'what went wrong' that fails to enable learner action 'how you can improve' is merely 'dangling data' that is unlikely to motivate learning. Research into learners' experiences of

feedback has highlighted the value placed on feedback that is actionable (38). Actionable feedback contains a roadmap for learner development; it provides explicit suggestions for building on strengths or addressing weaknesses in performance.

The theoretical concept of 'scaffolding' by tutors has been well developed by Wood et al. in their constructivist model of learning (39). They based this on Vygotsky's many studies in children of how the learner is helped to develop into their 'zone of proximal development' (beyond their current ability) by social interaction with tutors or peers (40). In the social interactions of adult learning the scaffolding concept can also be helpful (27,41–43).

The tasks of scaffolding as described by Wood et al. are:

1. Orient the learner to the task
2. Simplify into steps
3. Motivate to maintain effort to achieve the goal
4. Highlight critical features of the task
5. Control frustration and the risk of failure
6. Provide a model of the required actions

For trainees with a low level of competence, scaffolding involves giving directive feedback or specific instructions; for trainees with a high level of competence scaffolding can be less directive i.e. suggestions, hints and tips for (further) improvement (facilitative feedback) (27).

9. **Do attend to trainee motivation when discussing strategies for improvement**
   In studies of educational psychology in children, motivation was a separate facet of the scaffolding of a challenging task (39). Learning takes place at the edge of the comfort zone (40). To prevent a child from giving up their efforts the teacher needs to encourage the child to believe that mastering the task is both possible and important. In adult learners, motivation is more likely to be internally generated (44) but it is no less important to learning, and is influenced by feedback (28,34,45). In aiming for sufficient motivation to learn to do the task and sufficient self-efficacy that their effort is likely to succeed, clinical supervisors should check trainee response to their feedback as they go along. Trainee response depends on perceptions of the advice – does it challenge their way of doing things? (I need to change) and is the emotional impact of feedback positive? (I want to change and believe I can change). The trainee who will pay attention to the formative advice in feedback is the one who thinks they need to and can improve. It may be that the trainee had not identified the need for improvement before they got feedback from a credible source which alerts them to the need. The question then is whether they acknowledge that need and seem to want to improve. Clinical tutors can enhance motivation by making the suggestions in the feedback align with the trainee's goals and therefore seem relevant (27).

10. **Regardless of the specific approach to feedback that is used, do engage the trainee in a reflective conversation that marries their self-assessment with your observations and elaborations.**
    By involving trainees in a discussion, supervisors can raise their awareness of their performance relative to their goals of quality performance through reflection-in-action and

reflection-on-action (22,32). Coaching then includes confirming or challenging the trainee's self-assessment, while recognizing that a challenge to the self-assessment of a junior learner whose understanding of the task is still superficial should differ from the challenge made to a more experienced trainee. Junior learners being less familiar with quality performance will rely more on the opinions of others (supervisors, peers) to make their self-assessment, and may need to be allowed an inflated self-efficacy and to receive the challenge step by step in order to keep trying. There are many factors which influence the effect of feedback, and the choice of how to deliver the feedback will depend on the task, the recipient and the feedback relationship (23). Feedback should be 'A supported sequential process rather than a series of unrelated events' (32).

Several approaches to feedback have been described in the literature (sandwich, Pendleton, reflective feedback conversation, agenda-led outcome-based analysis, feedforward), but no single approach has been established to be the most effective. Rather, the likely best approach varies according to the learner, the teacher-learner relationship, and the context. The approaches mentioned are:

The feedback sandwich (in which the supervisor describes what went well, what can be improved, then re-emphasizes what went well) (46) harnesses the psychological effect of praise to enable the reception of criticism. This approach is thought helpful especially in the delicate start of a feedback relationship, but unnecessary once the relationship is robust. Evidence of its effectiveness is lacking.

Pendleton (47) outlined a method for giving feedback aiming to engage the learner in self-reflection and to balance positive and critical feedback. He suggested a series of 'rules':

1. Check the learner wants and is ready for feedback.
2. Let the learner give comments/background to the material that is being assessed.
3. The learner states what was done well.
4. The observer(s) state(s) what was done well.
5. The learner states what could be improved.
6. The observer(s) state(s) how it could be improved.
7. An action plan for improvement is made.

The rules are intended to promote a safe and supportive environment, to encourage and incorporate self-assessment, and to generate recommendations rather than criticisms. The rules have been criticized as clunky and formulaic, but the framework can be helpful for learning to give and receive feedback.

Cantillon and Sargeant's concept of the **'reflective feedback conversation'** (48) is grounded in empiric work on the role of reflection as a critical link between receiving and using feedback.

The reflective feedback conversation unfolds like this:

1. The teacher asks the learner to share concerns about performance.
2. The learner describes concerns and what they would have liked to have done better.

185

3. The teacher provides views and offers support, then asks the learner what might improve the situation.
4. The learner responds, then the teacher elaborates on that response, correcting if needed, and checking understanding.

This approach focuses on the essential goals of feedback, encouraging learners to reflect, and motivating subsequent performance improvement. Importantly, the conversation should be viewed as a process rather than an event; revisiting and follow-up are often required.

Agenda-led outcome-based analysis (starts with the trainee's agenda, looks at the outcomes they were aiming for, encourages self-assessment and problem-solving, provides balanced feedback and suggests alternatives). This method is described in Kurtz, Silverman and Draper's Calgary Cambridge method for teaching communication skills (49) and is a learner-centred way of identifying the most helpful focus for a feedback discussion.

By contrast, the **feedforward interview** (34) is not actually a technique for feedback. It aims to avoid creating a discrepancy between a preferred standard and the actual state of affairs (seen as a key element of feedback, but also recognized as problematic for trainees who have low self-esteem) by focusing learners on their best performances. The trainee recalls peak moments in his/her performance and is asked to reflect on what conditions in themselves and their surroundings made that possible, then considers strategies to ensure sustainable peak performance. Kluger and van Dijk recommend periodical feedforward interviews with trainees about their peak experiences, partly in order to prepare the ground for necessary feedback to be received with a 'promotion' approach.

11. **Don't assume that a single approach to feedback will be effective with all trainees or in all circumstances. As the players and the contexts change, so too does the most useful approach to feedback.**
   Don't assume
   - **You know what a trainee wants to learn**
   - **You know why a trainee is struggling**
   - **You know if or why a trainee wants feedback**
   - **You know what information a trainee takes out of a situation or feedback conversation**

Individuals vary in their orientation toward clinical and educational tasks. Responses to feedback also differ between learners, even regarding similar performance on similar tasks. Dijksterhuis showed individual variability in the acceptance and responsiveness to feedback (50). Kluger and van Dijk (34) proposed that regulatory focus theory might explain some of the observed variability in feedback responses, and Watling et al.'s naturalistic exploration of the usefulness of this theory showed it offered some insights into feedback responses in clinical learning situations (25).

Variability in the impact of feedback extends beyond the individual. Responses to feedback are also shaped by learning culture, and the norms and expectations it creates for feedback (38), And context, including the relational element of feedback, is increasingly recognized as influential; Telio (51) has recently highlighted the contextual influence of the 'educational

alliance' that develops between teacher and learner on the feedback that is exchanged. In the face of such variability, teachers must develop versatile approaches to feedback that are grounded in an understanding of the learner. The feedback exchange is perhaps at its most effective when teachers' and learners' goals are aligned (38). Alignment requires engagement and dialogue.

12. **Don't provide feedback without follow-up. Trainees are unlikely to be influenced by feedback that is not followed by an opportunity for them to demonstrate improving performance.**
Sargeant et al. (52) explored physicians' reflective processes after they received multisource feedback. Reflection was found to influence not only the assessment and assimilation of feedback, but also the processing of their emotional responses to feedback. Furthermore, facilitated reflection was found to be useful in terms of enhancing the acceptance and use of feedback. The process of reflection, however, was often an extended one, especially when the feedback was perceived as negative or was in conflict with self-perception.

13. **Don't provide feedback that is poorly informed (or is based on hearsay); doing so diminishes the value that trainees assign to feedback in general.**
Surveys have demonstrated that trainees value feedback in principle, and value the provision of feedback as a desired quality of clinical teachers (53). In reality, however, the quality of the feedback received in medical training is often reported as low, and poorly informed due to factors including limited direct observation of performance. As a consequence, trainees may begin to devalue external feedback in general, relying instead on self-assessment (10,11). It is encouraging that this need not be the case, and the quality of feedback improves after specific training of clinical faculty (29,54,55).

14. **Don't underestimate the emotional impact of feedback that is perceived as negative. Emotional distress may be a barrier to acceptance and use of feedback.**
Feedback intervention theory (28) posits that feedback which threatens self-esteem is much less likely to be effective. Sargeant provided a sobering example of this theory in action. In a study done two years after practising doctors received multisource feedback, she found that those who had received negative feedback that conflicted with their self-assessment experienced distressing and long-lasting emotions that limited their ability to accept and act upon the feedback (31). Eva showed that the interpretation and acceptance of feedback was influenced by a complex interplay of emotions, including confidence and fear, and highlighted the importance of allowing the learner to maintain their self-concept when delivering feedback (2).

15. **Don't give grades without explaining the criteria for allocation of grades and providing strategies for improvement.**
The mini-CEX and other workplace assessments are most valuable as instruments for learning, rather than as a formal assessment of competence, but all too often grades are given with the

comment boxes left blank (56). In studies of the impact of grades in formative assessment, participants who reported that low grades motivated them to find strategies to improve did however need an explanation of the grade in order for it to be useful to them (26,57). Because of the potential for grades to demotivate or to reduce effort, it has been suggested that it might be wisest to avoid giving grades except when formally assessing the learner (in infrequent 'high stakes' assessments) (58).

What we don't know

16. **What determines the credibility of feedback?**
Credibility is a fundamental determinant of the ultimate impact of feedback on a learner. How trainees make judgements about feedback's credibility, and how well those judgements serve them educationally, deserve careful study (10).

17. **How much is the right amount of content when giving feedback?**
How does the supervisor determine how many items of feedback are optimal (both strengths and weaknesses)? We do have some evidence from higher education studies which suggests that more is less, and that increasing complexity can even reduce the effect of feedback (27). Recall of feedback is partial and selective (26). According to cognitive load theory cognitive architecture leads to a working memory that is limited in its capacity when it has to deal with novel information (59). A review on the cognitive load effects of visual and verbal instructions concluded that instructions that contain redundant information (for instance verbally stating what has already become visually obvious) more often inhibit than enhance learning (60). A set of studies in various clinical training contexts could be helpful.

18. **What determines the 'open and safe interaction' in the feedback conversation?**
Many, including Pendleton (47) , have highlighted the importance of a safe and supportive climate for the exchange of feedback. But the specific constituents of a safe climate remain poorly understood, as are the ways in which individuals and organizations can promote it.

19. **What influences the trainee's response? (constructive or destructive outcomes)**
Regulatory focus theory may explain some of the individual variability in feedback responses (25,34). What we don't know is how regulatory focus interacts with other influences on feedback's impact, such as credibility. We also don't know how regulatory focus can best be primed in order to enhance the impact of feedback.
How do the issues of vulnerability (self-efficacy), motivation to improve or to prevent harm, and credibility interact to give shape to constructive or destructive feedback in a workplace learning situation? How do we help trainees to believe that they can improve?

Responses to feedback are driven by individual traits and preferences *and* by values embedded within the learning culture. How these influences interact is inadequately understood, making it challenging to know where to focus our energies. Workplace learning

theorists (e.g. Eraut (61), Billett (62)) have highlighted the need to understand how individual and the sociocultural influences on learning interact. Billett emphasizes the notions of affordances and agency; a learning environment offers a range of affordances, or opportunities to learn, but an individual learner must exercise agency to engage with those affordances. Feedback challenges may lie with either affordance (is good feedback made available to learners?) or agency (do learners choose to engage with feedback?), or both; the way these factors interact merits further study, as it has implications for where, and how, educators and institutions should channel their energies to improve feedback.

20. **Is overt comparison with peers when made by the supervisor helpful to the trainee? Indeed, is overt comparison with required performance standards helpful?**

The evidence is rather conflicting on these two related questions, so although there is a lot of evidence we have decided that it may depend on the context and on what comparison is made.

