Stakes in the eye of the beholder: an international study of learners' perceptions within programmatic assessment

Suzanne Schut,[1,2] Erik Driessen,[1,2] Jan van Tartwijk,[3] Cees van der Vleuten[1,2] & Sylvia Heeneman[1,4]

OBJECTIVES Within programmatic assessment, the ambition is to simultaneously optimise the feedback and the decision-making function of assessment. In this approach, individual assessments are intended to be low stakes. In practice, however, learners often perceive assessments designed to be low stakes as high stakes. In this study, we explored how learners perceive assessment stakes within programmatic assessment and which factors influence these perceptions.

METHODS Twenty-six learners were interviewed from three different countries and five different programmes, ranging from undergraduate to postgraduate medical education. The interviews explored learners' experience with and perception of assessment stakes. An open and qualitative approach to data gathering and analyses inspired by the constructivist grounded theory approach was used to analyse the data and reveal underlying mechanisms influencing learners' perceptions.

RESULTS Learners' sense of control emerged from the analysis as key for understanding learners' perception of assessment stakes. Several design factors of the assessment programme provided or hindered learners' opportunities to exercise control over the assessment experience, mainly the opportunities to influence assessment outcomes, to collect evidence and to improve. Teacher–learner relationships that were characterised by learners' autonomy and in which learners feel safe were important for learners' believed ability to exercise control and to use assessment to support their learning.

CONCLUSIONS Knowledge of the factors that influence the perception of assessment stakes can help design effective assessment programmes in which assessment supports learning. Learners' opportunities for agency, a supportive programme structure and the role of the teacher are particularly powerful mechanisms to stimulate the learning value of programmatic assessment.

[1]Faculty of Health, Medicine and Life Sciences, School of Health Professions Education, Maastricht University, Maastricht, the Netherlands
[2]Department of Educational Development and Research, Faculty of Health, Medicine and Life Sciences, Maastricht University, Maastricht, the Netherlands
[3]Department of Education, Utrecht University, Utrecht, the Netherlands

[4]Department of Pathology, Faculty of Health, Medicine and Life Sciences, Maastricht University, Maastricht, the Netherlands

Correspondence: Suzanne Schut, Faculty of Health, Medicine and Life Sciences, Educational Development and Research, Maastricht University, PO Box 616, 6200 MD Maastricht, the Netherlands. Tel: +31 (0)43 388 5726; E-mail: s.schut@maastrichtuniversity.nl

INTRODUCTION

Programmatic assessment as a new approach to assessment is emerging rapidly within medical education.[1–3] This approach is used in various medical school programmes around the world, ranging from undergraduate to postgraduate.[4–6] Programmatic assessment can be used as a framework when designing assessment programmes that are aimed at optimising both the learning and the decision-making function of assessment.[7] A growing body of evidence to support the value of programmatic assessment is emerging, and although research shows the first positive results that this assessment approach might be beneficial for supporting the development of self-regulated learning,[5,8,9] implementing this approach is a challenge and many of the principles are still uncertain in practice.[4,5,10] There is an urgent need for empirical verification of the principles and concepts underlying the theoretical model of programmatic assessment.

One of the important concepts within programmatic assessment is that assessment is proposed as a continuum with a proportional relationship between what is at stake and the number of individual assessments.[7] Each individual assessment itself has limited consequences for the learner (i.e. is low stakes) but the consequences of the evaluation of the aggregated assessments can be substantial when they are used for a decision about, for instance, graduation or promotion (i.e. high stakes). Lowering the stakes of the individual assessment is supposed to optimise and benefit the learning potential of programmatic assessment, and provide learners with a continuous flow of information about their performance.[11] However, researchers have reported a mismatch between the designers' intentions to develop low-stakes assessments to stimulate and optimise learning, and learners' perceptions of these assessments as high stakes and summative.[4,5] This potentially leads learners to focus on each individual assessment as a hurdle and not as a learning opportunity.[12,13] Furthermore, it raises the question of whether the meaning of assessment stakes as defined in the theoretical model of programmatic assessment (the consequences following an assessment) is the same for learners whose performance is being assessed.

