Perceptions of purpose, value, and process of the mini-Clinical Evaluation Exercise in anesthesia training

Perceptions du but, de la valeur et du processus du *mini-Exercice d'évaluation clinique* dans la formation en anesthésie

Damian J. Castanelli, MBBS · Tanisha Jowsey, PhD · Yan Chen, PhD · Jennifer M. Weller, MD

Received: 16 May 2016/Revised: 15 August 2016/Accepted: 13 September 2016/Published online: 23 September 2016
© Canadian Anesthesiologists' Society 2016

Abstract

Introduction Workplace-based assessment is integral to programmatic assessment in a competency-based curriculum. In 2013, one such assessment, a mini-Clinical Evaluation Exercise (mini-CEX) with a novel “entrustability scale”, became compulsory for over 1,200 Australia and New Zealand College of Anaesthetists (ANZCA) trainees. We explored trainees’ and supervisors’ understanding of the mini-CEX, their experience with the assessment, and their perceptions of its influence on learning and supervision.

Methods We conducted semi-structured telephone interviews with anesthesia supervisors and trainees and performed an inductive thematic analysis of the verbatim transcripts.

Electronic supplementary material The online version of this article (doi:10.1007/s12630-016-0740-9) contains supplementary material, which is available to authorized users.

D. J. Castanelli, MBBS (✉)
Department of Anaesthesia and Perioperative Medicine, Monash Medical Centre, 246 Clayton Road Clayton, Melbourne, VIC 3168, Australia
e-mail: Damian.Castanelli@monash.edu

D. J. Castanelli, MBBS
Department of Anaesthesia and Perioperative Medicine, Monash University, Melbourne, VIC, Australia

T. Jowsey, PhD · Y. Chen, PhD · J. M. Weller, MD
Centre for Medical and Health Sciences Education, Faculty of Medical and Health Sciences, University of Auckland, Auckland, New Zealand

J. M. Weller, MD
Department of Anaesthesia, Auckland City Hospital, Auckland, New Zealand

Results Eighteen supervisors and 17 trainees participated ($n = 35$). Interrelated themes concerned the perceived purpose of the mini-CEX, its value in trainee learning and supervision, and the process of performing the assessment. While few participants saw the mini-CEX primarily as an administrative burden, most focused on its potential for facilitating trainee improvement and reported positive impacts on the quantity and quality of feedback, trainee learning, and supervision. Finding time to schedule assessments and deliver timely feedback proved to be difficult in busy clinical workplaces. Views on case selection were divided and driven by contrasting goals – i.e., receiving useful feedback on challenging cases or receiving a high score by choosing lenient assessors or easy cases. Whether individual mini-CEXs were summative or formative was subject to intense debate, while the intended summative use of multiple mini-CEXs in programmatic assessment was poorly understood.

Conclusion Greater clarity of purpose and consistency of time commitment are necessary to embed the mini-CEX in the culture of the workplace, to realize the full potential for trainee learning, and to reach decisions on trainee progression.

Résumé

Introduction L'évaluation fondée sur le lieu du travail fait partie intégrante de l'évaluation programmatique dans le cadre d'un programme de formation fondé sur la compétence. En 2013 une telle évaluation, le mini-CEX (mini-exercice d'évaluation clinique), qui comportait une nouvelle « échelle mesurant la capacité à se voir confier un geste » (ou ‘entrustability scale’), a été rendue obligatoire pour plus de 1200 stagiaires du Collège des anesthésistes d'Australie et de Nouvelle-Zélande (ANZCA). Nous avons

cherché à savoir comment les stagiaires et les superviseurs comprenaient le mini-CEX, quelles étaient leurs expériences avec cette évaluation, et quelles étaient leurs perceptions de son influence sur l'apprentissage et la supervision.

Méthode Nous avons réalisé des entretiens téléphoniques semi-structurés avec des superviseurs et des stagiaires en anesthésie avant de procéder à une analyse thématique inductive des transcriptions verbales de ces entretiens.

Résultats Dix-huit superviseurs et 17 stagiaires ont participé à notre évaluation ($n = 35$). Les thèmes interconnectés comprenaient le but perçu du mini-CEX, sa valeur dans l'apprentissage et la supervision des stagiaires, et le processus de réalisation de l'évaluation. Alors que les participants furent peu nombreux à percevoir le mini-CEX principalement comme un fardeau administratif, la plupart ont mis l'emphasis sur le potentiel de cet outil à aider les stagiaires à progresser plus facilement et ont rapporté des impacts positifs sur la quantité et la qualité des rétroactions, sur l'apprentissage des stagiaires, et sur la supervision. Il s'est avéré difficile de trouver le temps de planifier des évaluations et d'offrir des rétroactions opportunes dans les milieux de travail clinique très occupés. Les opinions sur la sélection des cas étaient divisées et motivées par des objectifs différents – soit le fait de recevoir des rétroactions utiles à propos de cas difficiles ou de recevoir une bonne note en choisissant des évaluateurs indulgents ou des cas faciles. Les opinions quant à l'aspect sommatif ou formatif des mini-CEX individuels étaient également très partagées, alors que l'utilisation sommative projetée de plusieurs mini-CEX pour l'évaluation programmatique a été mal comprise.

Conclusion Une plus grande clarté de l'objectif et un investissement de temps plus constant sont nécessaires pour intégrer le mini-CEX dans la culture du lieu de travail, pour réaliser son plein potentiel pour l'apprentissage des stagiaires, et pour parvenir à des décisions quant aux progrès des stagiaires.

