Enhancing Formative Feedback in Orthopaedic Training: Development and Implementation of a Competency-Based Assessment Framework

Natalie Wagner, BSc, [*,¶] Anita Acai, MSc, [*,¶] Sydney A. McQueen, MSc, [*,†] Colm McCarthy, MD, MSc, [*] Andrew McGuire, MD, [*,‡] Brad Petrisor, MD, FRCSC, [*] and Ranil R. Sonnadara, PhD [*,†,¶]

[*]Department of Surgery, McMaster University, Hamilton, Ontario, Canada; [†]Faculty of Medicine, University of Toronto, Toronto, Ontario, Canada; [‡]Department of Surgery, Queen's University, Kingston, Ontario, Canada; and [¶]Department of Psychology, Neuroscience & Behaviour, McMaster University, Hamilton, Ontario, Canada

OBJECTIVE: The purpose of this study was to develop, implement, and evaluate the effectiveness of an assessment framework aimed at improving formative feedback practices in a Canadian orthopaedic postgraduate training program.

METHODS: Tool development began in 2014 and took place in 4 phases, each building upon the previous and informing the next. The reliability, validity, and educational impact of the tools were assessed on an ongoing basis, and changes were made accordingly.

RESULTS: One hundred eighty-two tools were completed and analyzed during the study period. Quantitative results suggested moderate to excellent agreement between raters (intraclass correlation coefficient = 0.54-0.93), and an ability of the tools to discriminate between learners at different stages of training (p's < 0.05). Qualitative data suggested that the tools improved both the quality and quantity of formative feedback given by assessors and had begun to foster a culture change around assessment in the program.

CONCLUSIONS: The tool development, implementation, and evaluation processes detailed in this article can serve as a model for other training programs to consider as they move towards adopting competency-based approaches and refining current assessment practices. (J Surg Ed 76:1376–1401. © 2019 Association of Program Directors in Surgery. Published by Elsevier Inc. All rights reserved.)

Correspondence: Inquiries to Ranil Sonnadara, PhD, Office of Education Science, Department of Surgery, McMaster University, 1280 Main Street West, A.N. Bourns Science Building, Rm. 131, Hamilton, Ontario L8S 4M1, Canada; e-mail: info@skillslab.ca

KEY WORDS: Assessment, Formative, Feedback, Orthopaedics, Medical Education, Tool development

COMPETENCIES: Practice-Based Learning and Improvement, Medical Knowledge, Patient Care, Professionalism, Interpersonal and Communication Skills

BACKGROUND

The implementation of competency-based medical education (CBME) frameworks around the world means that trainee progress will now rely upon the demonstration of competence rather than length of time spent in training.[1] As such, robust assessment practices have become a focal point for medical educators. Assessment described as "summative" is typically infrequent, high stakes, and serves to determine if competence has been attained.[2] To date, this type of assessment has been the focus of most residency programs. As an example, end-of-rotation evaluations are, in many cases, accompanied by little to no feedback, and may be completed well after the rotation has ended.[3] Assessment described as "formative," on the other hand, is typically lower stakes and takes place on an ongoing basis, guiding learners along the path to competence through immediate feedback.[4] Formative feedback has been identified as "one of the most powerful influences on learning,"[5] as it can enhance learning by helping learners track their progress, target teaching and resources to learners' needs, and identify learners who need extra support early on.[6]

Despite its theoretical benefits, Hattie, Timperley[5] note that in practice, it takes a great deal of skill to deliver feedback effectively, and not all educators are

  1931-7204/$30.00 https://doi.org/10.1016/j.jsurg.2019.03.015

good at doing this. For example, Ahmed et al.[7] found that feedback given to surgical trainees is often minimal, given in less than 50% of cases, and when it is given, the feedback is typically unidirectional and nonspecific, making it ineffective in providing learners with concrete strategies on how to improve. Other research has found that faculty are not willing to assign low marks to poorly performing trainees, creating a "failure to fail" phenomenon.[8] McQueen et al.[9] conducted interviews with 22 physician educators from 4 institutions, and uncovered a variety of reasons for the failure to fail phenomenon, including insufficient documentation of trainee performance, a lack of support from the program, insufficient observation by faculty, competing time demands, fear of appeals and legal action, and fear of reciprocated poor faculty evaluations. McQueen et al[9] concluded that in addition to being good practice, formative feedback may help to address the "failure to fail" phenomenon by normalizing discussions about performance and allowing poor performance to become an opportunity for learning as opposed to a more serious issue later on.