The impact of any assessment system on learning is mediated by learner's perceptions.[13,14] Insight into these perceptions is therefore crucial for understanding what low and high stakes mean to the learner, and how and why assessment enables, or fails to enable, learners to optimise and self-regulate their learning using these assessments. Therefore, the current study aims to gain more insight into how assessment stakes are perceived by learners and which factors influence learners' perceptions.

METHODS

Sample

We purposively selected different assessment programmes and interviewed learners from multiple institutes, countries and educational phases about their assessment experiences within programmatic assessment. We used an open and qualitative approach to data gathering and analyses, inspired by constructivist grounded theory.[15,16] The inclusion criteria were: (i) a programmatic approach to assessment is used, including low-stakes assessments aiming to provide learners with information about their progress, and high-stakes decisions regarding learners' progress are based on the evaluation of the aggregation of multiple low-stakes assessments; and (ii) there is a stable implementation of programmatic assessment over a longer period to minimise interference in the perceptions of assessment stakes due to suboptimal implementation issues. It was expected that learners' experiences with and views on assessment and the stakes involved could vary with their level of training and the process of enculturation into different learning communities. Therefore, programmatic assessment practices from pre-clinical and clinical phases were purposively selected: from pre-clinical undergraduate education (setting A); from clinical undergraduate education (setting B); and from clinical postgraduate medical education (setting C). In all selected programmes there is diversity in assessment formats including a portfolio. In these portfolios, learners collect, combine and reflect on the assessment information with the aim of self-regulating their learning, supported by a mentor. The structure and characteristics of the different assessment programmes are presented in Table 1. Within these different programmes, participants were selected based on their assessment experience: all participants have experienced at least one full feedback loop (i.e. the process of multiple low-stakes assessments and at least one high-stakes decision).

Table 1 Summary of characteristics of the selected assessment programmes

| | Pre-clinical setting (*n* = 11) | | Clinical setting (*n* = 15) | | |
	A1	A2	B1	C1	C2
Institute	Cleveland Clinical Lerner College of Medicine, Cleveland, Ohio, USA	Faculty of Health, Medicine and Life Science, Maastricht University, Maastricht, the Netherlands	Faculty of Health, Medicine and Life Science, Maastricht University, Maastricht, the Netherlands	Dalhousie University Department of Family Medicine, Halifax, Nova Scotia, Canada	Maastricht University Medical Centre, Maastricht, the Netherlands
Programme	5-year graduate-entry programme, physician investigator	4-year graduate-entry Masters programme, physician-clinical investigator	6-year Bachelor-Masters programme, medicine	2-year family medicine residency programme	3-year family medicine residency programme
Course/ phase	Years 1 and 2	Year 2	Year 2 of the Masters phase, the 12 weeks clinical rotation for family medicine (last of five clinical rotations in the Masters phase)	Year 2 of the residency programme	End of year 1 or year 3 of the residency programme
Interviews*	17 (M), 18 (M), 19 (M), 20 (F), 21 (F), 22 (M)	1 (F), 2 (F), 3 (M), 4 (F), 9 (F)	5 (F), 6 (F), 7 (F), 8 (F)	11 (F), 12 (F), 13 (F), 14 (M), 15 (F), 16 (F)	10 (F), 23 (M), 24 (M), 25 (F), 26 (F)
Year group	32	50	330	15	17
Low-stakes assessments	Weekly SAQs and CAPPs, PBL (peer) evaluations, direct observations, OSCEs, Journal Club, periodic reviews	Knowledge (in- and end-of-block) tests, progress tests, OSCEs, direct observations, scholarly projects, variety of assignments	Formative knowledge test, progress tests, case-based discussions, workplace-based performance evaluation forms (mini-CEXs, field-notes), variety of assignments	Evaluation objectives, field notes, reflective discussions, narrative, OSCEs, presentations, scholarly project, ITAR, periodic reviews	Formative evaluations (knowledge and skills), national knowledge progress test in family medicine, scholarly project, presentation, self-evaluation, video assessments

SAQs = self-assessment questions; CAPPs = open book concept appraisals; PBL = problem-based learning tutorials; OSCE = objective structured clinical examinations; mini-CEX = mini clinical evaluation exercise; ITAR = narrative in-training assessment reports.
* M, male; F, female. The number represents the order in which the interviews were conducted.