Comparison with a standard of performance is part of one accepted definition of feedback in clinical education - 'specific information about the comparison between a trainee's performance and a standard, given with the intent to improve the trainee's performance' (5) - but while this comparison must be going on in the mind of the feedback giver, it may or may not be helpful to the trainee receiving the feedback to be aware of their position relative to the standard.

According to Kluger and de Nisi's meta-analysis, some feedback recipients feel content to be 'good enough' or become helpless when told they are not making the grade, to the detriment of their performance (28). In the studies described, feedback is more likely to have a positive than a negative effect, but what we cannot be sure of as feedback providers is which of these is more likely in a given feedback situation, although there are predictive factors (23). In a competency-based programme such as medical training it seems logical to reference the feedback given to required standards of competence. Trainees are anxious to know whether they are 'making the grade'. Enabling support of learners to self-monitor in relation to competency requirements is an important goal (63,64) and may be seen as such by our trainees which might explain why they desire and value grades. But do comparisons with standards help them to improve, or is it better for each trainee to strive for personal excellence? How can we determine which learners in which circumstances will find comparisons motivating, as opposed to disheartening?

What about comparison with peers? There is evidence that feedback becomes less effective as its focus moves away from the task and toward the self (13,28). Both self-referenced and other-referenced feedback (in)directly focus the attention to the self. Unfavourable comparisons with others may threaten self-esteem and promote a performance goal orientation, potentially hindering learning (17). But despite these concerns, some research has suggested value in comparisons: one group showed that undergraduate medical students can be motivated by and can learn from self-comparison with peers (65).

This leads us to question does the feedback sign affect trainee clinicians learning of clinical skills in the same ways as it does psychology students' performance writing essays (57), or is this effect context-dependent?

Eventually it boils down to the way the trainee's psychology is affected. Do they feel they need to change? Want to change? Know how to change? The way these desires and understandings are shaped is an area of study which is still producing conflicting results so deserves further careful study.

21. **Does a written summary of the feedback discussion enhance learning?**
    Medical students have been found to value informal verbal feedback more than formal workplace-based assessment (WBA) with written feedback (66,67). One explanation is that feedback works best soon after the event, especially for a complex task such as consulting with a patient (11,13). The value of the written summary is therefore secondary but could include:
    - Aiding reflection on the feedback at a later date
    - Aiding discussion between tutor and trainee at a later date
    - Enhancing tutor effort at the time of generating the feedback

The optimal role for written feedback represents an area for study.

Background evidence to guidelines for the learning culture (what elements of learning culture support the exchange of meaningful feedback, and what elements constrain it?)

22. **Do have a systems approach, building feedback into the learning processes**
    Institutions can create opportunities for longitudinal teacher-learner relationships to flourish, such as extended placements (68–70). Supervision of a trainee can have built-in and protected routines of supervisor observation of trainee performance followed by feedback (32) and expectations of recurrent feedback following multiple assessment tasks over time (22). Institutional expectations of supervision can include that written feedback is more than ticking boxes and ensure that the feedback instruments used enable specific explanations of the trainee's position relative to required goals, and encourage the supervisor to suggest how to attain the goals (13,71). Expectations of the trainee might be reflection-on-feedback with some system of reinforcing implementation of strategies for improvement (32). New trainees will require induction into the rules of the particular academic community.
    In order to ensure a climate of feedback, an institution should provide a system of regular feedback not only for trainees but also for supervisors (32,72).
    In addition to providing faculty development courses, educational support can be offered to supervisors and the supervisors' social networks can be used and supported to facilitate acceptance and use of feedback (73).
    In the new movement towards programmatic assessment, progress and learning from feedback is emphasized and built into the system (74,75). This has been successful (76) although it has also met some difficulties in implementation (69), and when summative

judgements are seen to be based on the formative assessments the feedback given may be less critical (77).

Some of these elements of a systems approach are further developed in the following guidelines.

23. **Do support the development of longitudinal, trusting supervisor-trainee relationships in medical training; influential feedback thrives in the context of trusting relationships.**

When trainees can build a relationship with their supervisors, it allows them to trust the credibility of the feedback they receive and the alignment of the teacher's goals with their own. As Bok et al. showed, durable teacher-learner relationships also prompt learners to seek feedback more readily (69).

Bates et al. (67) explored medical students' perceptions of assessment and feedback in a longitudinal integrated clerkship – a setting that enables the development of durable, trusting, teacher-learner relationships. They found that such relationships afforded 'constructive interpretation of critical feedback' (p.366); students were able to interpret even challenging or corrective feedback as supportive.

Within a trusting and supportive relationship, feedback is also more likely to be viewed as credible (50,78). Recognizing the centrality of relationship in the feedback process, the concept of the 'educational alliance' has been proposed as a framework for understanding the links between the teacher-learner relationship and the impact of the feedback generated within it (51,79).

24. **Do use video review with feedback as a component of training.**

The main advantage of video is that the trainee can review what they did and as well as getting feedback. The supervisor's feedback may not differ whether following direct observation or following video observation but the trainee will be able to confirm the strengths and weaknesses in their own performance.

Supervisors differ considerably in the feedback they give after reviewing the same videotaped consultation (80). This raises the question of whether the supervisor's feedback adds value to the trainee self-assessing their own videoed consultation. In a systematic review Hammoud et al. concluded that video review with self-assessment alone was not found to be generally effective for medical students, but when linked with expert feedback it was superior to traditional feedback alone (81). This is a strong argument in favour of building video review with feedback into educational programmes especially to address the important but less self-evident problems.

Potential disadvantages include the relative complexity of arranging filming and viewing and that if videos are being selectively proffered for feedback the trainee may choose their best performances.

25. **Do promote communities of practice in clinical workplaces in which feedback is routine, regular and valued.**
This can be a helpful approach in turning the workplace into a powerful learning environment when it can otherwise be a frustratingly hard place to change (6). If the people working together in a workplace realize that everyone is also a learner and that feedback is a powerful way of learning, an environment is created in which providing feedback is considered 'normal'. This would mean, for example, that trainees are encouraged to give feedback to their supervisors (82,83). It has also been found in the training of athletes and musicians that critical feedback is exchanged more readily when it is normalized by a learning culture (38,84). These studies provide at least indirect support for the idea that when feedback becomes a routine part of a learning culture, it may be more readily taken up and used by learners. And part of becoming 'routine' is that feedback, including critique, is exchanged very frequently.

26. **Make sure that those who have a formal role in a workplace's educational system are aware of that role and understand what learners' educational objectives should be.**
In a study of residents' expectations of their clinical teachers, Boor et al. found that, next to the importance of a good relationship, learners value clinical supervisors who are aware of the educational system and expectations and who can apply that knowledge to the individual learner (85). Van der Vleuten's comments on programmatic assessment are useful here: 'If a programme of assessment is to provide meaningful outcomes, all the players should understand what they are doing, why they are doing it, and why they are doing it *this way.'* (86) If we substitute 'feedback' for 'assessment', the comment rings equally true.

27. **Make sure that the team give feedback regularly, reflect on the practice of giving feedback, and follow refresher courses to maintain and improve competency in providing feedback.**
Lack of faculty insight in the assessment process remains an issue (87). The feedback landscape described by Evans (1) indicates the need for tutor training: the tutor must accurately diagnose academic and social needs; understand and empathize with the learner's perspective, and have skills to employ appropriate scaffolding tools. Although no one technique of giving feedback has proven superiority and different individual trainees may respond to different approaches, there is evidence that it has been helpful to train supervisors in techniques of providing feedback constructively, and their behaviour changes in providing more useful feedback (73).

28. **Don't rely exclusively on faculty development to improve the effectiveness of feedback.**
Historically, faculty development in feedback delivery has been the primary approach to improving the quality and effectiveness of feedback (29,73). This focus on how feedback is given ignores the important element of how it is received by learners (88). The crucial role of learning culture in making effective feedback possible, normalizing constructive criticism, and establishing the value of feedback for learning is also missed by an approach focused on individual teachers (38). Faculty development is important but not sufficient; attention must also be paid to learners' receptivity to feedback and to the elements of the learning culture that support or constrain the feedback exchange.

29. **Don't allow formal assessments of clinical skills, such as the mini-CEX, to be completed without observation and feedback**

   Although designed to rely on observation of at least one clinical encounter and including space for documentation of feedback discussions, paper instruments such as the mini-CEX are frequently used as tick-box exercises to enable progression of trainees (56)

What we don't know about the feedback culture

30. **What are the vital components that ensure a constructive system of workplace learning that caters to trainees, workers, and the educational system? How can the institution nourish a climate which encourages the provision and seeking of feedback?**

   Although we found several Do's relating to the system approach to learning in the workplace, these feedback approaches are largely limited to individuals, despite ways of working in health care that increasingly demand competent team function. There are few studies on the impact of the provision of feedback to teams of individuals and the outcomes are variable, as described in a review by Gabelica et al. (89). They raised an interesting paradox: 'On the one hand… feedback might impact a huge diversity of critical team processes (amongst which the three most frequent variables: motivation, team goal, and collaboration/cooperation) and emergent states (among which the most frequent variables: collective efficacy, cohesion, outcome expectations, and task concern/interest) and occasionally have a direct effect on team performance (in 23 studies overall). On the other hand, some studies confirmed that feedback might not always lead to significant or at least measurable changes and thus not fulfil its function as a leverage point that can be used to support teams.' They conclude that the real question is not whether feedback works, but under what circumstances is works best. A model is provided that highlights key factors that might enhance and support feedback effectiveness. Feedback about and during the <u>process</u> of teamwork was more reliably effective than feedback about <u>performance</u> given to the team or to individuals within the team. They recommended further research into what makes for effective feedback about team processes - how teams communicate, interact, establish their team atmosphere, define team objectives and strategies, monitor performance, come to a common understanding of the task and its requirements, build on each other's expertise, make team decisions and coordinate in an efficient way.

31. **Is it most effective to give feedback to individuals alone or in a group setting?**

   In group learning of clinical skills, feedback to the trainee(s) who have experimented with a task is generally given by and in front of the group. This can include group feedback on a videoed real patient consultation. The advantages to this approach are that a range of feedback perspectives are gained, feedback-giving is role-modelled, and observers learn vicariously. The disadvantages are reduced control over content and volume of feedback, plus the risk of a negative emotional impact. In situations where it might be possible to give feedback either individually or in a group setting, we do not know whether the advantages

outweigh the disadvantages. There are studies which have found learner preferences for group feedback (90) and for individual feedback (1) and it is clear that the context matters (28).