Workplace-based assessments (WBAs) have become ubiquitous with the move to competency-based postgraduate medical education.¹ Workplace-based assessments were initially adopted to increase the authenticity of assessment by shifting its focus from testing what trainees know to testing what trainees do.² Subsequently, their use has evolved into a decision-making tool for determining trainees' preparedness for progression or certification.³⁻⁵

In parallel, the emphasis on observation, assessment, and feedback to facilitate learning^{3,4} has led to a

conceptual shift from “assessment of learning” to “assessment for learning”.⁶ The main purpose of assessment of learning is to “determine if the trainee met summative (‘pass/fail’) performance standards or successfully completed a course or study”.⁷ In contrast, with assessment for learning, “the assessment process is inextricably embedded within the educational process, which is maximally information-rich, and which serves to steer and foster the learning of each individual student to the maximum of his/her ability”.⁶

Assessment for learning is an extension rather than a replacement for assessment of learning; proponents of assessment for learning have acknowledged that summative decisions are still required.⁸ Programmatic assessment has been proposed to fulfill these dual purposes in a systematic manner, where individual assessments are focused on maximizing trainee learning, but aggregate information from all assessment data is compared with a performance standard to facilitate decision-making.⁹

Nevertheless, difficulties have been reported with the implementation of WBAs in a programmatic approach to assessment, with assessment of learning and compliance with administrative requirements dominating the experience of learners and supervisors.¹⁰⁻¹² These concerns have resulted in the introduction of alternatives such as the Supervised Learning Events, which, like WBAs, require observation, judgement, and feedback. Unlike WBAs, however, they are not intended to contribute to progression decisions.¹³⁻¹⁵

The mini-Clinical Evaluation Exercise (mini-CEX) has been adopted into anesthesia training programs around the world.¹⁶⁻¹⁸ In 2013, the Australian and New Zealand College of Anaesthetists (ANZCA) introduced a competency-based curriculum.¹⁹ ANZCA selected the mini-CEX as part of a suite of WBA tools, consistent with the principles of assessment for learning and programmatic assessment discussed above:

“Workplace-based assessment provides a framework to support teaching and learning in the clinical environment and promotes a holistic view of a trainee's clinical practice. Trainees have the opportunity to assess their own learning and use feedback from these assessments to inform and develop their own practice. While the goal of workplace-based assessment is to aid trainee learning, they can be used to create a record to demonstrate development and inform the regular review of trainee progression.”¹⁹

Preliminary work investigating the mini-CEX in a volunteer sample of anesthesia trainees had suggested a very positive effect on observation, feedback, and learning.²⁰ An initial trial using a traditional scale (unsatisfactory, satisfactory, superior) showed very low reliability, which improved markedly when the scale was changed to one where supervisors made judgements on the

level of supervision required.^{21,22} The ANZCA scale for the mini-CEX asks supervisors to score the trainee's ability to manage the case independently,²² i.e., based on the need for direct or more distant supervision. This independence scale is an example of what has been referred to as an "entrustability scale".²³ We wanted to know if the benefits of improved observation and feedback evident in our volunteer pilot studies^{20,22} would be retained in the real world of the ANZCA training program.

We aimed to explore how ANZCA trainees and supervisors of training considered their experience with the mini-CEX 18 months after their compulsory introduction into anesthesia training.

Methods

This study was approved by the University of Auckland Human Participants Ethics Committee (reference number 011108) and Monash University Human Research Ethics Committee (reference number CF14/1668 – 20140007960).

Given that our aim was to explore the experience of participants using the mini-CEX, our study was designed following an interpretivist approach which seeks to understand and explain people's lived experiences.²⁴

Context

The ANZCA is responsible for the training and assessment of all anesthesia trainees in Australia and New Zealand. Prior to implementation, the ANZCA instituted a faculty development program to prepare anesthetists for their role as WBA assessors. This program used a "snowballing" technique to train a large number of participants effectively within a brief period of time across a geographically dispersed training program. Supporting resources were developed centrally and "champions" were trained. The champions subsequently became responsible for delivery of training in their local regions. More than 800 anesthetists, or approximately 20% of the WBA assessors, received training through this program (O. Jones, ANZCA, personal communication).

The ANZCA trainees are required to submit a minimum number of WBAs during each stage of their training, with specialist anesthetists acting as supervisors. These WBAs are in turn available to their supervisors of training (SoTs), a subset of supervisors officially appointed by the ANZCA to be responsible for training in their departments. They oversee each trainee's clinical performance and WBAs, perform regular clinical placement reviews, and confirm

progression of trainees through the training program. Supervisors, trainees, and SoTs access the ANZCA mini-CEX online, but a written version is available²⁵ (Appendix 1; available as Electronic Supplementary Material).

This study took place approximately 18 months after implementation, with 1,224 ANZCA trainees having completed 7,808 mini-CEXs in the prior twelve months.

Sampling

Given the potential role of the mini-CEX in progression decisions, we selected SoTs and trainees as the most suitable subjects to explore the impact of the mini-CEX. The sampling approach involved a purposive, criterion-based, and maximum-variation sampling frame to maximize the potential for conceptual generalizability.^{26,27} This sampling technique involves recruitment of subjects as required until no new ideas emerge from the interview data.²⁶ The ANZCA Clinical Trials Network staff selected potential interviewees from confidential ANZCA records to ensure they represented trainees at different training stages and that trainees and SoTs came from diverse practice locations. Recruitment was monitored to ensure representation from rural and metropolitan hospitals of different sizes.

Interviews

After obtaining informed consent, we conducted semi-structured telephone interviews with each participant for approximately 30 min. Semi-structured interviews are "encounters between the researcher and informants directed towards understanding the informants' perspectives on their lives, experiences, or situations as expressed in their own words".²⁸ The semi-structured interviews began with open questions to facilitate participants' description of their personal experiences and to allow exploration of their individual responses. The interview guide was developed by the investigators based on literature review and themes from a previous qualitative study²⁰ (Appendix 2; available as Electronic Supplementary Material).