Data collection

E-mails inviting learners to participate in one-to-one interviews were sent by local faculty members to all selected participants. A convenience sampling approach was taken based on learners' availability at predetermined times. A total of 26 respondents participated in individual, semi-structured

interviews. Open-ended questions were posed by one interviewer (SS), who asked participants to describe their assessment experiences, including if and why they considered the assessment meaningful for their learning and what they perceived the stakes to be, and to reflect on the consequences that followed based on their performance. When the participant did not mention high-stakes assessments, the interviewer asked him or her to reflect on an assessment with consequences designed to be high stakes, for example a certification examination or progress decisions based on a portfolio, and if or how this was different, in order to fully understand the participant's assessment experience. All sessions were recorded and transcribed verbatim. Interviews and analyses were conducted iteratively, allowing early insights, conceptual ideas and unexpected findings to shape subsequent data collection.[15] Data were collected between April 2016 and November 2016. Participants received a small compensation (a $10 gift card). Ethical approval was obtained from the Dutch Association for Medical Education Ethical Review Board (NVMO-ERB668 on 1 March 2016), the Dalhousie Health Sciences Research Ethics Board (REB#2016-3882 on 25 July 2016) and the Cleveland Clinic Institutional Review Board (IRB#16-1261 on 21 September 2016).

Data analysis

Interview data were analysed using a constant comparative approach.[15] Independent analysis of the first four transcripts using an open coding strategy was carried out by SS and SH. During this process, coding results and relations between codes were discussed constantly. Differences were discussed until consensus was reached. This process resulted in initial codes and preliminary themes, which were used by the first author (SS) for coding of the next four transcripts. When new codes and themes emerged, these transcripts too were independently analysed by the second researcher (SH) to test the fit and relevance of the new codes and themes. Necessary adaptations to the interview questions were made for the subsequent interviews. Through coding and constant comparison, data were organised around two main categories: programme factors and (inter) personal factors. Several discussions with all members of the research team were organised in order to reach consensus on the themes that emerged, on the depth of the preliminary analysis and on the relationships between codes and categories in order to raise the analytical level from categorical to conceptional.

Furthermore, two members of the research team (ED and JvT) read two additional transcripts to review the data and to ensure a fit with the codes and discussed themes. Data collection and analysis continued until theoretical sufficiency was reached, defined as 'the stage at which categories seem to cope adequately with new data without requiring continued extensions and modifications'.[17] Theoretical sufficiency was proposed by Dey[17] and offers a more nuanced alternative to saturation to deal with issues concerning the sense of completeness and certainty implied by theoretical saturation.[18] The following criteria were used: (i) new data could be fitted in categories that were already developed; (ii) no new insights, themes, issues or counter-examples or cases arose; and (iii) consensus within the research team was reached about the notion of sufficiency with the collected and analysed data.[15–17] All interviews were then re-read by the first researcher to ensure that no relevant information was missed.

Reflexivity

We acknowledge that data in this study are co-constructed by interactions with the participants, as are the interpretations and meaning we gave to these data.[15] To prevent biases as much as possible, we brought together a multidisciplinary research team: SS and ED have a background in educational sciences, CvdV in psychology, JvT in sociology and SH in biomedical sciences. SS, ED, CvdV and SH all have involvement in programmatic assessment in medical education. To avoid tunnel vision in our interpretation of the data, we brought in an outsider perspective: JvT works in the social sciences and in teacher education and is not directly involved in medical education.