32. **Does the use of formative assessment outcomes for summative purposes (such as having supervisors provide formative feedback that at the end of a rotation is also used for a summative assessment) corrupt a well-intentioned educational system?**

    Programmatic assessment (a system of frequent formative assessments also used for end-of-year summative judgements) is designed to optimize learning and reduce exam stress (86). Evidence is now emerging from qualitative evaluations of programmatic assessment curricula which raises questions about the mixing of formative and summative assessment. A qualitative study with clinical undergraduate veterinary students and their supervisors highlighted that both struggled with formative assessments that are used as 'data points' for a final summative judgment. As a result, the formative assessments did not play the powerful assessment-for-learning role they are meant to have in a curriculum based on programmatic assessment (69).
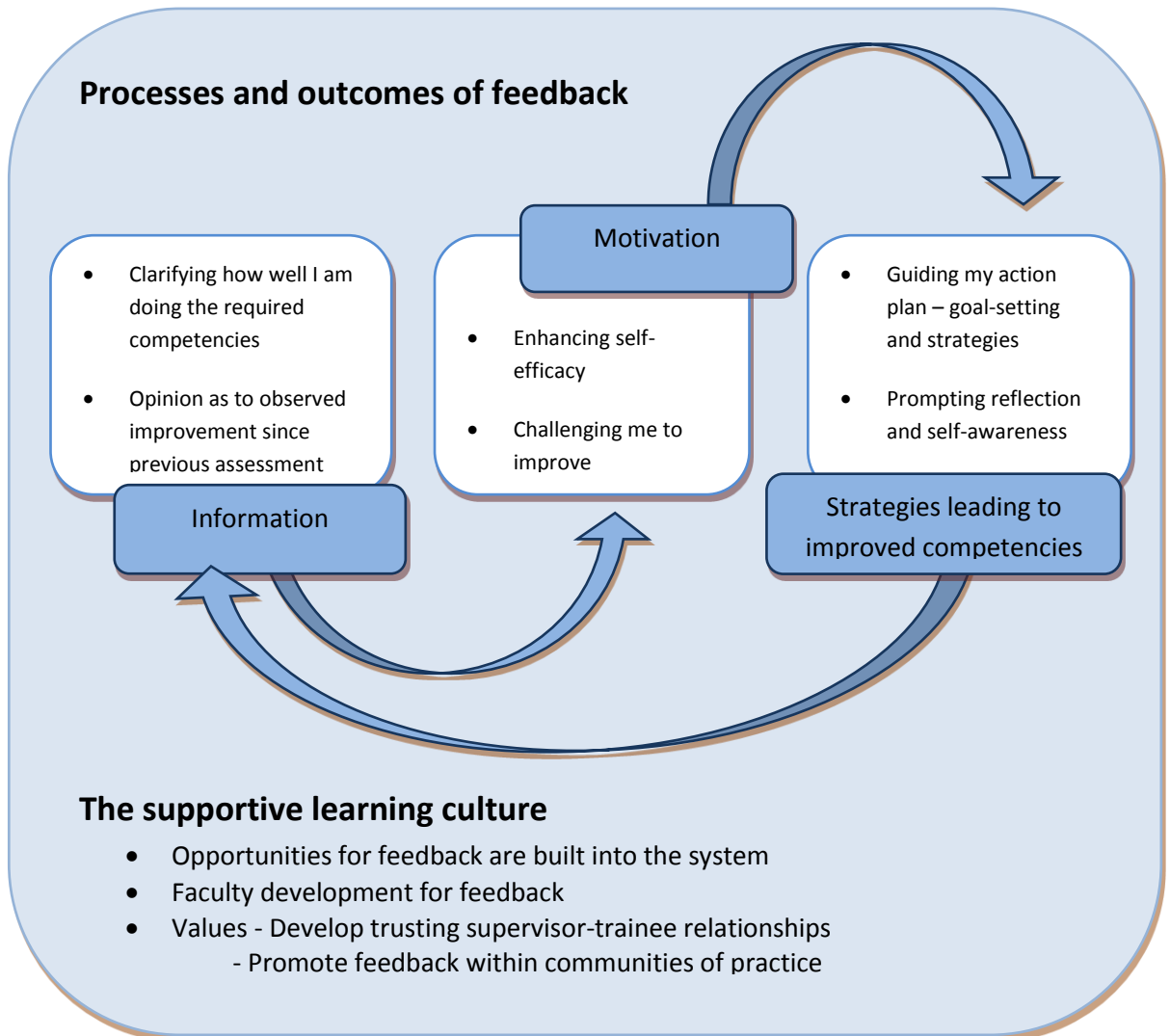
    Medical education not only blurs the line, at times, between summative and formative assessment, but also blurs the line for its teachers between the roles of coach and assessor. Although these roles are distinct - coaches provide formative feedback while assessors make summative judgements – the same teacher is routinely expected to play both roles simultaneously and for the same learner. Recent literature has begun to challenge this approach, suggesting that the quality and impact of feedback may be compromised when the teacher is assigned this dual role (38,91,92). Exactly how feedback is impacted by this practice, and whether feedback would be more effective if the coaching and assessment roles were separated, remains unknown.

Summary

We have produced what we hope is a usable set of guidelines in an area that is central to teaching. Our work adds to the literature by interpreting a diverse and sometimes contradictory range of research and opinion for the clinical supervisor and his/her manager.

We have also developed a visual representation of the feedback process and outcomes (Fig. 1) which we offer as a summary of the guidelines from the viewpoint of the recipient of feedback. Trainees are looking for **information** about their performance and **motivation** to be/aim to be exemplary clinicians. The feedback process is incomplete if it does not result in the generation of **strategies for improvement** – either recommendations, or self-generated as a result of feedback. And the best feedback process loops back into a subsequent assessment with feedback about whether this has resulted in improved clinical performance. These processes and outcomes will flourish in the **supportive learning culture** of systematic dialogic feedback.

Figure 1: Feedback processes and outcomes – what the trainee wants from the feedback relationship.

Processes and outcomes of feedback

- Clarifying how well I am doing the required competencies
- Opinion as to observed improvement since previous assessment

Information

Motivation

- Enhancing self-efficacy
- Challenging me to improve

- Guiding my action plan – goal-setting and strategies
- Prompting reflection and self-awareness

Strategies leading to improved competencies

The supportive learning culture
- Opportunities for feedback are built into the system
- Faculty development for feedback
- Values - Develop trusting supervisor-trainee relationships
     - Promote feedback within communities of practice

Strengths and weaknesses of the paper

Our combination of perspectives and our iterative, consensus-building approach to creating these guidelines are strengths of this work. An obvious weakness is the lack of a systematic search method; as a consequence, we will have missed some useable evidence. We did, however, use the systematic reviews we know of to ensure that the evidence therein has contributed to these guidelines.

Conclusion

Feedback resists one-size-fits-all guidelines. The wealth of research on feedback paints a picture of a nuanced process, with a great potential to help learners in all sorts of circumstances, but also

a process that is fraught with variability and unpredictability, and influenced by individuals, contexts, and culture. In short, feedback is both an opportunity and a threat for teachers and learners. But we must not simply throw up our hands. Feedback may be complex, but it is essential to learning in medicine. We encourage supervisors to support best practices in feedback by embracing the Do's we have identified and banishing the Don'ts. And we invite researchers to explore the intriguing and critical Don't Knows of feedback, so that the field continues to advance and the next set of guidelines will be even more firmly grounded in empirical work. Our work has challenged us to reconsider the very definition of feedback in medical education. We offer a new definition that may help to shape future conversations*:*

Helpful feedback is a supportive conversation that clarifies the trainee's awareness of their developing competencies, enhances their self-efficacy for making progress, challenges them to set objectives for improvement, and facilitates their development of strategies to enable that improvement to occur.

References

1. Evans C. Making Sense of Assessment Feedback in Higher Education. Rev Educ Res. 2013 Feb 22;83(1):70–120.

2. Eva KW, Armson H, Holmboe E, Lockyer J, Loney E, Mann K, et al. Factors influencing responsiveness to feedback: on the interplay between fear, confidence, and reasoning processes. Adv Health Sci Educ Theory Pract. 2012 Mar;17(1):15–26.

3. Sargeant J, Mann K, van der Vleuten C, Metsemakers J. "Directed" self-assessment: practice and feedback within a social context. J Contin Educ Health Prof. 2008;28(1):47–54.

4. Watling C, Driessen E, Van der Vleuten CPM, Vanstone M, Lingard L. Beyond individualism: Professional culture and its influence on feedback. Med Educ. 2013;47:585–94.

5. Van de Ridder JMM, Stokking KM, McGaghie WC, ten Cate OTJ. What is feedback in clinical education? Med Educ. 2008 Feb;42(2):189–97.

6. Teunissen PW. Experience, trajectories, and reifications: an emerging framework of practice-based learning in healthcare workplaces. Adv Heal Sci Educ. Springer Netherlands; 2015;20:843–56.

7. Bing-You RG, Paterson J. Feedback falling on deaf ears: Residents' receptivity to feedback tempered by sender credibility. Med Teach. 1997;19(1):40.

8.   Sargeant J, Mann K, Ferrier S. Exploring family physicians' reactions to multisource feedback: perceptions of credibility and usefulness. Med Educ. 2005 May;39(5):497–504.

9.   Sargeant J, Mann K, Sinclair D, Van Der Vleuten C, Metsemakers J. Challenges in multisource feedback: Intended and unintended outcomes. Med Educ. 2007;41:583–91.

10.   Watling C, Driessen E, van der Vleuten CPM, Lingard L. Learning from clinical work: the roles of learning cues and credibility judgements. Med Educ. 2012 Feb;46(2):192–200.

11.   Watling C, Kenyon C, Zibrowski E, Schulz V, Goldszmidt M, Singh I, et al. Rules of engagement: residents' perceptions of the in-training evaluation process. Acad Med. 2008;83(10 Suppl):S97–100.

12.   Walsh C, Ling S, Carnahan H. Concurrent versus terminal feedback: it may be better to wait. Acad Med. 2009;84(10 Suppl):S54–7.

13.   Hattie J, Timperley H. The Power of Feedback. Rev Educ Res. 2007 Mar 1;77(1):81–112.

14.   Ivers N, Jamtvedt G, Flottorp S, Jm Y, Sd F, Ma OB, et al. Audit and feedback : effects on professional practice and healthcare outcomes. Cochrane Database of Systematic Reviews. John Wiley & Sons, Ltd; 2012.

15.   Higgins E. Promotion and prevention: regulatory focus as a motivational principle. In: Zanna MP, editor. Advances in Experimental Social Psychology. San Diego, CA: Academic Press; 1998. p. 1–46.

16.   Tuckey M, Brewer N, Williamson P. The influence of motives and goal orientation on feedback seeking. J Occup Organ Psychol. 2002;75(2):195–216.

17.   Teunissen PW, Bok HGJ. Believing is seeing: How people's beliefs influence goals, emotions and behaviour. Med Educ. 2013;47(11):1064–72.

18.   Morrison E, Cummings L. The impact of feedback diagnosticity and performance expectations on feedback-seeking behaviour. Hum Perform. 1992;5(4):251–64.

19.   Tsui AS, Ashford SJ. Adaptive Self-regulation: A Process View of Managerial Effectiveness. J Manage. 1994;20(1):93–121.

20.   Teunissen PW, Stapel DA, van der Vleuten CPM, Scherpbier AJJA, Boor K, Scheele F. Who Wants Feedback? An Investigation of the Variables Influencing Residents' Feedback-Seeking Behavior in Relation to Night Shifts. Acad Med. 2009;84:910.

21. Molloy E, Boud D. Seeking a different angle on feedback in clinical education: the learner as seeker, judge and user of performance information. Med Educ. 2013 Mar;47(3):227–9.

22. Boud D, Molloy E. Rethinking models of feedback for learning: the challenge of design. Assess Eval High Educ. 2013 Sep;38(6):698–712.

23. Van de Ridder JMM, McGaghie WC, Stokking KM, ten Cate OTJ. Variables that affect the process and outcome of feedback, relevant for medical training: a meta-review. Med Educ. 2015;49(7):658–73.

24. Hewson M, Little M. Giving feedback in medical education. J Gen Intern Med. 1998;13(2):111.

25. Watling C, Driessen E, van der Vleuten CPM, Vanstone M, Lingard L. Understanding responses to feedback: the potential and limitations of regulatory focus theory. Med Educ. 2012 Jun;46(6):593–603.

26. Lefroy J, Hawarden A, Gay SP, McKinley RK, Cleland J. Grades in formative workplace-based assessment: a study of what works for whom and why. Med Educ. 2015;49:307–20.

27. Shute VJ. Focus on Formative Feedback. Rev Educ Res. 2008 Mar 1;78(1):153–89.

28. Kluger A, DeNisi A. The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. . Psychol Bull. 1996;119(2):254.

29. Weston PSJ, Smith CA. The use of mini-CEX in UK foundation training six years following its introduction: Lessons still to be learned and the benefit of formal teaching regarding its utility. Med Teach. 2014;36(2):155–63.

30. Goodman JS, Wood RE, Hendrickx M. Feedback Specificity, Exploration, and Learning. J Appl Psychol. 2004;89(2):248–62.

31. Sargeant J, Mann K, Sinclair D, van der Vleuten C, Metsemakers J. Understanding the influence of emotions and reflection upon multi-source feedback acceptance and use. Adv Heal Sci Educ Theory Pr. 2008;13(3):275–88.

32. Archer JC. State of the science in health professional education: effective feedback. Med Educ. 2010 Jan;44(1):101–8.

33. Higgins E. Making a good decision: Value from fit. Am Psychol. 2000;55(11):1215.

34. Kluger A, Van Dijk A. Feedback, the various tasks of the doctor, and the feedforward alternative. Med Educ. 2010;44:1166.