The interviews were recorded and professionally transcribed. An investigator (Y.C.) with experience in qualitative interviewing, but without a medical background or ANZCA affiliation, conducted all interviews. This served to maintain anonymity and safety for interviewees and to minimize the potential for the relationship between anesthetist investigators and participants to influence the responses.

Analysis

Inductive thematic analysis was undertaken following Savin-Baden and Major and Morse and Field.^{29,30} After a close reading of the data, we iteratively developed a coding scheme.³¹ Each code was assigned a description explaining when to code a quote, inclusion/exclusion criteria, and examples of appropriate quotes.³¹ Two researchers from an academic consulting firm coded all data using QSR NVivo10 qualitative software (QSR International, Victoria, Australia).³² A member of our research team (T.J.) checked this coding periodically throughout the coding process and at the completion of the coding. Two meetings were held between the research team and one of the coders during the coding process to refine the coding scheme further and ensure accurate and consistent coding.

Upon completion of the coding, we identified the frequency of coding to each code (forming the basis for content analysis) and examined the associations between codes.³¹ We looked for patterns in quotes that were coded to more than one code and synthesized key messages from the codes into descriptive findings. These interconnected findings formed the basis for generating themes, which were then refined by comparison with the data until the researchers reached agreement that the themes expressed the meaning in the data. All members of the research team were involved throughout the analysis, except where indicated above.

The research team included anesthetists with experience as SoT (D.C.), clinical teachers, curriculum and assessment designers, and medical education researchers (D.C., J.W.), as well as non-clinicians involved in medical education with backgrounds in anthropology (T.J.) and psychology (Y.C.). This researcher triangulation ensured a diversity of insider and outsider perspectives to inform this work.

Results

We interviewed 18 SoTs and 17 trainees from August–December 2014. Nine SoTs and eight trainees were female, and nine SoTs and eight trainees worked in regional hospitals. Of the trainees, eight were in basic training (6–24 mth), seven were in advanced training (24–48 mth), one was an introductory trainee (0–6 mth), and one was a provisional fellow (48–60 mth). The major themes that emerged in the analysis relate to perceptions of the purpose of the mini-CEX, its value in trainee learning and supervision, and aspects of the process of performing the mini-CEX. Themes and subthemes are provided in Table 1, with exemplar quotes highlighted in the following sections.

Table 1 Themes and sub-themes

Purpose:
Trainee improvement
Documentation/ trivialization
Inform progression
Identify underperformance
Summative or Formative
Value:
Trainee Learning
Feedback
Supervision
Time:
Timing of case selection
Timing of feedback
Scoring:
Understanding
Variability
Leniency
Case/assessor selection:
Demonstrate competence
Maximize learning
Assessor characteristics

Purpose

The majority of participating SoTs and trainees appeared engaged in the WBA process and perceived the purpose of the mini-CEX as facilitating trainee learning and improvement.

“The main purpose is to assess competency ... feedback of things which have been done well and which aren’t...and looking for areas for improvement.” (Trainee 16)

Nevertheless, a sizeable minority of participants reported that the primary purpose was seen as administrative on some occasions, and the process was often trivialized. The mini-CEX was “*just another hoop that the College has established for us all to jump through.*” (Trainee 9) A trainee explained, “*I do them because I have to.*” (Trainee 10)

While both trainees and SoTs reported that there was a risk of not progressing without completing the required number of mini-CEXs, some described how the mini-CEX could provide a gauge of trainee performance to inform formal decision-making on trainee progression.

“Now if those mini-CEXs indicate that they may not be functioning satisfactorily at that level then I can hold them back ... and I will use repeat mini- CEX

plus other tools to assess their ability, to hold them back until they do.” (SoT 11)

A number of participating trainees and SoTs saw detection of underperformers as an important intended purpose of the mini-CEX. Nevertheless, they expressed scepticism regarding its usefulness for this purpose due to issues of leniency and variability in rating and trainee selection of assessor and case, with trainees reportedly gaming the system to find lenient supervisors.

“Most trainees are going to be able to game the system in order (to) find consultants (supervisors) who give them an easy ride.” (Trainee 2)

Nevertheless, mini-CEX assessments were reported to have been useful during remediation on at least one occasion.

“Asking a trainee to do a mini-CEX on something they were struggling with... it gave the trainee a target to work for as well as me a means to fully assess it.” (SoT 11)

Many participants expressed views on whether the intended use of the mini-CEX was as a formative or a summative assessment. It was often seen as formative.

“We don’t treat it as an exam. It’s not a nasty process. In fact, it’s really encouraged to be not a confrontational process. It is a formative process, not a summative assessment. They’re aware of that.” (SoT 11)

There was acknowledgement by some participants that there were different understandings of this aspect of the mini-CEX among both trainees and supervisors.

“You know they misunderstand that it’s formative, and they see it as a bit of a test, and they get anxious.” (SoT 13)

A complementary view expressed by other SOT and trainee participants was that the mini-CEX was actually a summative assessment and that confusion arose from mislabelling it as formative.

“There’s too much mud in the water talking about them being formative which they’re not.” (SoT 2)

Value

Both SoT and trainee participants reported benefits that the mini-CEX brings to trainees’ learning. The compulsory nature of the mini-CEX, along with the core skills assessed, was said to make it easier for the trainees to identify their own weaknesses and develop learning goals.