RESULTS

Overall, learners shared the same definition of stakes as defined within the model of programmatic assessment; that is, the consequences following an assessment. However, these consequences were not primarily considered as the proposed continuum in the programmatic assessment model, but rather as a dichotomy. Assessment comes with stakes (i.e. with consequences) or no stakes at all (i.e. no consequences); 'It doesn't count, nobody cares, it's not like you have to remediate or take a resit or whatever' (A2). Assessment as a continuum of stakes was recognised but appeared much more complex, encompassing more than just consequences

following an assessment. In all different programmes, learners' conceptualisation of assessment stakes as a continuum was strongly related to their perceived ability to act, control and make choices within the learning and assessment environment. Several design factors of the assessment programme influenced learners' opportunities to exercise control. Whether or not learners acted upon these opportunities depended on the interplay between experience and confidence, as well as the relationship with others in the assessment environment such as teachers. The results are presented as (i) the opportunities for learners' control within the assessment programme and (ii) factors influencing learners' believed ability to exercise control.

Opportunities for control within the assessment programme

Several programme design factors influenced learners' opportunities to exercise control and with that the perception of stakes. These aspects are described below, with participants' quotes to illustrate the themes.

Opportunities to influence outcomes

What was being assessed and which format was used for the assessment influenced the perceived stakes. In the case of making progress in generic competencies (e.g. communication, collaboration and professionalism), learners experienced multiple perspectives on the requirements of these competencies, often without a clear standard or norm, resulting in a perception of more influence on the required outcomes. By contrast, most learners considered standardised knowledge tests as high stakes and associated these assessment tasks with a fixed norm to be achieved. They experienced success in such assessments as being able to 'find the correct answer', according to a pre-constructed test and answer key, which led to a feeling of being highly dependent on the content, quality and relevance of the specific test. This caused a perception of little to no control over the assessment and outcomes, especially when this type of assessment resulted in grades:

> You either know it or you don't when it comes to knowledge. Whereas I guess when you're talking to your preceptor about a field note, it's less measurable outcomes. So it's more so about reflecting and just talking through something, it's more fluid than grades. (C1)

Furthermore, the opportunity to interact with the assessor (e.g. during an oral examination, or when an assessor would interact with the learner during direct observation) was perceived as a potential influence on the assessment outcome. Learners indicated that interaction with the assessor provided more opportunities to show their progress and abilities, and made them feel more in control over the process and the outcome of the assessment. This lowered the perceived stakes. However, interaction with the assessor could also raise the perceived stakes: learners thought this carried the risk of losing face, especially when the assessor was intimidating, an important role model or worked in a discipline of interest.

Opportunities to collect evidence

In all programmes, learners collected evidence within a portfolio, with the aim of monitoring and showing their progress. However, programmes varied in the freedom learners had to collect and select evidence. Some programmes gave learners the opportunity to initiate an assessment, for example by allowing learners to assess their knowledge development by taking formative self-tests or encouraging learners to ask for direct observation on the learners' own terms. This feeling of control not only lowered the perceived stakes, but more importantly also seemed to make the assessment feel more relevant:

> You have more control over the assessment [initiating a Mini-CEX] and then you can focus the assessment to what is important to you. You can tailor it to what you need at that moment. That makes it low-stake and more meaningful. (C2)

The perceived stakes were lower when results or follow-ups were not automatically accessible to others and learners could control what was shared. Learners experienced more choice and felt more in control when given the opportunity to select their own evidence for the portfolio:

> To me those [own evidence] would be the lowest stakes, because you are not expected to collect them, they are not expected to be in there [the portfolio]. (A1)

Opportunities to improve

The procedures offered by programmes to improve earlier insufficient performance influenced learners'

perceived control of the impact of each individual assessment. An important factor was whether or not opportunities for improvement were integrated into the educational programme. When this had to be done next to the regular curriculum or assessment activities, the time investment needed for improvement felt like an overload, the stakes became higher and learners were more motivated to avoid this:

> It [curriculum] is already overloaded [...] For me that is also the incentive to just want to pass and get it over with, I think otherwise the other things will be in jeopardy. (A2)

Most programmes provided multiple complementary assessments that were meant to facilitate more opportunities for learners to show progress and improvement. This lowered the stakes of the individual assessment. Not being solely dependent on one individual 'snapshot' gave learners more feeling of being in control, because of multiple opportunities to show and improve on their performance, especially when the focus was on trends or reoccurring feedback messages:

> I have lots of different people talking about my professionalism and so, each additional one has less impact. It's okay to get some negative feedback, because you have a lot. So you have some negative and some positive, just, you know that's how it tends to balance out. (B1)

Although the number of complementary assessments influenced the perception of assessment stakes positively, this also came with reaching a point of so-called 'overkill', in which the assessment became meaningless and a checkbox activity to meet the requirements of the programme:

> It almost becomes a hunt on evaluations. And it's not about the quality or their usefulness anymore, but just about the quantity. (B1)

Nonetheless, learners did not always recognise the coherence or complementary nature of multiple assessments, causing them to perceive the assessment as isolated and therefore high stakes. When grades were used for individual assessments, and learners could correct insufficient performance by ways of averaging multiple results, this contributed to learners' understanding of the coherence. However, receiving grades also contributed to competition amongst learners, anxiety and a performance orientation, which raised

the stakes. The following is an example of a learner reflecting on his transition from a previous assessment environment with grades to his current environment without grades:

> In undergraduate studies a test would make me nervous and anxious and worried about how well I was going to do. But because these assessments don't have the same consequences [receiving grades], because they are just to help me identify what to study, I don't feel the same nervousness that I did before. It's the good without the bad. (A1)

Factors influencing learners' believed ability to exercise control

Learners used opportunities provided by the assessment programme for control, when they believed they had the ability to exercise control. This belief was influenced by personal attributes, as well as the relationship experienced with teachers. In effect, the factors influencing learners' believed ability to exercise control are presented as (i) the interplay between learners' experience and confidence and (ii) the influence of teachers.

The interplay between experience and confidence

Previous experience of assessments influenced the perceived stakes within all programmes. Most learners were accustomed to defining success as being top of their class and getting high scores or grades:

> Assessments, they were always a bit stressful for me. [...] There was always something connected to them, like proceeding to the next year for example. And also, like with my parents and grandmothers, if I got a good grade, then there is a certain reward. And I think that is also like a bit of conditioning. (C2)

In programmes previously attended, this was often rewarded and even viewed as a necessity. An example is the situation when admission to a medical school required a high secondary school grade point average. Assessment was then associated with pressure for high performance, insecurities and fear of failing. Such assessment experiences had a strong impact. New experiences were required before these associations were replaced with a more learning-oriented perception of assessment. Learners had to gain confidence in the meaning and consequences of the low-stakes assessment,

which contributed significantly to the perception of stakes. First-time experiences with low-stakes assessments were unanimously perceived as high stakes, especially when learners did not fully understand what was expected, or what could happen when they were unable to meet the demands: 'I think a lot of the anxiety was caused by us not knowing exactly what was going to happen [if we would perform poorly on an OSCE]' (A1). The more familiar learners became with such assessments, the less anxious they felt.

The influence of teachers

The believed ability to exercise control and therefore the perception of assessment as low stakes seemed strongly dependent on learners' relationship with their teachers. When learners felt the teacher was their advocate, facilitated learning, and allowed them to experiment and to take control, they felt safe and able to interpret low-stakes assessments as low stakes and meaningful for learning. The assessment environment was then described as a safe place to learn and experiment: 'I feel very comfortable looking stupid' (A1) and 'I think it's hard to feel the fear of failing' (C1). However, some learners felt that the relationship was characterised by an unequal power balance that influenced their perception of assessment stakes:

> So he [the teacher] has all the power. That's how it feels to me. He has a lot to say about it. The things he considers important, he picks them out and focuses on them. And the consequence they have, I think are much bigger than the consequences such a test should actually have. (B1)

For learners to take control, the teacher needed to provide them with the opportunity to exercise control.