35. Bok HGJ, Teunissen PW, Spruijt A, Fokkema JPI, van Beukelen P, Jaarsma D a DC, et al. Clarifying students' feedback-seeking behaviour in clinical clerkships. Med Educ. 2013 Mar;47(3):282–91.

36. Higgins R, Hartley P, Skelton A. The Conscientious Consumer : reconsidering the role of assessment feedback in student learning. Stud High Educ. 2002;27(1):53–64.

37. Sadler DR. Beyond feedback: developing student capability in complex appraisal. Assess Eval High Educ. 2010 Aug;35(5):535–50.

38. Watling C, Driessen E, van der Vleuten CPM, Lingard L. Learning culture and feedback: An international study of medical athletes and musicians. Med Educ. 2014;48:713–23.

39. Wood D, Bruner J, Ross G. The role of tutoring in problem-solving. J Child Psychol Psychiatry. 1976;17:89–100.

40. Vygotsky L. Thought and Language. MIT; 1986.

41. Smith SE, Tallentire VR, Cameron HS, Wood SM. The effects of contributing to patient care on medical students' workplace learning. Med Educ. 2013 Dec;47(12):1184–96.

42. Carless D, Salter D, Yang M, Lam J. Developing sustainable feedback practices. Stud High Educ. 2011 Jun;36(4):395–407.

43. Lingard L, Garwood K, Schryer CF, Spafford MM. A certain art of uncertainty: case presentation and the development of professional identity. Soc Sci Med. 2003 Feb;56(3):603–16.

44. Sitzmann T, Ely K. A Meta-Analysis of Self-Regulated Learning in Work-Related Training and Educational Attainment: What We Know and Where We Need to Go. Psychol Bull. 2011;137(3):421.

45. Overeem K. "Paying it forward": performance improvement through feedforward interviews. Med Educ. 2010 Dec 17;44(12):1159–61.

46. Dohrenwend A. Serving up the feedback sandwich. Fam Pract Manag. 2002;9(10):43–6.

47. Pendleton D, Schofield T, Tate P, Havelock P. The new consultation Developing doctor-patient communication. Oxford University Press; 2003.

48.   Cantillon P, Sargeant J. Giving feedback in clinical settings. BMJ. 2008;337:1292–4.

49.   Kurtz S, Silverman J, Draper J. Teaching and Learning Communication Skills in Medicine Second Edition. Oxford: Radcliffe Publishing Ltd; 2005.

50.   Dijksterhuis MGK, Schuwirth LWT, Braat DDM, Teunissen PW, Scheele F. A qualitative study on trainees' and supervisors' perceptions of assessment for learning in postgraduate medical education. Med Teach. 2013;35(8):e1396–402.

51.   Telio S, Ajjawi R, Regehr G. The "Educational Alliance" as a Framework for Reconceptualizing Feedback in Medical Education. Acad Med. 2015;90(5):609–14.

52.   Sargeant JM, Mann K V., van der Vleuten CP, Metsemakers JF. Reflection: a link between receiving and using assessment feedback. Adv Heal Sci Educ. 2008;14(3):399–410.

53.   Schultz KW, Kirby J, Delva D, Godwin M, Verma S, Birtwhistle R, et al. Medical Students' and Residents' preferred site characteristics and preceptor behaviours for learning in the ambulatory setting: a cross-sectional survey. BMC Med Educ. 2004;4:12.

54.   Minehart R, Rudolph J, Pian-Smith M, Raemar D. Improving Faculty Feedback to Resident Trainees during a Simulated Case. Anesthesiology. 2014;120(1):160–71.

55.   Sox CM, Dell M, Phillipi CA, Cabral HJ, Vargas G, Lewin LO. Feedback on Oral Presentations During Pediatric Clerkships: A Randomized Controlled Trial. Pediatrics. 2014;134(5):965–71.

56.   Norcini J, Burch V. Workplace-based assessment as an educational tool: AMEE Guide No. 31. Med Teach. 2007;29(9-10):855.

57.   Lipnevich AA, Smith JK. Effects of differential feedback on students' examination performance. J Exp Psychol Appl. 2009 Dec;15(4):319–33.

58.   Driessen E, Scheele F. What is wrong with assessment in postgraduate training? Lessons from clinical practice and educational research. Med Teach. 2013;35(7):569–74.

59.   Van Merriënboer JJG, Sweller J. Cognitive load theory in health professional education: Design principles and strategies. Med Educ. 2010;44(1):85–93.

60.   Kalyuga S. Instructional benefits of spoken words: A review of cognitive load factors. Educ Res Rev. Elsevier Ltd; 2012;7(2):145–59.

61.   Eraut M. Learning from other people in the workplace. Oxford Rev Educ. 2007;33(February 2015):403–22.

62.   Billett S. Toward a Workplace Pedagogy: Guidance, Participation, and Engagement. Adult Educ Q. 2002;53(1):27–43.

63.   GMC. Tomorrow ' s Doctors The duties of a doctor registered with the General Medical Council. 2009.

64.   Eva K, Regehr G. Self-assessment in the health professions: a reformulation and research agenda. Acad Med. 2005;80:46.

65.   Raat J, Kuks J, Cohen-Schotanus J. Learning in clinical practice: Stimulating and discouraging response to social comparison. Med Teach. 2010;32(11):899–904.

66.   Urquhart LM, Rees CE, Ker JS. Making sense of feedback experiences: a multi-school study of medical students' narratives. Med Educ. 2014 Feb;48(2):189–203.

67.   Bates J, Konkin J, Suddards C, Dobson S, Pratt D. Student perceptions of assessment and feedback in longitudinal integrated clerkships. Med Educ. 2013;47:362–74.

68.   Veloski J, Boex JR, Grasberger J, Evans A, Wolfson DB. Systematic review of the literature on assessment, feedback and physicians' clinical performance: BEME Guide No. 7. Med Teach. 2006;28(2):117.

69.   Bok HG, Teunissen PW, Favier RP. Programmatic assessment of competency-based workplace learning: when theory meets practice. BMC Med Educ. BMC Medical Education; 2013;13(1):123.

70.   Hägg-Martinell A, Hult H, Henriksson P, Kiessling A. Students perceive healthcare as a valuable learning environment when accepted as a part of the workplace community. Educ Heal Chang Learn Pract. 2014;27(1):15–23.

71.   Lefroy J, Thomas A, Harrison C, Williams S, O'Mahony F, Gay S, et al. Development and face validation of strategies for improving consultation skills. Adv Heal Sci Educ. 2014;19(5):661–85.

72.   Jippes E, Steinert Y, Pols J, Achterkamp M, van Engelen J, Brand P. How do social networks and faculty development courses affect clinical supervisors' adoption of a medical education innovation? An exploratory study. Acad Med. 2013;88(3):398–404.

73.   Jippes E, Van Luijk SJ, Pols J, Achterkamp MC, Brand PLP, Van Engelen JML. Facilitators and barriers to a nationwide implementation of competency-based postgraduate medical curricula: A qualitative study. Med Teach. 2012;34(8):e589–602.

74. Van der Vleuten CPM, Schuwirth LWT. Assessing professional competence: from methods to programmes. Med Educ. 2005 Mar;39(3):309–17.

75. Schuwirth LWT, Van der Vleuten CPM. Programmatic assessment: From assessment of learning to assessment for learning. Med Teach. 2011 Jan;33(6):478–85.

76. Driessen EW, van Tartwijk J, Govaerts M, Teunissen P, van der Vleuten CPM. The use of programmatic assessment in the clinical workplace: a Maastricht case report. Med Teach. 2012 Jan;34(3):226–31.

77. Bok HGJ, Jaarsma D a. DC, Spruijt A, Van Beukelen P, Van Der Vleuten CPM, Teunissen PW. Feedback-giving behaviour in performance evaluations during clinical clerkships. Med Teach. Informa UK Ltd; 2015;00(00):1–8.

78. Nasca TJ, Weiss KB, Bagian JP. Improving Clinical Learning Environments for for Tomorrow's Physicians. N Engl J Med. 2014;370(11):2014–6.

79. Wearne S, Dornan T, Teunissen PW, Skinner T. General practitioners as supervisors in postgraduate clinical education: An integrative review. Med Educ. 2012;46(12):1161–73.

80. Alves de Lima A, Conde D, Costabel J, Corso J, van der Vleuten C. A laboratory study on the reliability estimations of the mini-CEX. Adv Heal Sci Educ. 2013;18(1):5–13.

81. Hammoud M, Morgan, Edwards, Lyon, White. Is video review of patient encounters an effective tool for medical student learning? A review of the literature. Adv Med Educ Pract. 2012;3:19–30.

82. Fluit C V, Bolhuis S, Klaassen T, DE Visser M, Grol R, Laan R, et al. Residents provide feedback to their clinical teachers: reflection through dialogue. Med Teach. 2013;35(9):e1485–92.

83. Fokkema J, Scheele F, Westerman M, van Exel J, Scherpbier A, van der Vleuten C, et al. Research Reports Perceived Effects of Innovations in Postgraduate Medical Education: A Q Study Focusing on Workplace-Based Assessment. Acad Med. 2014;89(9):1259–66.

84. Watling C, Driessen E, van der Vleuten CPM, Vanstone M, Lingard L. Music lessons: Revealing medicine's learning culture through a comparison with that of music. Med Educ. 2013;47(8):842–50.

85. Boor K, Teunissen PW, Scherpbier AJJA, van der Vleuten CPM, van de Lande J, Scheele F. Residents' perceptions of the ideal clinical teacher-A qualitative study. Eur J Obstet Gynecol Reprod Biol. 2008;140(2):152–7.

86.   Van der Vleuten CPM, Schuwirth LWT, Driessen EW, Dijkstra J, Tigelaar D, Baartman LKJ, et al. A model for programmatic assessment fit for purpose. Med Teach. 2012 Jan;34(3):205–14.

87.   Massie J, Ali JM. Workplace-based assessment: a review of user perceptions and strategies to address the identified shortcomings. Adv Heal Sci Educ. 2015;(online first DOI 10.1007/s10459-015-9614-0).

88.   Eva KW, Regehr G. Exploring the divergence between self-assessment and self-monitoring. Adv Health Sci Educ Theory Pract. 2011 Aug;16(3):311–29.

89.   Gabelica C, Bossche P Van Den, Segers M, Gijselaers W. Feedback, a powerful lever in teams: A review. Educ Res Rev. Elsevier Ltd; 2012;7(2):123–44.

90.   Rowe AD, Wood LN. What feedback do students want ? [Internet]. Australian Educational Researcher. 2007 [cited 2015 Aug 6]. Available from: http://www.aare.edu.au/publications-database.php/5501/what-feedback-do-students-want

91.   Cavalcanti RB, Detsky AS. The Education and Training of Future PhysiciansWhy Coaches Can't Be Judges. JAMA. 2011;306(9):993–4.

92.   Kogan JR, Conforti LN, Bernabeo EC, Durning SJ, Hauer KE, Holmboe ES. Faculty staff perceptions of feedback to residents after direct observation of clinical skills. Med Educ. 2012;46(2):201–15.

Chapter 11

Discussion

11.1 Critical discussion of methodologies

As our research strategy we adopted a blend of action research and realist stances. The justification for this is outlined in this section, and is followed by a critical discussion of the methodologies and methods used.

Action research is a methodology aiming to solve practical problems and to contribute to knowledge in a participatory process (1). Action research methodology has been used to effect societal change and institutional development, with a variety of underlying critical, interpretivist or positivist assumptions determining choice of methods of enquiry (2,3). Kember (1) points out that while recognised schools of action research differ in the degree to which they are influenced by critical theory (that critical reflection upon practice is how changes in attitude can occur), they are similar in being:

> concerned with social practice
> aimed towards improvement
> a cyclical process
> pursued by systematic enquiry
> a reflective process
> participative (although some accept action research as individual problem-solving and reflection
> and determined by the practitioners.