“If you do a mini-CEX it may expose significant deficiencies in knowledge or other professional attributes, ... it is good because it’s better to expose those weaknesses than to cover them up and never to improve.” (Trainee 4)

The SoT participants valued the different domains on the mini-CEX, reporting that they facilitated the provision of more structured and in-depth feedback. They reported personal benefits, such as having a record of their teaching and feedback and being obliged to “step back and watch” (SoT 2) or “critically observe” (SoT 3). They also reported applying the structure of the mini-CEX to cases that were not being assessed.

The different aspects of the mini-CEX form were reported to prompt supervisors to broaden the scope of areas being assessed to include non-technical skills and other aspects of trainee performance previously not examined.

“We are forced into looking at those various different non-technical skills and other aspects of their performance that we maybe didn’t look at before.” (SoT 19)

Trainees confirmed this broadening of the scope of feedback and, in addition, reported actively seeking feedback from supervisors in specific practice areas.

Participants particularly valued the comments that supervisors recorded in the free-text fields.

“The number scales are useful generally. But a lot more useful is the individual information you write in the free field boxes about what they did well, what they didn’t do well.” (SoT 11)

Most participants reported that the quality of supervision was improved more generally as supervisors were more often critically observing trainees’ performance against explicit criteria.

“It is an opportunity for the consultant (supervisor) and the trainee to discuss each other’s practice standards and to make sure that what they are doing is the best that they can do.” (SoT 3)

The overall impression was that the mini-CEX fostered positive interactions between supervisors and trainees, with a resultant greater engagement in teaching and an increase in the dialogue between trainees and supervisors.

Time

In their responses regarding the use of the mini-CEX, participants emphasized three key aspects of *time*: *timing* of case selection, *making time* for feedback (including

duration and scheduling), and *timeliness* of providing feedback. These three aspects of time often overlapped in participant accounts, with a background of time pressure and lack of time.

The SoTs and trainees held different views regarding the timing of case selection. The SoT participants suggested that the mini-CEX assessments held greater potential for learning when trainees planned them in advance, and many SoTs were frustrated when trainees requested a mini-CEX without warning. Trainees reported that some supervisors were willing to retrospectively complete assessments on past cases but that most were not. The SoT participants reported refusing to do mini-CEXs on past cases, as it was difficult to provide specific feedback if they had not prioritized observing the trainee during the case.

"I think most of the time on most lists we're so time precious that whilst we're aware that the trainee is with us ... we're not intentionally watching what they're doing." (SoT 1)

Trainees reported reluctance to request a mini-CEX in a busy list; however, both SoT and trainee participants noted that, regardless of workplace constraints, curriculum requirements meant that they were obliged to make the time.

The SoT participants reported that feedback was most easily timed immediately following the period of formal observation.

"I ... pretty much do it in theatre at the time that the period of observation has been completed." (SoT 3)

Supervisors noted that, when feedback was not provided immediately, they found it difficult to recall the details of the exercise and this increased the likelihood that the mini-CEX could be trivialized. They found that the formal feedback after the mini-CEX could take "a long time" and that "fitting" it into their busy schedules was a considerable time burden and extra work.

Scoring

During the interviews with participants, a focus of discussion was how scores were awarded on the mini-CEX, particularly participants' interpretation or understanding of the scoring scale, variability in scoring, and issues of leniency. The SoTs reported that the independence scale reflected "*how close to being a consultant they are*", and hence, it was "*appropriate for all levels of experience*". (SoT 19)

"It would be quite appropriate for a more junior trainee to require considerable consultant (supervisor) input to a more complex case, which

would normally mean that the assessor would tick boxes one, two or three." (SoT 3)

Some trainee participants also expressed this view that the score reflected "*progression to independent practice*". (Trainee 5)

"You might do very well in your case, but it scored quite low because you don't have much experience, or it was a very complex case." (Trainee 7)

Nevertheless, participating SoTs understood that communicating this message to supervisors and trainees was important but difficult, as trainees were accustomed to "succeeding" and gaining high grades.

"As long as you feel comfortable in your head explaining to your registrars about why ... some of those scores might be low, but it's completely appropriate for them to score low at their particular level of training, then I think the form's designed to work at all levels. It's just whether you are prepared to use it in that way, and the registrars prepared to accept it like that. Registrars never want to be marked low under any circumstances." (SoT 19)

Some trainees agreed that there was potential for difficulty in adapting to the scoring system.

"Trainees are very used to pass/fail systems, so it actually takes a little bit of getting used to if you get a score that's not between seven and ten." (Trainee 8)

While participating SoTs and trainees generally understood how the scale should work, their experience was that supervisors' understanding was variable and that there was an ongoing confusion about how they should be assessing trainees.

"Constant debate about (whether) we're assessing the trainee in comparison with where we think they should be at for the level of training or we're assessing the trainee in comparison to where they should be at when they finally complete training. So you get a wide disparity in terms of the numbers." (SoT 2)

Trainees consequently reported marked variability in their supervisors' scoring.

"If you're a brand new beginner you're supposed to score a one or a two but that doesn't seem to be generally how people score it. People score it as sort of a percentage...." (Trainee 9)

Both groups agreed that, when the independence scale was understood, it facilitated giving low scores to a trainee

when they required close but appropriate supervision with a case.

"It's easier to say that a candidate needs supervision rather than say they're performing at a lower level."
(SoT 6)

Case and assessor selection

Participants described several factors as informing their decision-making and practices concerning case and assessor selection, including the perceived opportunity that the case would offer to demonstrate competence or maximize learning. A number of both SoTs and trainees reported that they aimed to complete an assessment when they thought the trainee had reached a level of competence that would be reflected in a good mark using the mini-CEX, rather than to seek feedback to facilitate improvement.

"They want to do them when they feel they're ready and are going to do well on them." (SoT 9)

Nevertheless, to increase the value of the mini-CEX to facilitate learning, some SoTs reported that the case complexity should extend trainees to the limits of their capabilities.