In light of the issues around delivering formative feedback in surgical training and the impending need for more robust assessment frameworks to align with the upcoming shift to competency-based curricula, the purpose of this study was to develop and implement a new formative assessment framework in an orthopaedic residency training program at a Canadian postgraduate medical training center. The assessment framework was also designed to help facilitate formative feedback and align with the Royal College of Physician and Surgeons of Canada (RCPSC)'s Competency by Design (CBD) initiative well before the program's formal transition to the new curriculum, thus providing the opportunity for faculty and trainees to become more comfortable with the new language and ideas underpinning CBD. This article will detail the processes of tool development and implementation and share early results pertaining to their effectiveness, thus serving as a useful case study and providing recommendations for other programs as they move toward adopting competency-based approaches and refining current assessment practices.

MATERIALS AND METHODS

Tool development took place in 4 phases. The first phase established the development process and basic guidelines for how the tools should look. The second, third, and fourth phases built upon this initial framework through the systematic collection and incorporation of user feedback, ultimately resulting in the development of 24 novel assessment tools to be used for the formative assessment of technical and nontechnical skills in orthopaedic surgery. Each phase of the development process is described below, and an example of a tool from each phase is provided in Appendices A to D, to illustrate the evolution of the assessment framework.

Context

This study took place within a 5-year orthopaedic postgraduate training program at a Canadian academic center. The program is comprised of approximately 25 residents per year, 3 to 10 fellows per year, and 38 faculty members, 26 of who regularly supervise residents, across 4 different hospital sites. The study was initiated after the training program identified a need to improve formative feedback practices and was conducted in collaboration with a number of education scientists. Ethics approval was obtained from the Hamilton Integrated Research Ethics Board (HiREB-0629).

Phase One

Two orthopaedic procedures, femoral and tibial intramedullary nailing and primary total knee arthroplasty, were selected as the initial procedures with which to begin the tool development process as they are some of the most common procedures seen by residents, with minimal procedural variation amongst faculty. They also encompass both elective and trauma surgeries. Tool development began with the generation of key procedural steps, or a task-specific checklist (TSC), by a resident, faculty surgeon, and education scientist. This initial list was used as a starting point for a focus group discussion. A subset of faculty surgeons who performed these procedures were invited, resulting in 8 attendees for the intramedullary nail focus group and 5 attendees for the total knee arthroplasty focus group (a 100% response rate). Using the TSC template provided, steps were revised, removed, and added to the list by focus group participants. A description of what constitutes "in training" versus "competent" was defined for each step; a "not assessed" rating category was also included for each step. In cases of procedural variation, language was modified to be general enough that different faculty members felt comfortable using the tool, while still including enough detail to provide meaningful feedback to trainees. In addition to the TSC, global rating scales (GRSs) were developed to assess the RCPSC's seven CanMEDS roles that constitute the essential qualities of a competent physician: medical expert, communicator, collaborator, leader, health advocate, scholar, and professional.[10]

Once complete, a draft of each tool was emailed to the focus group participants for review. Feedback was integrated in an iterative fashion until all participants agreed on a final version of each tool (Appendix A). Subsequent piloting of the tools took place across 3 of

the 4 sites and relied upon the voluntary participation of residents, fellows, and faculty. Following the pilot period, a chi-squared analysis was used to evaluate whether the overall performance GRS could discriminate between learners at different stages of training as a measure of construct validity (this was also done for the tools described in phases 2-4). Additionally, one-on-one interviews were conducted with 4 residents and 2 faculty surgeons who had used the tools during the piloting process. The interview transcripts were analyzed for emergent themes using thematic analysis, a flexible mode of "identifying, analyzing, and reporting patterns (themes) within [qualitative] data."[11]

Phase Two

In the second phase of tool development, tools were refined based on feedback from phase 1 (e.g., the addition of GRSs measuring technical skill, surgical judgment, operative flow, and overall surgical performance based on the Objective Assessment of Technical Skills [OSATS; Martin et al.[12]] and a preoperative planning section at the beginning of each tool). Tools for 10 additional procedures (ankle open reduction and internal fixation [ORIF], dynamic hip screw, distal radius fracture, elbow arthroscopy, knee arthroscopy/anterior cruciate ligament repair, open shoulder, total shoulder arthroplasty, trauma shoulder arthroplasty, shoulder arthroscopy, and primary total hip arthroplasty) were developed in a manner similar to phase 1.