McKearnan in his handbook of methods and resources for the reflective practitioner wishing to conduct curriculum action research (4) describes typologies and models of action research. Type 1 is described as scientific action research following Lewin's model : plan, fact-find, execute and analyse (5). Type 2 is practical-deliberative action research involving less monitoring by measurement and more immediate problem-solving by negotiation, aiming for a 'moral' improvement after Schon (6). Type 3 is critical-emancipatory represented by Carr and Kemmis quoted in Chapter 1 (7) following the Deakin model which is politically empowering for participants who are seen as social reformers of education. It stresses equipping practitioners with analytical skills and follows the same plan-act-observe-reflect cycle advocated by Lewin. The form of action research most current recently is that employed by practitioners to answer the question 'How do I improve my professional practice?'. This is done in collaboration with others who act as critical friends. It aims to help other people and therefore has values and virtue at its core as outlined by Jean McNiff (8) in Fig 1.

Our approach to action research was initially most akin to type 1, what Carr and Kemmis would call 'technical action research' (7) being quality improvement by problem-solving or innovation without any wider emancipatory character. Later as we widened participation, published our findings and affected the beliefs of a broad circle of clinical tutors, our action research could be described as more type 3 critical-emancipatory. As Kember comments on critical theory, action research doesn't guarantee a change in beliefs but it does provide a mechanism for participants to change their perspectives, especially through regular meetings (1).

Fig 1. Action research principles:
Action research for professional development - Concise advice for new action researchers.
Jean McNiff 2012

- You decide to investigate what you are doing with a view to improving it.
- This will help you to understand the situation more fully.
- Your developed understanding will help you to evaluate your work and change it as necessary.
- Your way of working might influence others; how can show this?
- You do this by checking your perceptions of what is happening against theirs.
- You change your way of working in light of their perceptions. You negotiate this with them.
- Your collective agreement about these things helps you all to understand the situation better.
- You learn from colleagues, and they learn from you. They decide to try things out for themselves.
- They invite you to become their critical friend, and help them evaluate their work.
- Collectively you are now a community of enquirers. You have changed your social situation, and this is bound to have consequences for wider social contexts.
- Your individual 'I-enquiry' has turned into a collective 'we-enquiry'. You have moved from 'I' to 'we'.

Our choice of methods of enquiry was largely determined by the research question we were addressing in each study and was drawn from the accumulated prior research experience of the action researchers in the group. As a group of mainly clinical scientists, our tendency to positivist assumptions made us attempt to pose generic questions (for example, what are the core generic consultation skills?) and to seek generalizable answers. The assumption that there can be a generic set of consultation skills may have been presumptuous but it was convenient for the purposes of assessment. We were aware that positivist questions would probably not suit enquiry about social interactions, and that many of our answers could only be 'true' for those who had spoken them. A more critical-interpretive constructivist paradigm was adopted for the understanding of student classroom and placement experiences.

At a critical mid-point in the ten year history of this action research, we were introduced to the realist paradigm which lies between positivist and constructivist stances. Realist philosophy is 'a methodological orientation or a logic of inquiry' (9) which can be chosen as one of the possible approaches to data collection and analysis in action research (10). Realist evaluation has an explanatory focus and involves participants deliberately in answering the question "what works for whom in what circumstances and why?" (11). Realist principles can be incorporated into action research by asking at the reflective stage in the action research cycle not just "has it worked?" but "for whom has it worked and not worked?" "In what circumstances does it work and not work?" and for both "why?". In realist analysis (11–13), an initial programme theory is composed from the literature plus the experience of the researchers. A programme theory expresses what is thought to be producing the outcomes of a complex system with diverse

participants, and how that might be happening. The initial programme theory is then tested and refined in analysis of the data by extracting middle range theories from coded CMO (context-mechanism-outcome) configurations in the data. A middle range theory in realist evaluation is a judgement about the repeatability of one or a set of related context-mechanism-outcome (CMO) configurations. Middle range theories "lie between the minor but necessary working hypotheses that evolve in abundance during day-to-day research and the all-inclusive systematic efforts to develop a unified theory" (Merton 1967 p39) They need to be trustworthy enough to justify making changes to the programme theory.

There are advantages of bringing a realist approach to action research. One is that by paying attention to the possibility that groups within a student cohort will react differently from each other to the same conditions, a more nuanced curricular approach might be developed. Another advantage is that while both action research and realist approaches emphasise the importance of context, to action research context is important as a situation in which the outcomes are embedded. This limits the transferability of findings except to another similar context. Realist evaluation aims to understand which are the elements within the context that influence which mechanisms can fire, and therefore determine the possible outcomes. This may enable interventions to be adapted for use in other contexts if the influence of context is well understood. The research questions we had were practical curriculum improvement questions. The strategy most likely to produce useful results for our institution was one in which changes made were evidence-based and critically appraised. This led us to choose action research methodology for our programme of curriculum reform, and realist evaluation to explore some of the puzzling questions which arose at the critical reflection stage in each cycle of action research. Other questions were addressed by modified grounded theory (before we had discovered the realist approach).

Action research is sometimes criticised as not 'proper' research because it is seeking solutions rather than explanations (1,2). Action research often adopts a social constructivist understanding of how we create meaning by interacting with the realities around us. The constructivist approach does not aim to establish causality in the same terms that positivist research may, nor does it seek the generation of general laws. The implications of this for this programme of research are that the findings will not be universal. Although they arise from real life experiences and not from experimental conditions, and therefore have validity for the people studied, our recommendations to others can only be tentative and must be adapted to the context and circumstances of the reader. To counter the problem that context may set limits on the generalizability of outcomes, the modified Delphi study developing GeCoS (Chapter 5) involved eight UK medical schools. Clinical tutors (hospital and GP) from each school proposed and amended the list of consultation domains and competencies.

Rather than attempting to gather data without being noticed, in action research and the realist approach researchers are encouraged to participate with individuals as they construct meaning within the reality of their situation. Participants are involved by researchers in analysing their

experience in order to understand it by their own explanation rather than relying solely on researcher interpretation. The rationale for this approach to curriculum development is that curricular policies and practices tend to be developed based on the interpretations of planners and can bear little relationship to student understanding and interpretation of their experiences (14,15). The implications of this research approach are that the participants' thought processes can be influenced by the researchers' own construction of meaning, and the interpretation of the data most certainly will be. We adopted a transparent reflexive approach to acknowledge and mitigate this. Group work may also reduce the effects of individual researcher bias due to assumptions. For example, our action research group developing the system of workplace assessment and asking about the impact of grading (Chapter 7) included a medical student co-researcher and co-author who double-coded the transcripts of interviews.

Outcome measures available to action researchers of curriculum development are limited. Student likes or dislikes for an educational intervention and self-evaluation of learning may be biased by factors other than the quality of the learning experience. A convincing level of proof can however be reached from a body of evidence and mixed methods may be used to achieve this. We used a number of research methods both qualitative and quantitative, approaching the research questions from various angles in order to triangulate results for greater trustworthiness. Triangulation of methods and sources is normal procedure in action research, not only to enhance credibility, but also to give a better depth of understanding by providing different types of information, and information from different perspectives (1). For example, in the first cycle of action research rolling out the first year of the curriculum (Chapter 2) the stakeholder action research group collected observational notes from the teachers, student evaluations from all students and focus group transcript data about the communication skills sessions and students' experiences of being allowed to set the emotional temperature.

One question about programme design arose which the action research team was so uncertain about, that we decided to try to test it experimentally, moving away from the usual conduct of action research. The study on the impact of grading in formative workplace assessment (Chapter 7) used a cross-over format in which each student was their own control, randomised to receive their first assessment feedback with or without grades, the second assessment with the other condition, and allowed to choose whether to have grades or not with the third. This enabled us to involve students in trying to understand and explain their individual perspective and experiences of what the grades meant to them. We could have used quantitative measures of outcome in this study (such as number of feedback items recalled under the two conditions, or position on a self-efficacy or satisfaction scale after each feedback was received). We felt, however, that we would get a better understanding from a realist evaluation which explicitly accounts for context and provides nuanced explanations of what has happened to individuals under the two conditions. The design was therefore pseudo-experimental, in that students had experienced both conditions before being interviewed.

Choice of methods of enquiry - advantages and disadvantages of the individual study methods used and consideration of some possible alternatives

Focus group evaluation (Chapters 2 and 9)

Focus groups enable exploration of concepts in depth with further refinements of students' thoughts (16). There is an economical advantage to researchers in being able to access larger numbers of participants than individual interviews could in the same time. Purposive sampling can create groups in which reticent individuals will feel able to express the views they might normally hide. By hearing the thoughts of others, individuals may become more aware of and able to express their own ideas.   There is a tendency in focus groups to over-emphasise consensus, however, and silence the minority view. This may not be mitigated by holding more groups if the minority feel unable to speak out in any of the groups.

The principles driving our analysis – modified grounded theory – allowed us to be reflexive about our assumptions based on educational theory, and to deliberately seek out dissonant data in order to refine that theory.

Interviews with thematic analyses (Chapters 3, 8, 9)

Semi-structured interviews suit this type of enquiry in which students and tutors are involved in explaining their actions, their preferences and the impact of educational interventions. They develop rich understanding of what has been experienced by different individuals within a cohort. The depth of understanding enabled is greater than would usually be possible with a questionnaire-based survey as responses can be explored. Thematic analysis can also be used to build on existing theory. However it is open to participant biases such as self-serving bias, social desirability bias and attribution bias. It means that the findings reflect participants' perceptions of the underlying phenomena, which have not themselves been tested. If participants' perceptions are biased then so will our findings be. Analysis requires reflexivity as it is also open to researcher bias.

Modified Delphi studies (Chapters 5 and 6)

The Delphi group consensus process can bring geographically distant panel experts together without direct confrontation and power play to condense their opinions into a few clear statements by an understood democratic process which has time-limited stages (17). The disadvantages are the likelihood of busy participants becoming less engaged if surveys are long or the process of consensus has many stages, also concerns about the reliability of the technique (partly because of the variety in methods of sampling and determining consensus). The positive features of our study design were: the involvement of representative clinical tutors, assessors and students; that being online was convenient for dispersed and busy participants of the Delphi exercise; and that we had purposive balanced representation of GPs and hospital doctors among both participants and round table researchers. An alternative approach would have been to use a nominal group process but this would have made it practically difficult to involve faculty from other UK medical schools.

One particular strength of the project to develop strategies for improving consultation skills (Chapter 6) was the involvement of students to validate and improve the strategies. This was done by a nominal group technique using 'audience response' electronic key pads, so that students voted without seeing each other's opinions, thus reducing bias from social conformity. Weaknesses are: that the starting point for both Delphi exercises was an existing set of concepts. These may have stifled original thought. Nor did we test any of the opinions gathered with the panel of experts (although we did test student opinion of the strategies, and both studies involved a round table discussion of mixed hospital and GP researchers at every stage).

Realist evaluation (Chapter 7)

Realist evaluation starts with a programme theory – what the researchers think is going on within the system being studied, and involves extracting middle range theories from the coded CMO (Context–Mechanism-Outcome) configurations in the data. A middle range theory in realist evaluation is a judgement about the repeatability of one or a set of related context-mechanism-outcome (CMO) configurations. Middle range theories "lie between the minor but necessary working hypotheses that evolve in abundance during day-to-day research and the all-inclusive systematic efforts to develop a unified theory" (Merton 1967 p39 (18)). They need to be trustworthy enough to justify making changes to the programme theory.

We have understood realist evaluation better as we have conducted the study of grades in formative workplace assessment and subsequent research. Reporting standards for realist evaluations have since been proposed (19). One criticism of our use of realist methodology is that the trustworthiness of the middle range theories we developed might have been confirmed by respondent validation.

11.2 What this body of work has added to existing knowledge

In this section the hindsight lens is applied to each study to see what contribution it has made to understanding of learning theory and development of curriculum, teaching practice, and medical education research methods. Citations have been examined to this end.

A critical lens has also been used. Each publication already contains its own critical commentary. Mention is made in the evaluation of each study, if in hindsight, that critique should be revised.

11.2.1 Contributions to the literature about medical student learning of consultation skills made by this action research

"Some like it hot: Medical student views on choosing the emotional level of a simulation" (Chapter 2).

In this study we gave first year medical students a choice of three levels of emotional challenge of role play interviews in an introductory communication skills class. 55% of the students chose the strongest emotional level; 28% chose moderate and 17% chose mild.