"I look at my list and the ASA status and the type of surgery the patient's having and then decide which surgery and patient would put the trainee on the learning edge... So we'll stretch them to the point where they're actually being challenged." (SoT 1)

Trainee participants also reported targeting their cases for mini-CEX to enhance their learning.

"The trainee needs to select an appropriately complex case that they don't feel confident managing and where they are genuinely interested to know about a better way of doing things." (Trainee 2)

In addition to case selection, assessor selection was influenced by desirable assessor characteristics, including familiarity with the trainee, good teacher, willing or engaged with the mini-CEX process, leniency, or provider of quality feedback.

Discussion

Our exploration of experiences with the mini-CEX in the ANZCA training scheme revealed key findings – i.e., the purpose of the assessments is variously perceived, and this confusion affects the perceived value of the assessments

and their feasibility. We have chosen to concentrate on these most significant aspects of our results in considering their broader implications.

Purpose: summative or formative

The differing perceptions of the purpose for the mini-CEX underscore participants' diverse range of experiences with the exercise. On the one hand, where participants' perceptions align with the underlying "assessment for learning" philosophy, it is more likely that difficulties are seen as surmountable, and case selection and scoring is guided by educational value, with the opportunity to provide feedback and promote trainee learning.³³ On the other hand, where this purpose is not understood or acknowledged, the mini-CEX is seen as an administrative burden, leading to the strategic selection of case or assessor with the risk of meaningless "tick box" assessments.¹¹ Others have documented this deleterious effect of the perception of individual WBAs as summative.^{10,11,33-35}

There is the opportunity to establish a virtuous cycle, where supervisors and trainees experience meaningful assessments that motivate them to overcome the obstacles and realize the potential of the mini-CEX to facilitate trainee learning. Conversely, there is the risk of establishing a vicious cycle of meaningless assessments where doubt or scepticism is reinforced and where the potential benefits for trainees are not realized and result in significant opportunity costs in terms of learning and practice. These contrasting cycles are represented graphically in Fig. 1.

Participants' views on whether the mini-CEX is a formative or summative assessment were disparate and often strongly held. These terms have been prevalent in assessment discussions in the past, yet their meaning is still the subject of academic debate.³⁶ Our participants saw summative and formative in terms of the intended purpose of the assessment. While summative implies that the purpose of the mini-CEX is to contribute to decisions regarding trainee progression, formative implies that the purpose is trainee learning. This is consistent with the historical use of these terms in the health professions education community.³⁷

Taras³⁶ suggests that all assessment is summative as it requires a judgement to be made, and that it becomes formative only if the learner uses the information generated to improve subsequent performance. In contrast, our participants generally interpreted summative and formative as mutually exclusive characteristics of assessments. This absolute dichotomy has been disputed for some time in the education literature.³⁸ For example, in 2003, Ramsden wrote that "the two separate worlds of

Fig. 1 Virtuous vs vicious cycle. In a virtuous cycle, a focus on trainee learning motivates participants to overcome barriers, select challenging cases, and generate useful feedback to maximize trainee learning. In contrast, a focus on completion of assessments leads to trivialization, with a vicious cycle of lenient scoring of easy cases and little feedback or learning which discourages future engagement

Fig. 2 Assessment for learning. On the left, the supervisor and trainee use the individual mini-Clinical Evaluation Exercise (mini-CEX) to maximize trainee learning, while on the right, each mini-CEX contributes with other workplace-based assessments (WBAs) to the aggregate data used by the Competence Committee to inform judgements on the trainee's progression through training

assessment called ‘formative’ and ‘summative’ do not exist in reality.”³⁹ Similarly, Boud’s words from 2000 seem prophetic: “Every act of assessment we devise or have a role in implementing has more than one purpose. If we do not pay attention to these multiple purposes we are in danger of inadvertently sabotaging one or more of them.”⁴⁰

As discussed above, the intention in adopting programmatic assessment is for individual evaluations to inform teachers and learners and to facilitate learning. In contrast, aggregate data from multiple assessments should be used only in high-stakes decisions by appropriately qualified personnel.^{9,41} The distinction between the roles of supervisor and supervisor of training is crucial. Our results suggest there is a lack of clarity in differentiating the role of the SoT in making decisions on trainee progress using aggregate data and the role of supervisors and trainees in using individual assessments to maximize information for trainee learning. Furthermore, conflating these roles has contributed to the disparate views of the purpose of the assessments. Importantly, the Royal College of Physicians

and Surgeons of Canada, in their Competence by Design framework, specifies that a competence committee should be established to perform the summative function rather than having an individual residency coordinator or program director assume this role.⁴² If the ANZCA were to adopt such a solution, it might result in a clearer delineation of these separate functions, i.e., it would create a greater separation between those completing the individual mini-CEXs and those using the aggregated data for summative purposes. A diagrammatic representation of the different but connected uses of the mini-CEX for these purposes is represented in Fig. 2.