The finalized tools (Appendix B) were subsequently made available for voluntary piloting at 4 sites. Telephone interviews with 2 residents and 4 faculty surgeons were conducted to gain feedback on implementation and explore content validity 6 weeks postimplementation. Data collection and analysis followed the same method as described in phase 1. Additionally, 4 evaluators independently rated videos of resident performance for a subset of tools (shoulder arthroscopy, knee arthroscopy, and ankle ORIF) to determine inter-rater reliability. Absolute agreement between evaluators was calculated using a 2-way, random intraclass correlation coefficient (ICC). This model assumes that each subject is assessed by each rater, and raters have been randomly selected. ICC values from 0.50 to 0.75 indicate moderate reliability, values between 0.75 and 0.90 indicate good reliability, and values greater than 0.90 indicate excellent reliability.[13]

Phase Three

In the third stage of tool development, project leaders convened to review the tool development process to date, including feedback received in the first 2 phases and next steps in the project. The interview feedback had revealed that neither residents nor faculty found assessment of the CanMEDS roles informative, as it proved difficult to rate broad-ranging skills such as leader or professional during a single surgical encounter. Thus, the GRSs on the CanMEDS roles were removed, and only the GRSs based on the OSATS remained. The exception was the "overall performance" category, which was changed to "overall entrustment" as the concept of using trust to gauge competence on a given activity was gaining popularity amongst medical educators and the RCPSC.[14,15]

Several new tools were developed at this stage, including 4 tools on nontechnical skills (chief resident, informed consent, requesting a consult, and handover) and 8 tools on technical skills not included in the first 2 phases (lower limb amputation, benign bone tumor resection, open biopsy, club foot, lumbar spine discectomy, slipped capital femoral epiphysis ORIF, soft tissue resection, and supracondylar fracture fixation). The addition of these new tools expanded the utility to more advanced users and addressed user concerns that they did not have tools for certain rotations (e.g., pediatrics). Faculty with expertise in the procedures refined the new tools through individual, face-to-face meetings or via email. Once each new or refined tool was complete (Appendix C), it was subjected to a final review by those involved in its creation and the residency program director.

Phase Four

In the fourth stage of tool development, project leaders reconvened to review the tool development process to date, tool uptake, remaining barriers to tool completion, and alignment with the RCPSC CBD initiative. The tools continued to evolve with respect to their format. In response to faculty feedback that the dichotomous scale used in the TSC ("In training" or "Competent") lacked sensitivity, the research team looked to the literature for other scales. As the RCPSC CBD initiative had recently adopted the Ottawa Surgical Competence Operating Room Evaluation (O-SCORE; Gofton et al.[16]) anchors for their entrustable professional activity assessment forms,[17] the TSC anchors were adapted to reflect this change and ensure alignment. Other changes in this phase included updates to the formatting and layout of the tools, including the addition of sections to indicate the complexity of the surgical procedure being assessed and the resident's role (i.e., primary or first-assist). The prompt for narrative feedback was also changed to "Give at least 1 specific suggestion the resident should work on improving," in order to address previously documented challenges with the quality and specificity of narrative feedback (Appendix D).

Prior to the implementation of the phase 4 tools, the program decided that residents would be expected to complete and submit 1 tool per week. To facilitate this, the assessments were collated into binders and placed in the operating room (OR) and/or the OR lounge at each site, along with a locked drop box. Chief residents at each site were expected to bring the drop box to grand rounds on a biweekly basis so the research team could collect the assessments. The division chair and program director informed residents and faculty surgeons of these expectations via email while the research team presented the plan to residents at grand rounds. It was also decided that monthly summaries of the number of evaluations completed should be sent to the program on an ongoing basis to address uptake barriers.

RESULTS

Tool Completion

Tool completion rates for each phase of the study are described in Table 1. As stated earlier, tool completion in phases 1, 2 and 3 was for pilot purposes and thus, voluntary; however, residents were required to complete 1 assessment per week in phase 4. This increased completion rates considerably.

Reliability

ICC values for the 3 representative procedures examined are shown in Table 2. These values suggest good to excellent agreement between responses for the shoulder arthroscopy and knee arthroscopy procedures, and moderate agreement for the ankle ORIF procedure.

Construct Validity

As phase 1 and 3 did not have assessments completed by residents across all 5 postgraduate years (PGYs), chi-squared analyses were only conducted on the assessments completed during phase 2 and 4. The results of the chi-squared analyses suggest that when data were available from all PGYs, the tools were able to discriminate between learners' PGYs of training (Table 3).