DISCUSSION

The theoretical assumption underlying the proposed continuum of assessment stakes within programmatic assessment, is that low-stakes assessments create learning opportunities and generate a continuous flow of information for learners that can be used to self-regulate their learning.[7,11] This requires assessment that is intended or designed to be low stakes, to be perceived as such by the learner.[12–14] This study identified the feeling of being in control to be

essential for learners' perception of assessment stakes, and identified factors that allowed or hindered learners' opportunity to exercise control. This is strongly linked to the concept of agency, referred to as learners' perceived ability to act, control and make choices within the learning and assessment environment.[12,13,19] The value and importance of learners' agency for continuous development using assessment has already been highlighted by others[1,8,12,19–21] and affects the ability or willingness to learn from assessment.[12,13,22] What this study contributes is insight into how learners' agency is negotiated in the context of programmatic assessment and what enables and constrains its emergence.

Different programme factors provided or hindered learners' opportunities to take control over the assessment experience. Standardised assessments provide little opportunity for learners' agency. Although necessary and understandable, standardisation places the control at the programme level, leaving little space for the individual learner to exercise control. This might even alienate learners from their learning and assessment experience.[23,24] A sense of agency was, however, encouraged when the programme allowed learners to initiate their own assessment, and when learners were enabled to select evidence of progress. This has the potential to engage learners more actively in the assessment process.[12,19,25]

Increasing the number of opportunities for learners to monitor and show progress, even with standardised knowledge tests, can be another strategy to lower the stakes. Programme designers should take care, however, not to create an assessment overload for both learners and faculty members. In addition, the link between individual low-stakes assessments must be clear for the learner in order to provide direction on how low-stakes assessment can and should be used to support learning.[26] Although using grades and opportunities to compensate can highlight the coherence amongst complementary assessments, this may have the adverse effect of encouraging different, less desirable, study strategies and behaviours.[27] Providing grades has the implicit risk of encouraging a focus on outcomes and competition rather than stimulating a focus on continuous improvement.[12,28] Not providing grades seemed to enable a shift to a learning orientation, described as one in which the learner's goal is to improve.[29,30] Linking complementary low-stakes assessments was better accomplished in

programmes that highlighted the influence of low-stakes assessments on learners' improvement plans. The focus should be on using information generated by low-stakes assessment to analyse and reflect upon learning progress and how this should direct future learning. We could consider giving more opportunity for the learner to control the appropriate number of assessments needed to show progress and improvement, rather than setting up quantity requirements. This can make the assessment experience a more personal inquiry and create ownership over the plan of improvement. Thinking this way does not take away the need for some type of consequences: when learners perceived little incentive to address their weaknesses or to act upon information concerning their strengths and weaknesses, low-stakes assessment rarely led to a focus on improvement. The implementation of the so-called post-assessment process (the follow-up activities or reflective tasks) is therefore essential,[31] both in the design of a supportive programme structure (i.e. facilitating room for improvement) and in the role of the teacher (i.e. valuing and stimulating improvement versus performance).

Last, learners' confidence and their believed ability to exercise control seemed to increase over time. Novices within programmatic assessment needed time to adjust to and get familiar with the new assessment approach. Associations and experiences with high-stakes assessment need to be phased out and teachers need to adjust the level of guidance and direction to the experience level of learners.[32] Moreover, when the teacher–learner relationship can be characterised as safe and with autonomy for the learner, learners are more likely to use assessment to support their learning. Within programmatic assessment, learners should be allowed independence and control over the assessment process. Only then, will learners perceive assessment as low stakes.