Value for trainee learning

Despite the participants’ diverse understanding of the purpose of the mini-CEX, there are nevertheless indications in our data that use of the mini-CEX can help make a valuable contribution to trainee learning. We have provided further evidence that use of the mini-CEX does

lead to increased observation and feedback from supervisors, thus providing the expert guidance that facilitates learning from practice.⁴³ We have examples of trainees and supervisors selecting challenging cases at the boundaries of their capabilities, consistent with Vygotsky's concept of the zone of proximal development.^{44,45}

Some participants valued narrative feedback more highly than scores, as others have reported.⁴⁶ Scores act as summaries and may conceal the details behind assessor judgements. With the understanding that different assessor judgements may indicate legitimate unique perspectives rather than error,⁴⁷ narrative data become important in capturing disparate views on performance to guide both learning and assessment. Recognition of the importance of narrative feedback is consistent with the emerging understanding that competence is highly contextual and more likely to be captured by descriptions of assessments that are rich in information.⁴⁸

The challenge of the feasibility of the mini-CEX and the multiple tasks asked of supervisors performing the assessments suggest that further exploration of the relative contributions of scoring and narrative feedback in the mini-CEX would be useful.⁴⁹

Value for decisions regarding trainee progression

The reports of supervisors and trainees strategically selecting cases, together with assessor leniency and variability in scoring, make it unsurprising that few participants thought the mini-CEX contributed meaningfully to decisions regarding trainee progression. Others have reported challenges in using WBAs for this purpose,¹² even when dedicated assessment panels are used. The ability of SoTs to make these judgements has been assumed, but we would contend that the complexity of this task is not currently recognized⁵⁰ and requires further investigation.

Programmatic assessment relies on accumulating evidence of competence to support robust decisions on trainee progression and attainment of fellowship.^{9,51} Hence, if the mini-CEX and other WBAs are to contribute to these decisions, they cannot be purely formative in nature. The dilemma is that our participants are not alone in reporting that perceptions of a summative use impair engagement and encourage trivialization, compromising the validity of the assessments and their value in contributing to decisions on trainee progression.^{10-12,33,34,46} Our findings support what van der Vleuten *et al.* predicted, that "without formative value the summative function would be ineffective".³⁵ If we are to make use of the mini-CEX and other WBAs in aggregate to inform summative decisions, we must first maximize their formative value as individual assessments.

Table 2 Recommendations for mini-CEX implementation

Key recommendations for trainees:

- Scores reflect your supervisor's estimate of your degree of independence for the case observed- low scores may be appropriate for complex cases
- Focus on creating learning opportunities by selecting challenging cases at the cusp of your ability
- Encourage supervisors to provide specific detailed performance-focused feedback
- Use your own self-assessment and the feedback you receive to plan future learning experiences
- Trust that the supervisor of training will take case complexity and your experience and approach to learning into account when using individual assessments as part of the evidence for decision-making on your progress through training

Key recommendations for supervisors:

- Scores reflect your estimate of the trainee's degree of independence for the case observed - low scores may be appropriate for complex cases
- Your assessment and feedback are based on what you observe, not on a comparison with a standard or expectation, so there is no pass/fail decision
- Focus on creating learning opportunities by encouraging trainees to select challenging cases at the cusp of their ability
- Provide specific detailed performance-focused feedback and help trainees plan future learning experiences
- Trust that the supervisor of training will consider the case complexity and the trainee's experience and approach to learning when using individual assessments as part of the evidence for decision-making on trainee progress through training

Key recommendations for supervisors of training:

- Assessment decisions are based on all available information, and the more important the decision the more evidence required to support it.
- Interpreting individual mini-CEX when aggregating assessment data requires knowledge of the context of the assessment – i.e., the experience and stage of training of the trainee and the complexity of the case.
- Encourage engagement in learning by rewarding trainees who select challenging cases and enact plans to improve in subsequent cases.

mini-CEX = mini-Clinical Evaluation Exercise

Recommendations for implementation of the mini-CEX and other WBAs are summarized in Table 2.

Feasibility

Consistent with the findings of Fokkema *et al.*, the participants described the difficulty of fitting the mini-CEX into a busy clinical workload as significant.⁵² Those who are most engaged in the innovation did not see this as a problem, and while others saw it as a problem that could be overcome, the least-engaged participants saw this as justification for trivializing the assessment. The difficulty could be overcome by either increasing the available time or increasing the motivation of supervisors and trainees to

circumvent it. Faculty development experience highlights that participants and non-participants perceive the same barriers to participation, yet one group can overcome them.⁵³ This implies that there is the potential for increased engagement to motivate more trainees and supervisors to overcome the existing barriers to performing the mini-CEX and other WBAs.

Clinical teachers^{12,34,54,55} and trainees⁴⁶ have previously identified lack of time as a barrier to educational engagement. While improved engagement may partly address this issue, we would argue that the greatest barrier to improving the quality of training is the environment in which training takes place. The quality of postgraduate training can have a persistent effect on subsequent health outcomes.^{56,57} While current patient care is of paramount importance, ensuring that reflection and feedback are encouraged and trainee learning and assessment are maximized would better serve the health of future patients. Achieving a high-quality training system requires re-defining our working conditions to embed learning and assessment in our work practices and create time and intellectual space for training.

Limitations

Although we purposively sampled for maximum variation amongst participants and continued interviewing until new ideas did not emerge, the number of participants is a small proportion of trainees and supervisors of training. As an exploratory interview study, our aim was to capture the variation in participants' experiences using the mini-CEX; however, this methodology does not allow us to define the proportion of trainees and supervisors who share the various views described. This would be an avenue for further research. Our study is confined to Australia and New Zealand; however, developments in curriculum design and challenges in supervisory practices in anesthesia training are widespread. The insights gained in exploring the implementation of the mini-CEX likely have relevance in other settings.

Conclusion

An attitudinal shift away from the dichotomous view of assessment as summative or formative to a clear focus on individual assessments designed for improving trainee performance and hence future patient care will require explanation and education. It will be challenging to optimize the potential for enhanced learning while still satisfying the requirement to make sound decisions on trainee progression using aggregate data. Nevertheless, our results suggest that focusing only on the summative

purpose of the mini-CEX and other WBAs carries a substantial risk of trivialization.