This study gave insights into the benefits of student self-directedness in simulation. The intention was that students would pick a level of challenge which would be within their zone of proximal development (20). Students could choose how far out of their prior competence they wished to go, making use of preliminary coaching and getting help from the group if they were struggling and needed to pause the interview. We found that there was a benefit to the group as well as to the individual, because some chose to be challenged beyond the level at which tutors would have set the emotional level. This broadened the learning experience for the observers.

This study has been cited as an example of the application of self-directed learning theory to improve education (21), and in an AMEE guide as an example of the use of focus groups in medical education research (22).

The educational process in this study - enabling individual choice in a group learning setting - is a fascinating mesh of self-determination with peer influence. This type of learning is worthy of further study. We made a number of observations about what might or might not be found in other contexts. The focus groups enabled us to gain insight into the complex set of social influences in operation when students were making their choice of level of challenge. Peer pressure, challenge seeking and fear of failure were all evident. We were aware of social constructivist theory such as Vygotsky's which stress that social group learning is useful as students will find it easier to believe that they can do something which they have seen a peer do (self-efficacy for learning) (23–25). This was reported by our participants. We also discovered the negative side of peer comparison as evidenced by students or their peers saying they were "chicken" when they chose to face the mild emotional level. Our call for tutors to provide "scaffolding" (20,26) of the individual and group choice of levels of challenge would apply to any attempt to mix self-determination and group learning. The impact of the group on self-determination is fascinating and could be explored by further study, along with the role of the tutor in balancing group and individual needs. A realist evaluation of what works in this context would be valuable.

"Minding the gap between communication skills simulation and authentic experience" (Chapter 3)

Simulation of communication skills is often conceptualised as good workplace preparation, 'bridging' or 'closing' the gap between classroom and clinical practice. In integrated curricula medical students are required to handle *contemporaneous* experiences of simulated and real patient interactions in classes and on placements back and forth within the same year. In this study we addressed two questions: 'How do new undergraduate medical students understand contemporaneous interactions with simulated and authentic patients?' and 'What is the impact of differences between simulated and authentic patient interactions on student learning of communication skills?'

From secondary analysis of three data sets of interviews and focus groups with junior medical students we deduced that there were gaps between simulation and clinical practice which tutors would do well to 'mind' in discussion with their students. For example, students perceived that simulated patients were less forthcoming with information than real patients, and confidentiality, while emphasised in the classroom appeared not to be an issue in clinical practice. Minding the gap would enable students to construct synergistic learning from simulated and authentic patient interactions. Our students' difficulties were turned round into a set of guidelines for tutors aiming to help their students who experienced this gap (Box 1).

Box 1: Guidelines for tutors (a precis of the paper's section 'practical implications for educators')

Don't ignore a gap as this risks paradoxical meaning-making, rejection of ideals in the face of reality, creation of dichotomies and misunderstandings.

Manage the gap: educators who recognise and understand gaps can work collaboratively with students to discuss perceived differences and make constructive comparisons. Facilitate student meaning-making including encouraging students to theorise about how and why identified differences occur. Encourage adaptability and self-directedness on placement over and above that which students may have needed in the classroom. Support students when possible during placements and debrief them afterwards, not relying solely on preparedness.

Use it - being 'mindful' of the gap: Tutors need to be mindful of how they portray their side of the gap to students who are on the other side, and of their influence as role-models (positive and negative).

This paper has been cited by Cushing (27) to illustrate the importance of students witnessing doctors role modelling the exploration and discussion of patient perspectives. It has not (yet, as far as we are aware) been cited as a call to move from 'competitive contrast' to 'constructive

comparison' of difference. This would be the paper's main contribution to educational theory and may need validating in another study.

"Designing whole-task learning opportunities for integrated end-of-life care: a practitioner-derived enquiry" (Chapter 4)

End-of-life care is one area of practice for which newly qualified doctors feel unprepared (28). What we did was to develop a study day which combined interactions between students and simulated patients with practical tasks including prescribing and use of best practice clinical management tools. The aim was to help students to integrate knowledge, skills and behaviours needed when encountering patients near the end-of-life including during transitions between primary and secondary care settings. Analysis of student reflective writing which drew together learning from classroom and workplace demonstrated learning in multiple domains.

The whole-task model is one which seeks to incorporate complexity and which avoids fragmentation or compartmentalising parts of a multi-task goal. Our work demonstrates that whole task theories can be applied when developing strategies for integrating learning in complex areas of practice such as end-of-life care. It has been cited in support of using whole-task learning to help students learn to function in complex and variable settings (29) and in a systematic review of palliative care in undergraduate medical education (30).

We also found that only a few students explicitly reflected in their writing on the actual death of their patient. We conclude that the strength of societal taboo about mentioning death may not have been overcome by these learning experiences.

This study did not have the scope to demonstrate effect on patient care. Observational studies of students and junior doctor interactions with patients and their families and other care-givers and healthcare workers would be of value in this respect.

11.2.2 Contribution to the literature about assessment of consultation skills

"Development and face validation of an instrument to assess and improve clinical consultation skills" (Chapter 5).

In this study the action research team was purposefully drawn from GP and hospital tutors at Keele University. We reviewed the Calgary Cambridge guide to the medical interview (31,32) and the Leicester Assessment Package (LAP) (33) and identified concepts common to both or only represented in one or the other. We revised the Calgary Cambridge guide to include concepts it did not contain ('Management', 'Clinical reasoning', 'Record keeping' and 'Case presentation') and populated it with competencies generated from the GMC's Tomorrow's Doctors guide (34), the LAP and the Calgary Cambridge guide. This was validated in a two-stage Delphi study across eight UK medical schools. The Delphi study made 24 amendments to the suggested set of competencies. The resulting instrument, the **Ge**neric **Co**nsultation **S**kills assessment framework

(GeCoS), is best seen as a toolkit for the student from which they select the processes and content which they want to apply to the task facing them. Its nine domains can be sampled by summative assessment blueprints and cover all the competencies which might be expected by clinical supervisors observing students in workplace-based assessments.

The GeCoS domains and competencies have been used since 2010 as the framework for workplace based assessment and also for summative consultation skills assessments at Keele University School of Medicine. This enables competency-based feedback to be given as global scores for performance in the domains of the consultation, plus related verbal feedback (35,36). This work has the strengths of being a multi-institution exercise and being generic - drawing on the conceptualisation of consultation competencies of undergraduate clinical tutors from both primary and secondary care. The resulting framework and instrument has had a major institutional impact at Keele, forming a robust core for the spiral consultation skills curriculum throughout the five year undergraduate medical course, providing alignment of classroom teaching, workplace assessment and OSCEs. GeCoS has underpinned a systematic approach to feedback on consultation skills, involving formal workplace assessments and i-pad capture of OSCE examiner scores and audio feedback. Both of these rich forms of feedback are stored in the electronic feedback portal of the individual student who can review it at will. Following the introduction of this feedback system, student satisfaction scores with feedback on their work rose sharply from low national ranking to the highest among UK medical schools (35,36).

GeCoS has not (yet) been adopted by any of the other institutions, even those involved in its genesis, and this raises questions. Is the instrument really generic enough to suit other medical school curricula? Is the lack of uptake by other institutions more about how such work is disseminated? The 59 competencies in GeCoS are as few as could be achieved and are generally acceptable to assessors who have used it in various contexts. The clinical reasoning domain does not map well onto the Keele curriculum for Higher Consultation Skills however, since the latter was developed after GeCoS had been developed, and used a different terminology. GeCoS has also been felt to lack a domain for assessment of professionalism. Since it is now 8 years old, it could be time for a further Delphi study to renew the instrument.

11.2.3 Contribution to the literature about feedback in the clinical workplace

"Development and face validation of strategies for improving consultation skills" (Chapter 6)

This set of strategies for improving each competency in the GeCoS framework was developed by a modified Delphi study with hospital and GP tutors from across the Keele faculty. The strategies for improvement suggested by tutors were face validated by a modified nominal group study involving medical students. In addition to face validating the strategies as useful to students, student comments were used to improve the strategies.

The resulting set of strategies for improving consultation skills contributes to the feedback literature in that it can be considered a 'primer' for consultation skills. The set of strategies have

been offered as a resource to other medical schools interested in supporting their tutors to improve the quality of the feedback they give their students. We are not aware of any other medical school's having adopted the GeCoS feedback instrument thus far, but the paper on strategies for improving consultation skills has been downloaded over 500 times (as logged on the Journal website by Sept 2017).

"Grades in formative workplace-based assessment: a study of what works for whom and why" (Chapter 7)

In this study 78% of participating students chose to have grades with their formative feedback summary. A realist evaluation was conducted to explore what was working for these students and why. We found a number of explanations which challenge the view that grades detract from feedback and may deserve consideration in the context of healthcare professional training where competence is a required outcome as well as aspiration to excellence (Box 2).

Box 2: Student perceptions of what grades can contribute to workplace assessment:

- Grades help students to locate themselves
- Grades calibrate their self-assessment, promoting self-awareness of competence and limits
- Grades help them track progress, especially if referenced to an understood standard such as that expected of the graduating doctor
- In the context of constructive feedback from a trusted assessor, students will usually create constructive meaning out of low grades
- In students who have well-developed self-regulation, low grades can focus attention and effort
- Satisfactory and high grades can enhance self-efficacy

However, we also found that 22% of students preferred not to have grades for a number of reasons, some feeling that they didn't need them as they knew where they were, others didn't trust the judgement of their assessor, and many wished to avoid potential harm (Box 3).

Box 3: Student perceptions of the potential harmful effects of grades:

- For these students grades were concrete, powerful and blunt and needed explanation for students to create helpful meaning from them.
- Grades could be discouraging.
- Grades were also described as potentially creating complacency in individuals who focus on the grade and ignore the accompanying advice.

Realist evaluation aims to develop our capacity to predict which individuals might react to context in which ways. This study might help us to identify in a future cohort those students likely to be harmed by receiving grades. Because of the complexity of the reasons for harmful effects, the researchers recommended that the institutional response to this study was not to attempt to identify such students, but instead to allow students to self-select. Some students in the study who chose not to have grades, perceived that grades did not galvanise effort for them, a response which some may have developed because of previous experiences. Such individuals appear to have barriers and filters, created by previous experience, and feedback with grades may be perceived as threatening rather than formative. In applying the findings of this study to subsequent years and allowing students to choose to have no grades if they prefer, we felt that we might be avoiding potential harms to these self-identified students.

A strength of this study is its cross-over design so that students could speak from both experiences about what having or not having grades did to them and why. It was the first study directly comparing feedback from workplace-based assessment with and without grades in undergraduate medical education.

The metaphor of the semi-permeable membrane was introduced to explain how students self-regulated to reduce potential negative effects and enhance the positive effects of feedback with grades by selective filtering and pumping. The diagram illustrating this metaphor was intended to help clinical tutors (who understand semi-permeable membranes) to understand the educational psychology theory developed in this study, and thus maybe to understand their learners better.

Given the prevalence of grades in medical education, this study has wide implications. It would need to be replicated in other settings, for example in summative assessment before any conclusions could be drawn about contexts other than the single school workplace assessment system within which is was studied.

This study has been cited 14 times to date, notably as an example of detailed description of data analysis in the RAMESES II set of reporting standards for realist evaluations (19). It has also informed a similar study of student perceptions of workplace assessment and feedback, in which the findings were somewhat different. In that study, grades were usually given at the end of the block and were therefore felt to be distantly related to the observations made (37). This reminds us as researchers to be cautious in the conclusions we draw.

"Positive impacts of written feedback: Does formal workplace based assessment add value to informal feedback?" (Chapter 8)

In this secondary analysis of the interviews about grades in formative workplace-based assessment we found that students valued the verbal feedback discussion highly and that they often considered the written summaries superfluous. This might initially lead us to conclude that provision of written feedback is simply a bureaucratic exercise which served little educational purpose. However, we also observed that the written format seemed to influence assessors' approach to giving verbal feedback. The act of preparing written feedback augmented the feedback discussion and tutors adopted the language of the formal instrument in the verbal feedback and free text written feedback. As a result we hypothesised that verbal feedback is enhanced by the requirement to produce written feedback; it may act as a mechanism to enhance

quality even if it is, in itself, superfluous. But this theory is speculative and needs further study. This potential means we haven't abandoned written feedback in practice, but these relationships require further investigation.