Our results fit well with the calls to create a shared responsibility between learners and teachers within the assessment process[1,33] and with the need to create a learning environment where dialogue can flourish to engage learners actively with feedback and assessment.[21,34] This challenges teachers to reconcile their responsibility for stimulating and evaluating development.[1,12,20] Teachers are fundamental for creating this safe learning environment and utilising the potential of programmatic assessment.

Limitations

This study has several limitations. Firstly, learners within medical education are typically characterised as high achievers, selected through rigorous course admission procedures. Many learners in this study referred to themselves as such and only a few of them had experienced failure in relation to assessment or progress decisions. The perception of assessment stakes might work differently for low-achieving learners. Future work could explore the relevance of our concepts and mechanisms in other (academic) settings by including programmes outside medical education or purposefully selecting low-achieving learners.

Secondly, participants received a small compensation for their time and effort. All ethical committees involved approved this. Without exceptions, participants seemed motivated to further our understanding of their assessment experience and to contribute to the quality of the educational programme, but of course we can't exclude that the small reward may have biased our sample.

Thirdly, the number of interviews per programme was limited. However, this study was not designed to compare different implementations of programmatic assessment. We included a range from undergraduate to postgraduate medical education to understand the underlying mechanisms that appeared in different programmes and to identify influencing design factors. Although the decision to sample different phases of medical education, and staying within one discipline, was made on methodological grounds, this could have influenced the representation of certain themes. Differences between programmes, class sizes and disciplines may also impact the implications of our findings and recommendations for other practices. We therefore stress the importance of replicating our study in different contexts and learning cultures for further understanding and transferability.

Finally, given the importance of the role of the teacher, future studies should triangulate students' self-reported perceptions by exploring the perceptions of teachers regarding the stakes of assessment.

CONCLUSION

This study identified factors that influenced the stakes learners perceived within a programmatic

approach to assessment. Learners perceive assessments more as low stakes when we provide opportunities for learners to exercise control and create an assessment environment that embraces possibilities for improvement. Environments wherein learners can freely discuss their weaknesses and uncertainties, and teachers encourage learners to share learning needs, invest in their improvement plans, and provide guidance and direction needed for achieving progress and improvement. In summary, learners' opportunities for agency, a supportive structure and the role of the teacher are particularly powerful factors in stimulating the perception of assessment as low stakes and enhancing the learning potential of assessment. Knowledge and understanding of the identified factors can help educational developers to design effective programmes of assessment that could increase the learning value of assessment.

REFERENCES

1 Eva KW, Bordage G, Campbell C, Galbraith R, Ginsburg S, Holmboe E, Regehr G. Towards a program of assessment for health professionals: from training into practice. *Adv Health Sci Educ Theory Pract* 2016;**21** (4):897–913.

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

3 Schuwirth L, van der Vleuten C, Durning SJ. What programmatic assessment in medical education can learn from healthcare. *Perspect Med Educ* 2017;**6** (4):211–5.

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

5 Heeneman S, Oudkerk Pool A, Schuwirth LWT, van der Vleuten CPM, Driessen EW. The impact of programmatic assessment on student learning: theory versus practice. *Med Educ* 2015;**49** (5):487–98.

6 Dannefer EF, Henson LC. The portfolio approach to competency-based assessment at the Cleveland Clinic Lerner College of Medicine. *Acad Med* 2007;**82** (5):493–502.

7 Van der Vleuten CP, Schuwirth LW, Driessen EW, Dijkstra J, Tigelaar D, Baartman LK, van Tartwijk J. A model for programmatic assessment fit for purpose. *Med Teach* 2012;**34** (3):205–14.

8 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** (3):221–5.

9 Sundre DL, Kitsantas A. An exploration of the psychology of the examinee: can examinee self-regulation and test-taking motivation predict consequential and non-consequential test performance? *Contemp Educ Psychol* 2004;**29** (1):6–26.

10 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;**47** (7):734–44.