Our results have implications that extend beyond the implementation of the mini-CEX within ANZCA. The introduction of assessment for learning in a competency-based curriculum requires increasing engagement by trainees and supervisors in education and perhaps a greater sophistication in their understanding of current educational practice and intended outcomes of assessments such as the mini-CEX. This calls for an examination of our work practices and a rebalancing of the priorities of service and training.

Acknowledgements We gratefully acknowledge the assistance of the ANZCA Clinical Trials Network in recruiting interview participants.

Funding This work was supported by a project grant from the Australian and New Zealand College of Anaesthetists (S14/002).

Conflicts of interest None declared.

Author contributions Damian J. Castanelli, Yan Chen, and Jennifer M. Weller were involved in the study design. Damian J. Castanelli, Yan Chen, Jennifer M. Weller, and Tanisha Jowsey were involved in the data analysis. Damian J. Castanelli and Tanisha Jowsey were involved in manuscript writing. Yan Chen was involved in recruitment of participants and conducting the interviews. All authors contributed to the interpretation of the results and critical revision of the manuscript.

Editorial responsibility This submission was handled by Dr. Gregory L. Bryson, Deputy Editor-in-Chief, *Canadian Journal of Anesthesia*.

References

1. Iobst WF, Sherbino J, Cate OT, et al. Competency-based medical education in postgraduate medical education. *Med Teach* 2010; 32: 651-6.
2. Miller GE. The assessment of clinical skills/ competence/ performance. *Acad Med* 1990; 65(9 Suppl): S63-7.
3. Norcini J, Burch V. Workplace-based assessment as an educational tool: AMEE Guide No. 31. *Med Teach* 2007; 29: 855-71.
4. Scheele F, Teunissen P, Van Luijk S, et al. Introducing competency-based postgraduate medical education in the Netherlands. *Med Teach* 2008; 30: 248-53.
5. Holmboe ES, Sherbino J, Long DM, Swing SR, Frank JR. The role of assessment in competency-based medical education. *Med Teach* 2010; 32: 676-82.
6. Schuwirth LW, Van der Vleuten CP. Programmatic assessment: From assessment of learning to assessment for learning. *Med Teach* 2011; 33: 478-85.
7. Kogan JR, Holmboe E. Realizing the promise and importance of performance-based assessment. *Teach Learn Med* 2013; 25(Suppl 1): S68-74.
8. Schuwirth LW, van der Vleuten CP. General overview of the theories used in assessment: AMEE Guide No. 57. *Med Teach* 2011; 33: 783-97.