Justification for medical education interventions needs to look beyond the obvious or immediate. Here is one example of a possible secondary beneficial effect of an intervention.

"Utility of an app-based system to improve feedback following workplace-based assessment" (Chapter 9).

This study aimed to determine whether an app-based software system to support production and storage of assessment feedback summaries makes workplace-based assessment easier for clinical tutors and enhances the educational impact on medical students.

In it we monitored our workplace assessor (GeCoS) app's usage by Year 3 to 5 medical students in 2014-15 and conducted focus groups with Year 4 medical students and interviews with clinical tutors who had used the apps.

The study contributes to the literature about the utility of apps in clinical WBA in two ways. Firstly, we demonstrated that, given the choice, tutors tend to use the electronic platform with which they are familiar. This suggests that an app-based system will need to be flexible enough to accommodate different experience and skills in typing, dictating and incorporating technology into work and life.

The second contribution of this study is an insight into the complexity of the social interaction, the feedback conversation. When a mobile device or computer screen is used in a feedback conversation, it can become an intrusive third party.

Our critical reflection about this study is that the wide variation in student experience within the focus groups suggests that we may not have sampled all the student and tutor views and experiences, and that after tutor training for more consistency in use of the app in the feedback conversation a repeat of the study might well produce different findings.

"Guidelines: the do's, don'ts and don't knows of feedback for clinical education" (Chapter 10)

This review of the literature about feedback to trainees in the clinical workplace aimed to move the focus from how supervisors should give feedback to how trainees receive feedback, and how institutions can support a feedback culture. We found the strongest evidence for useful feedback being a two-way conversation, bespoke, specific and actionable, but concluded that feedback doesn't fit a single model. The research on feedback describes processes which vary from individual to individual and from context to context. Feedback is very important in medical education but can also threaten the trainee. A subsequent critical appraisal of the literature drew similar conclusions (38).

Our work on these guidelines moved us to offer a new definition: "Helpful feedback is a supportive conversation that clarifies the trainee's awareness of their developing competencies, enhances their self-efficacy for making progress, challenges them to set objectives for improvement, and facilitates their development of strategies to enable that improvement to occur". This definition has been quoted in a study of the quality of written feedback (39).

The criteria for inclusion of evidence for these guidelines were not those of a systematic review. We felt that a systematic review would have concluded that there were very few papers in the broad literature about feedback which met the inclusion criteria, and the reader would not gain very much. Instead, we reached consensus on what was important to us as medical educators with our individual teaching experience and awareness of the literature. This gave a more balanced and useful overview of what is there. The process is described in the paper. To make it clear that the criteria were different, we avoided using the ABC indicators of strength of evidence and devised our own indicators which combine the evidence with the authors' consensus (table 1). The aim was to produce practical guidance on what we know to work, what we know doesn't work, and what we don't yet know (D3- Do's, Don'ts and Don't Knows). This was the first of a series of guidelines in the journal and we were requested by the editors to write up our method of working in compiling them, for the benefit of future authors of D3 guidelines, who have now used them (40).

Table 1: Criteria for strength of recommendation

Strong	A large and consistent body of evidence
Moderate	Solid empiric evidence from one or more papers plus the consensus of the authors
Tentative	Limited empiric evidence plus the consensus of the authors

These guidelines seem to have addressed a need as they are being read and cited more frequently than any of the other publications in this body of work (34 citations to date). This is probably because of the nature of primary research which is detailed and answers only a few specific questions, whereas a summary of guidelines addresses a wider set of questions. It is also reader-friendly with the guidelines clearly outlined and then developed one by one with the supporting evidence.

The guidelines have been cited in an article by Eppich and colleagues about clinical event debriefings and coaching conversations, to emphasise the importance of eliciting learners' self-assessment and then confirming or challenging it to help them to improve (41). The guidelines have also been cited in studies of feedback for trainees developing their competency with entrustable professional activities (EPAs) (42,43); used to define the quality criteria in a study of the quality of online daily written feedback for medical students (44); cited to illustrate the importance of feedback in medical education (45) and in a study of postgraduate educational supervision (46). They have also been used as support for the Promoting Excellence and Reflective Learning in Simulation (PEARLS) approach to health care debriefing (47) and for the adoption of a two-way feedback process in a nurse-midwifery programme in California (48).

11.3 Action Learning: what this programme of action research has changed in me as a practitioner-teacher-researcher

In the spirit of action research for personal professional development (49), it is appropriate to include a section in this thesis reflecting on how I have changed as a result of this work. I have a portfolio career as a half-time GP and half time medical school lead for consultation skills teaching. My research interests have all been within medical education, and these three areas of work – clinical practice, teaching and research - have been mutually enhancing. My values are that I have always aimed to practice patient-centred medicine and to teach my students to do the same. This programme of research has enhanced my personal practice and teaching, as I have myself adopted the consultation skills 'toolkit' which we have developed and taught to students, and have become more aware of the student experience of learning and therefore better able to help my students to learn.

My attitude to medical education research has also changed. Jean McNiff describes three reasons for doing action research - to improve practice, to generate new knowledge and theory as practitioners, and to explain how you are contributing to new understandings for yourself and others (50). My foray into action research was initially for the first of these reasons - to make the best job of a curriculum development task by critical evaluation of our experience with reference to the literature, making and monitoring logical changes. I have always had collaborative tendencies and that was another reason why I was attracted to forming an action research group. As we went through the cycle of curriculum improvement and evaluation I found that action research can also advance knowledge and theory, and I started to get a desire for this and for disseminating our findings.

Working as a research team has made me very aware of the role of co-creation of evidence as exemplified by the way student involvement added a perspective to the studies in Chapters 2 and 6. Much of my medical training was informed by positivism, but as I have changed into a medical education researcher I have needed to adopt other epistemologies. I have a tendency to seek to prove universal truths and to regard the findings of others as transferable facts. I have been helped to change perspective by my social scientist colleagues and by reading work in these paradigms.

I have always been a reflective individual but as a researcher I have acquired a critically reflective approach to what I do and to the work of others. I enjoy peer reviewing medical education research and learn by doing that.

The other question is what impact my position has had on our interpretation of findings? Having lived with and invested a great deal in this curriculum over the past ten years, I am well placed to understand and disseminate findings but may be reluctant to perceive its deficiencies. In our action research team we always double coded data before discussion of our interpretation of findings, and then discussed with the whole action research team. This provided a check for individual researcher bias, especially when the double coding was done by a medical student co-

researcher whose biases were probably different from mine. As the convener of the action research group my opinions may have carried more weight than those of others. Whether or not this was the case, I found that the interpretations I had developed were sometimes changed and were always augmented and improved by discussion. This is not *my* research so much as *our* research, and I give full credit to my co-researchers as outlined in the acknowledgements section of this thesis.

11.4 Future research

Continuation plans are a key element in action research to keep the cycle of improvement going (3). The next cycle of this programme of action research is already underway with the introduction (suggested by a student focus group) of a facility for student reflection on feedback. What I aim to do in this final chapter section is to look at a broader set of focused researchable ideas arising from this body of work.

Future research about feedback in the clinical workplace
Our review of the literature about feedback concluded that it should be perceived as a conversation rather than as a commodity. This concept could shape research into the unanswered questions of clinical supervisors about how to give effective feedback, such as:
How much is the right amount of content when giving feedback? Students may often wish for more feedback but on occasion be swamped by unwanted advice which can make them feel inadequate. How does the supervisor determine how many items of feedback are optimal (both strengths and weaknesses)? To keep the learner in the 'zone of proximal development' they need to be stretched but not beyond their capacity. How is this achieved? What determines the 'open and safe interaction' in the feedback conversation?
A set of studies in various clinical training contexts could be helpful in answering these related questions. Triangulation of observational study with phenomenological inquiry might be the most helpful approach to these questions.

Motivation arising from workplace assessment and feedback is both promotion (of doing better) and prevention (of losing face). Regulatory focus theory may explain some of the individual variability in feedback responses (51,52). Regulatory focus is not fixed, but changes according to the situation. How can regulatory focus best be directed in order to enhance the impact of feedback? We have suggested that there is a semi-permeable membrane for such feedback and that the "settings" of the protective filters depend on prior experiences and promotion/ prevention focus but are influenced by the feedback relationship (Chapter 7). How do the issues of vulnerability (self-efficacy), motivation to improve or to prevent harm, and credibility interact to give shape to constructive or destructive feedback in a workplace learning situation? How do we help trainees to believe that they can improve? This is a big area of research.

One specific research question within this area is whether overt comparison with peers—when made by the supervisor – is helpful to the trainee? There is conflicting evidence on this and on a related question - do comparisons with standards help students on a competency-based training programme to improve, or is it better for each trainee to strive for personal excellence? How can we determine which learners in which circumstances will find comparisons motivating, as opposed to disheartening? It would be worth taking a realist approach to studying instances of constructive or destructive outcomes of feedback to understand what influences the trainee's response.

Future research about written feedback

Three potential implications of treating feedback as a conversation rather than a commodity are that feedback should be addressed <u>to</u> the individual not <u>about</u> them; that where appropriate, questions should be used rather than final vocabulary in order to stimulate formative action; and that there should be an opportunity for two-way discussion of performance. The research agenda relating to written feedback is whether this dialogic nature of feedback survives when it is written, and when it is also used for assessment purposes (so a third party is 'listening' to the feedback conversation).

The process of generating a written summary of feedback appears from our research to be worthwhile (Chapters 8 and 9). Establishing whether this is indeed the case through prospective study should be a future research aim. The educational impact on both tutors and students of providing pre-formulated advice within the app is worth further study.

Future research about the feedback culture in a workplace

Several broad questions about institutional culture were raised in the course of discussing the Do's, Don't and Don't knows Guidelines for feedback in clinical education. One specific testable question of interest in undergraduate education is whether it is most effective to give feedback to individuals alone or in a group setting? In group learning of clinical skills, feedback to the trainee(s) who have experimented with a task is generally given by and in front of the group. This can include group feedback on a videoed real patient consultation. The advantages to this approach are that a range of feedback perspectives are gained, feedback-giving is role-modelled, and observers learn vicariously. The disadvantages are reduced control over content and volume of feedback, plus the risk of a negative emotional impact. In situations where it might be possible to give feedback either individually or in a group setting, we do not know whether the advantages outweigh the disadvantages. There are studies which have found learner preferences for group feedback (53) and for individual feedback (54) and it is clear that the context matters (55). This question might be amenable to comparative studies where participants act as their own controls.

Future research about assessment of competency

Demonstrating competency in test conditions is not the same thing as demonstrating capability to use those competencies appropriately in real patient care, and medical educators are duty bound to try to measure the effectiveness of our educational programmes at enabling and improving patient care. For example, is there any longer term impact on end of life care (e.g. after commencing foundation training) arising from experiential learning with study day support in medical school? Workplace based assessment may be used as the basis of such measures at the level of the individual student or doctor, and competency assessment instruments continue to be developed for various settings (56) but the challenge is how to develop robust systems of assessment to draw meaning out of the workplace assessment data which are necessarily complex and often affected by trainer-trainee relationships, face-saving judgements and failure to fail (57–59). In programmatic assessment and clinical competence committees, judgements are made on the basis of accumulated data, often including workplace assessments. How to make appropriate use of workplace assessment will be an important question for such committees.

Future research about end of life care teaching and learning

Some research questions stem from our action research on the end of life care curriculum (Chapter 4). Development of understanding with respect to reluctance to 'name death' and refining of learning opportunities to address this appropriately would be a topic for a future cycle of action research by the team facilitating palliative care education. Investigation of 'what really happens' when students meet with professionals and patients engaged in end-of-life care processes e.g. how and why do they interact in certain ways could be done by observational studies with realist evaluation.

The final suggestion is that in action research, replication of the same studies - in other settings; sampling across other groups; and triangulated by other methods - is worthwhile in order to understand mediating conditions and limits to the applicability of findings.