11 Schuwirth LW, Van der Vleuten CP. Programmatic assessment: from assessment of learning to assessment for learning. *Med Teach* 2011;**33** (6):478–85.

12 Harrison C, Könings KD, Dannefer EF, Schuwirth LWT, Wass V, van der Vleuten CPM. Factors influencing students' receptivity to formative feedback emerging from different assessment cultures. *Perspect Med Educ* 2016;**5** (5):276–84.

13 Cilliers FJ, Schuwirth LW, Adendorff HJ, Herman N, van der Vleuten CP. The mechanism of impact of summative assessment on medical students' learning. *Adv Health Sci Educ Theory Pract* 2010;**15** (5):695–715.

14 Segers M, Nijhuis J, Gijselaers W. Redesigning a learning and assessment environment: the influence on students' perceptions of assessment demands and their learning strategies. *Stud Educ Eval* 2006;**32** (3):223–42.

15 Watling C, Lingard L. Grounded theory in medical education research: AMEE Guide No. 70. *Med Teach* 2012;**34** (10):850–61.

16 Corbin J, Strauss A. *Basics of Qualitative Research (3rd ed.): Techniques and Procedures for Developing Grounded Theory.* Thousand Oaks, CA: Sage publications, Inc. 2008. http://methods.sagepub.com/book/basics-of-qualitative-research.

17 Dey I. *Grounding Grounded Theory: Guidelines for Qualitative Inquiry.* San Diego, CA: Academic Press 1999.

18 Morse JM. The significance of saturation. *Qual Health Res* 2016;**5** (2):147–9.

19 Cilliers FJ, Schuwirth LW, van der Vleuten CP. A model of the pre-assessment learning effects of

assessment is operational in an undergraduate clinical context. *BMC Med Educ* 2012;**12**:9.

20 Watling C. The uneasy alliance of assessment and feedback. *Perspect Med Educ* 2016;**5** (5):262–4.

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

22 Wiliam D. *Embedded Formative Assessment.* Bloomington, IN: Solution Tree Press 2011.

23 Mann SJ. Alternative perspectives on the student experience: alienation and engagement. *Stud High Educ* 2001;**26** (1):7–19.

24 Kahu ER. Framing student engagement in higher education. *Stud High Educ* 2013;**38** (5):758–73.

25 Shepard LA. The role of assessment in a learning culture. *Educ Res* 2000;**29** (7):4.

26 Olupeliyawa A, Balasooriya C. The impact of programmatic assessment on student learning: what can the students tell us? *Med Educ* 2015;**49** (5): 453–6.

27 Harrison C, Konings KD, Schuwirth L, Wass V, van der Vleuten C. Barriers to the uptake and use of feedback in the context of summative assessment. *Adv Health Sci Educ Theory Pract* 2015;**20** (1):229–45.

28 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** (3):307–20.

29 Konopasek L, Norcini J, Krupat E. Focusing on the formative: building an assessment system aimed at student growth and development. *Acad Med* 2016;**91** (11):1492–7.

30 Dweck CS. Motivational processes affecting learning. *Am Psychol* 1986;**41** (10):1040.

31 Eva KW, Munoz J, Hanson MD, Walsh A, Wakefield J. Which factors, personal or external, most influence students' generation of learning goals? *Acad Med* 2010;**85** (10 Suppl):S102–5.

32 Wood D, Bruner JS, Ross G. The role of tutoring in problem solving. *J Child Psychol Psychiatry* 1976;**17** (2):89–100.

33 Ramani S, Post SE, Konings K, Mann K, Katz JT, van der Vleuten C. "It's Just Not the Culture": a Qualitative Study Exploring Residents' perceptions of the impact of institutional culture on feedback. *Teach Learn Med* 2017;**29** (2):153–61.

34 Bloxham S, Campbell L. Generating dialogue in assessment feedback: exploring the use of interactive cover sheets. *Assess Eval High Educ* 2010;**35** (3):291–300.