9. van der Vleuten CP, Schuwirth LW, Driessen EW, et al. A model for programmatic assessment fit for purpose. *Med Teach* 2012; 34: 205-14.
10. Bindal N, Goodyear H, Bindal T, Wall D. DOPS assessment: a study to evaluate the experience and opinions of trainees and assessors. *Med Teach* 2013; 35: e1230-4.
11. Bindal T, Wall D, Goodyear HM. Trainee doctors' views on workplace-based assessments: are they just a tick box exercise? *Med Teach* 2011; 33: 919-27.
12. Bok HG, Teunissen PW, Favier RP, et al. Programmatic assessment of competency-based workplace learning: when theory meets practice. *BMC Med Educ* 2013; 13: 123.
13. *The Foundation Programme*. The Foundation Programme Curriculum 2016 London, U.K.: Academy of Medical Royal Colleges; 2016. Available from URL: <http://www.foundationprogramme.nhs.uk/curriculum/> (accessed August 2016).
14. Cho SP, Parry D, Wade W. Lessons learnt from a pilot of assessment for learning. *Clin Med (Lond)* 2014; 14: 577-84.
15. Rees CE, Cleland JA, Dennis A, Kelly N, Mattick K, Monrouxe LV. Supervised learning events in the foundation programme: a UK-wide narrative interview study. *BMJ Open* 2014; 4: e005980.
16. Ringsted C, Ostergaard D, Scherpbier A. Embracing the new paradigm of assessment in residency training: an assessment programme for first-year residency training in anaesthesiology. *Med Teach* 2003; 25: 54-62.
17. *The Royal College of Anaesthetists*. Curriculum for a CCT in Anaesthetics 2010. Available from URL: <http://www.rcoa.ac.uk/> (accessed August 2016).
18. Boker AM. Toward competency-based curriculum: Application of workplace-based assessment tools in the National Saudi Arabian Anesthesia Training Program. *Saudi J Anaesth* 2016. Available from URL: <http://www.saudija.org/preprintarticle.asp?id=179097> (accessed August 2016).
19. *Australian and New Zealand College of Anaesthetists*. Anaesthesia Training Program Curriculum 2012. Available from URL: <http://www.anzca.edu.au/documents/anaesthesia-training-program-curriculum.pdf> (accessed August 2016).
20. Weller JM, Jones A, Merry AF, Jolly B, Saunders D. Investigation of trainee and specialist reactions to the mini-clinical evaluation exercise in anaesthesia: implications for implementation. *Br J Anaesth* 2009; 103: 524-30.
21. Weller JM, Jolly B, Misur MP, et al. Mini-clinical evaluation exercise in anaesthesia training. *Br J Anaesth* 2009; 102: 633-41.
22. Weller JM, Misur M, Nicolson S, et al. Can I leave the theatre? A key to more reliable workplace-based assessment. *Br J Anaesth* 2014; 112: 1083-91.
23. Rekman J, Gofton W, Dudek N, Gofton T, Hamstra SJ. Entrustability scales: outlining their usefulness for competency-based clinical assessment. *Acad Med* 2016; 91: 186-90.
24. Schwandt TA. Three epistemological stances for qualitative inquiry: interpretivism, hermeneutics and social constructionism. In: Denzin NK, Lincoln YS, editors. *Handbook of Qualitative Research*. 2nd ed. Newbury Park, CA: Sage; 2000 .
25. *Australian and New Zealand College of Anaesthetists*. Mini Clinical Evaluation Exercise (Mini-CEX) paper form, Feb 2012. Available from URL: <http://www.anzca.edu.au/documents/mini-cex.pdf> (accessed August 2016).
26. Liampittong P, Ezzy D. *Qualitative Research Methods*. 2nd ed. South Melbourne: Oxford University Press; 2005 .
27. Kitto SC, Chesters J, Grbich C. Quality in qualitative research. *Med J Aust* 2008; 188: 243-6.
28. Taylor SJ, Bogdan R. *Introduction to Qualitative Research Methods: The Search for Meaning*. NY: John Wiley & Sons; 1984 .
29. Savin-Baden M, Major CH. *Qualitative Research: the Essential Guide to Theory and Practice*: Routledge; 2013.
30. Morse JM, Field PA. *Qualitative Research Methods for Health Professionals*, 2nd ed. California, USA: Sage Publications; 1995: 254.
31. Saldana J. *The Coding Manual for Qualitative Researchers*. London: Sage Publications; 2009 .
32. QSR. NVivo Version 10. 2010. Available from URL: <http://www.qsrinternational.com/> (accessed August 2016).
33. Heeneman S, Oudkerk Pool A, Schuwirth LW, van der Vleuten CP, Driessen EW. The impact of programmatic assessment on student learning: theory versus practice. *Med Educ* 2015; 49: 487-98.
34. Massie J, Ali JM. Workplace-based assessment: a review of user perceptions and strategies to address the identified shortcomings. *Adv Health Sci Educ Pract* 2016; 21: 455-73.
35. van der Vleuten CP, Schuwirth LW, Scheele F, Driessen EW, Hodges B. The assessment of professional competence: building blocks for theory development. *Best Pract Res Clin Obstet Gynaecol* 2010; 24: 703-19.
36. Taras M. Assessment for learning: assessing the theory and evidence. *Procedia - Social and Behavioral Sciences* 2010; 2: 3015-22.
37. Downing SM, Yudkowsky R. *Assessment in Health Professions Education*. Routledge; 2009.
38. Boud D, Soler R. Sustainable assessment revisited. *Assessment & Evaluation in Higher Education* 2015. DOI:10.1080/02602938.2015.1018133.
39. Ramsden P. *Learning to Teach in Higher Education*. 2nd ed. London: Routledge; 2003 .
40. Boud D. Sustainable assessment: rethinking assessment for the learning society. *Studies in Continuing Education* 2000; 22: 151-67.
41. Olupeliyawa A, Balasooriya C. The impact of programmatic assessment on student learning: what can the students tell us? *Med Educ* 2015; 49: 453-6.
42. *Royal College of Physicians and Surgeons of Canada*. Meantime Guide: Competence Committee - Establish a Competence Committee - Competence by Design. Available from URL: <http://www.royalcollege.ca/rcsite/documents/cbd/meantime-guide-competence-committee-e.pdf> (accessed August 2016).
43. Ericsson KA. Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Acad Med* 2004; 79: S70-81.
44. Verenikina I. Scaffolding and learning: its role in nurturing new learners. In: Kell P, Vialle W, Konza D, Vogl G, editors. *Learning and the Learner: Exploring Learning for New Times*. Wollongong, NSW: University of Wollongong; 2008 .
45. Slater RJ, Castanelli DJ, Barrington MJ. Learning and teaching motor skills in regional anesthesia: a different perspective. *Reg Anesth Pain Med* 2014; 39: 230-9.
46. Driessen E, Scheele F. What is wrong with assessment in postgraduate training? Lessons from clinical practice and educational research. *Med Teach* 2013; 35: 569-74.
47. Gingerich A, Kogan J, Yeates P, Govaerts M, Holmboe E. Seeing the 'black box' differently: assessor cognition from three research perspectives. *Med Educ* 2014; 48: 1055-68.
48. Govaerts M, van der Vleuten CP. Validity in work-based assessment: expanding our horizons. *Med Educ* 2013; 47: 1164-74.
49. Govaerts M. Workplace-based assessment and assessment for learning: threats to validity. *J Grad Med Educ* 2015; 7: 265-7.
50. Tavares W, Eva KW. Exploring the impact of mental workload on rater-based assessments. *Adv Health Sci Educ Theory Pract* 2013; 18: 291-303.
51. Driessen EW, van Tartwijk J, Govaerts M, Teunissen P, van der Vleuten CP. The use of programmatic assessment in the clinical workplace: a Maastricht case report. *Med Teach* 2012; 34: 226-31.

52. Fokkema JP, Teunissen PW, Westerman M, et al. Exploration of perceived effects of innovations in postgraduate medical education. *Med Educ* 2013; 47: 271-81.
53. Steinert Y, Macdonald ME, Boillat M, et al. Faculty development: if you build it, they will come. *Med Educ* 2010; 44: 900-7.
54. Castanelli DJ, Smith NA, Noonan CL. Do anaesthetists believe their teaching is evidence-based? *Med Teach* 2015; 37: 1098-105.
55. Dijksterhuis MG, Schuwirth LW, Braat DD, Teunissen PW, Scheele F. A qualitative study on trainees' and supervisors' perceptions of assessment for learning in postgraduate medical education. *Med Teach* 2013; 35: e1396-402.
56. Asch DA, Nicholson S, Srinivas S, Herrin J, Epstein AJ. Evaluating obstetrical residency programs using patient outcomes. *JAMA* 2009; 302: 1277-83.
57. Johnston MJ, Singh P, Pucher PH, et al. Systematic review with meta-analysis of the impact of surgical fellowship training on patient outcomes. *Br J Surg* 2015; 102: 1156-66.