11.5 Conclusion – cohesion of this body of work and its impact on the field

This action research programme has been a curriculum quality improvement effort by various teams of faculty and students. Discrete projects in the development of classroom lesson plans, the consultation skills assessment instrument and the workplace consultation skills feedback system are united by the overarching aim of developing a consultation skills curriculum in which intended learning outcomes, methods of learning in the classroom and in clinical workplaces and assessment are all aligned and focused on graduating excellent patient-centred clinicians. The published papers are snapshots of studies and findings along the course of the programme of action research and cannot do full justice to the innovation, monitoring and improvement which has taken place over the past ten years.

The impact at a local level is that Keele University school of medicine have adopted the GeCoS set of generic consultation skills as a spine to the spiral of teaching and learning in the classroom, extending to the system of workplace assessments in GP placements and to OSCE assessments for all five years. Students and clinical tutors in all five years therefore use the same domain and competency descriptors, enabling feedback about consultation skills to be in a common language, and expectations to be more uniform. This consistency in curricular alignment follows an important educational principle powerfully set out by Biggs (60) but not yet adopted in all curricula.

A feedback system has been introduced at Keele. By designing the format for giving and receiving student-focused feedback in the introductory communication skills classes in year 1, the action research group set the institutional tone for feedback discussions from then onwards. This has been refined in subsequent cycles, and the guidelines reproduced in Chapter 10 are the culmination of my and the other authors' accumulated understanding of how feedback can be most useful to trainees.

By working together on the action research cycles of reflecting, planning, implementing and observing outcomes, the various teams of researchers have developed skills and knowledge. A community of practice has grown and cross-fertilised organically. This has strengthened the individual team members and the institution teaching and research capabilities.

The contributions of this published body of work to the field are:

  a validated set of competencies which have been disseminated at national and international conferences and in these publications,

  guidelines for giving feedback to trainees in the clinical workplace,

  and advances in our understanding about the use of technology for workplace feedback.

GeCoS is offered to other medical schools as a generic assessment instrument with the advantages that it has brought to the Keele curriculum.

225

These publications may also be seen as a contribution to the pool of knowledge concerning higher education in general. The contributions are:

an increased understanding of the impact of challenge in group learning,

more detailed understanding of the social interactions in feedback conversations and how these are affected by such requirements as grading,

insight into the value of capturing written summaries of feedback and the impact of using mobile devices for this.

References

1.   Kember D. Action Learning, Action Research: Improving the Quality of Teaching and Learning. London: Kogan Page; 2000.

2.   Davison RM, Martinsons MG, Kock N. Information Systems Journal : Principles of Canonical Action Research. Inf Syst J. 2004;14:65–86.

3.   Norton L. Action Research in Teaching & Learning. Oxford: Routledge; 2009.

4.   McKearnan J. Curriculum Action Research. 2nd ed. London: Kogan Page; 1996.

5.   Lewin K. Resolving social conflicts. New York: Harper; 1948.

6.   Schon D. Educating the Reflective Practitioner. Towards a New Design for Teaching and Learning in the Professions. Oxford and San Francisco: Jossey-Bass; 1991.

7.   Carr W, Kemmis S. Becoming Critical: Knowing Through Action Research. Lewes: Falmer; 1986.

8.   McNiff J. Action research for professional development: concise advice for new action researchers [Internet]. 2012 [cited 2018 Mar 4]. Available from: http://jeanmcniff.com/ar-booklet.asp

9.   Pawson R. Evidence-based policy: a realist perspective. London: Sage; 2006.

10.   Westhorp G, Stevens K, Rogers PJ. Using realist action research for service redesign. Evaluation. 2016;22(3):361–79.

11.   Wong G, Greenhalgh T, Westhorp G, Pawson R. Realist methods in medical education research: what are they and what can they contribute? Med Educ. 2012 Jan;46(1):89–96.

12.   Pawson R, Tilley N. Realistic Evaluation. London: Sage; 1997.

13.   Pawson R. The Science of Evaluation: A Realist Perspective. London: Sage; 2013.

14.   Denzin NK. Interpretive Interactionism. Newbury Park: Sage; 1989.

15.   Stringer E. Action Research in Education. 2nd ed. Upper Saddle River, New Jersey: Pearson Education Inc; 2008.

16.   Barbour RS. Making sense of focus groups. Med Educ. 2005;39(7):742–50.

17.   Humphrey-Murto S, Varpio L, Gonsalves C, Wood TJ. Using consensus group methods such as Delphi and Nominal Group in medical education research. Med Teach. 2017;39(1):14–9.

18.   Merton R. On sociological theories of the middle range. In: Merton RK, editor. On theoretical sociology: five essays, old and new. New York: The Free Press; 1967. p. 39–72.

19.   Wong G, Westhorp G, Manzano A, Greenhalgh J, Jagosh J, Greenhalgh T. RAMESES II reporting standards for realist evaluations. BMC Med. 2016;14(1):96.

20.   Bandura A. Social Learning & Personality Development. Holt, Rinehart & Winston, INC: NJ; 1975.

21.   Mazmanian P, Feldman M. Theory is needed to improve education, assessment and policy in self-directed learning. Med Educ. 2011;45(4):324–6.

22.   Stalmeijer RE, Mcnaughton N, Van Mook WNK a. Using focus groups in medical education research: AMEE Guide No. 91. Med Teach. 2014 Jul 29;(91):1–17.

23.   Vygotsky L. Thought and Language. MIT; 1986.

24.   Bandura A. Self-efficacy: the exercise of control. NY: Freeman; 1997.

25.   Schunk DH. Learning Theories: an educational perspective. Vol. 5th. NJ: Pearson Education Inc; 2008.

26.   Wood D, Bruner J, Ross G. The role of tutoring in problem-solving. J Child Psychol

Psychiatry. 1976;17:89–100.

27. Cushing AM. Learning patient-centred communication: The journey and the territory. Patient Educ Couns. Elsevier Ireland Ltd; 2015;98(10):1236–42.

28. Gibbins J, McCoubrie R, Forbes K. Why are newly qualified doctors unprepared to care for patients at the end of life? Med Educ. 2011 Apr;45(4):389–99.

29. Vandewaetere M, Manhaeve D, Aertgeerts B, Clarebout G, Van Merriënboer JJG, Roex A. 4C/ID in medical education: How to design an educational program based on whole-task learning: AMEE Guide No. 93. Med Teach. 2015;37(1):4–20.

30. Fitzpatrick D, Heah R, Patten S, Ward H. Palliative Care in Undergraduate Medical Education--How Far Have We Come? Am J Hosp Palliat Med. 2017;34(8):762–73.

31. Kurtz SM, Silverman JD. The Calgary-Cambridge Referenced Observation Guides: an aid to defining the curriculum and organizing the teaching in communication training programmes. Med Educ. 1996 Mar;30(2):83–9.

32. Kurtz S, Silverman J, Draper J. Teaching and Learning Communication Skills in Medicine Second Edition. Oxford: Radcliffe Publishing Ltd; 2005.

33. Fraser R, McKinley R, H Mulholland. Consultation competence in general practice : Establishing the face validity of prioritized criteria in the Leicester assessment package. Br J Gen Pract. 1994;44:109–13.

34. GMC. Tomorrow's Doctors The duties of a doctor registered with the General Medical Council. 2009.

35. Harrison CJ, Molyneux AJ, Blackwell S, Wass VJ. How we give personalised audio feedback after summative OSCEs. Med Teach. 2015;37(4):323–6.

36. Harrison CJ, Könings KD, Molyneux A, Schuwirth LWT, Wass V, van der Vleuten CPM. Web-based feedback after summative assessment: how do students engage? Med Educ. 2013 Jul;47(7):734–44.

37. Schopper H, Rosenbaum M, Axelson R. "I wish someone watched me interview:" medical student insight into observation and feedback as a method for teaching communication skills during the clinical years. BMC Med Educ. BMC Medical Education; 2016;16(1):286.

38. Kornegay JG, Kraut A, Manthey D, Omron R, Caretta-Weyer H, Kuhn G, et al. Feedback in Medical Education: A Critical Appraisal. AEM Educ Train. 2017;1(2):98–109.

39. Gulbas L, Guerin W, Ryder HF. Does what we write matter? Determining the features of high- and low-quality summative written comments of students on the internal medicine clerkship using pile-sort and consensus analysis: a mixed-methods study. BMC Med Educ. BMC Medical Education; 2016;16(1):145.

40. Yardley S, Westerman M, Bartlett M, Mark Walton J, Smith J, Peile E. The do's, don't and don't knows of supporting transition to more independent practice. Perspect Med Educ. 2018;8–22.

41. Eppich WJ, Mullan PC, Brett-Fleegler M, Cheng A. "Let's Talk About It": Translating Lessons From Health Care Simulation to Clinical Event Debriefings and Coaching Conversations. Clin Pediatr Emerg Med. Elsevier Inc.; 2016;17(3):200–11.

42. Duijn C, Welink L, Mandoki M, ten Cate O, Kremer W, Bok H. Am I ready for it? Students' perceptions of meaningful feedback on entrustable professional activities. Perspect Med Educ. 2017;6(4):256–64.

43. Peters H, Holzhausen Y, Boscardin C, Cate O, Carrie H. Twelve tips for the implementation of EPAs for assessment and entrustment decisions. Med Teach. Informa UK Ltd.; 2017;0(0):1–6.

44. Shaughness G, Georgoff PE, Sandhu G, Leininger L, Nikolian VC, Reddy R, et al. Assessment of clinical feedback given to medical students via an electronic feedback system. J Surg Res. Elsevier Inc; 2017;218:174–9.

45. Hayes V, Bing-You R, Varaklis K, Trowbridge R, Kemp H, McKelvy D. Is feedback to medical learners associated with characteristics of improved patient care? Perspect Med Educ. 2017;

46. Patel P. An evaluation of the current patterns and practices of educational supervision in postgraduate medical education in the UK. Perspect Med Educ. 2016;5(4):205–14.

47. Cheng A, Grant V, Robinson T, Catena H, Lachapelle K, Kim J, et al. The Promoting Excellence and Reflective Learning in Simulation (PEARLS) Approach to Health Care Debriefing: A Faculty Development Guide. Clin Simul Nurs. Elsevier Inc; 2016;12(10):419–28.

48. Myers K, Chou C. Collaborative and Bidirectional Feedback Between Students and Clinical Preceptors: Promoting Effective Communication Skills on Health Care Teams. J Midwifery &Women's Heal. 2016;61:S22–7.

49. McNiff J. Action research for professional development. Dorset: September Books; 2010.

50. McNiff J. Action Research: all you need to know. London: SAGE Publications Ltd; 2017.

51. Kluger A, Van Dijk A. Feedback, the various tasks of the doctor, and the feedforward alternative. Med Educ. 2010;44:1166.

52. Watling C, Driessen E, van der Vleuten CPM, Vanstone M, Lingard L. Understanding responses to feedback: the potential and limitations of regulatory focus theory. Med Educ. 2012 Jun;46(6):593–603.

53. Rowe AD, Wood LN. What feedback do students want ? [Internet]. Australian Educational Researcher. 2007 [cited 2015 Aug 6]. Available from: http://www.aare.edu.au/publications-database.php/5501/what-feedback-do-students-want

54. Evans C. Making Sense of Assessment Feedback in Higher Education. Rev Educ Res. 2013 Feb 22;83(1):70–120.

55. Kluger A, DeNisi A. The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. Psychol Bull. 1996;119(2):254.

56. Michels NR, Denekens J, Driessen EW, Van Gaal LF, Bossaert LL, De Winter BY. A Delphi study to construct a CanMEDS competence based inventory applicable for workplace assessment. BMC Med Educ. 2012;12(1):86.

57. Ginsburg S, Mcilroy J, Oulanova O, Eva K. Toward Authentic Clinical Evaluation : Pitfalls in the Pursuit of Competency. Acad Med. 2010;85(5):780–6.

58. Cleland JA, Knight L V, Rees CE, Tracey S, Bond CM. Is it me or is it them? Factors that influence the passing of underperforming students. Med Educ. 2008 Aug;42(8):800–9.

59. Yepes-rios M, Dudek N, Duboyce R, Curtis J, Allard RJ, Varpio L, et al. The failure to fail underperforming trainees in health professions education: A BEME systematic review: BEME Guide No. 42. Med Teach. 2016;38(11):1092–9.

60. Biggs J. Teaching for quality learning at univeristy. Vol. Second. Open University Press; 2003.