TABLE 2. This Table Describes the Absolute Agreement Among 4 Raters Who Independently Rated Performance on a Subset of the Tools From Phase 2

Procedure	ICC	Confidence Interval
Shoulder arthroscopy	0.91	0.87 to 0.93
Ankle ORIF	0.69	0.54 to 0.79
Knee arthroscopy	0.87	0.80 to 0.91

TABLE 3. This Table Describes the Chi-Squared Values (χ^2) for Phases 2 and 4

Phase	Tools Completed	χ^2(df)	p Value
2	45	56.00 (4)	<0.01
4	116	23.95 (4)	<0.01

Qualitative Interview Data

Assessment Tool Use

Participants reported that the tools were quick and easy to use, and took less time once faculty became more familiar with the tools.

> *"Usually the first time we sit down with the faculty and do it, it takes about ten to fifteen minutes, and the second or third time you do it usually it takes between five to fifteen minutes" (Resident 3, 2016)*

Most often the assessments were completed in the OR or the surgeons' lounge immediately following the case. However, participants noted that this was not always possible, especially when cases ended in the middle of the night or other priorities arose. Under these circumstances, assessments were completed at the end of the day, or on the following day.

> *"[The faculty's] preference is to do it right after the operation, but for one [case] at like 2 in the morning we had a bunch of other stuff to do . . . so, he said, 'Why don't we meet first thing tomorrow morning in the clinic,' and that's where we did it" (Resident 2, 2015)*

TABLE 1. This Table Describes the Number of Assessments Completed, the Number of Residents and Evaluators Who Participated, the Implementation Window, and Average Number of Tools Completed per Month in Each Phase

Phase	Tools Completed	Participating Residents	Participating Evaluators	Implementation Period	Average Tools Completed /Month
1	9	5/26	6/32	Jan 2015 to March 2015	3.00
2	45	8/31	10/32	Jan 2016 to Dec 2016	3.75
3	12	5/30	8/32	April 2017 to Aug 2017	2.40
4	116	22/31	30/32	Sept 2017 to Sept 2018	8.92

Utility of the Task-Specific Checklist

Many participants described the TSC as the most useful component of the assessment tools. By having a detailed description of what a competent performance looks like for each step, the TSCs served as a reminder that each procedure is comprised of multiple, individual competencies (i.e., milestones), something that is often overlooked when using a GRS.

> *"This is a new way of thinking about how we learn to operate. Previously, I mean some people think about different steps, but a lot of the time, it's just the [procedure as a] whole . . . now it's explicit that there's multiple different steps and they can be learned [and evaluated] as individual steps" (Resident 3, 2015)*

Additionally, the specificity of the TSCs meant the assessment tools could be incorporated into the curriculum as a teaching tool for novice learners, clarifying what is expected of all trainees.

> *"I think it's very useful as both a teaching tool as well as an evaluation tool . . . I want to allow them to do [the procedures] but I won't let them do them until they know all the steps . . . so, it's useful in that way as well" (Faculty 2, 2015)*

> *"I think it provides clarity of what is expected in the OR, and then the surgeon can expect it" (Resident 1, 2015)*

Most importantly, the TSCs provided faculty with a framework to structure the feedback they gave to residents at the end of a procedure. Residents reported that this not only increased the amount of feedback they received but also the quality, as faculty were prompted to consider each individual step and comment on both strengths and weaknesses throughout the procedure.

> *"They give you feedback as they go through each step a lot of the times so you get feedback on things you wouldn't always be getting feedback on" (Resident 3, 2016)*

Lastly, as the tools were often completed as a collaborative effort between the faculty and resident, the TSCs provided trainees with the opportunity to reflect on their own performance and set goals for the future.

> *"We've got lots of areas we might have struggled with in the past but didn't really think about it afterwards that much. So, just doing these assessments highlights the areas [where] we might have struggled a bit, and [we] can read up on [those things] and maybe improve on them for the next procedure" (Resident 2, 2016)*

Culture Change Around Assessment

Residents reported that using the assessment tools to provide feedback and reflect on their performance was beginning to foster a culture change around assessment in the program. When faculty decided to use a tool, residents said that it identified education as one of the priorities for the day, thus creating an environment that welcomed feedback.

> *"[Using the tools] kind of declares the objective, at least the objective of part of the case, as being focused on education . . . It opens up feedback for not only the surgeon, but anyone else that is assisting. They kind of feel prompted to give harder questions and feedback within the OR" (Resident 1, 2015)*

Both residents and faculty reported that the assessments empowered residents to ask if they could participate in a procedure when they wanted feedback on something previously identified as an area for improvement or wanted to receive feedback on something they had never been evaluated on before.

> *"So, for the next case they can go to their faculty person and say, 'Look I've already done X, Y and Z, I need to focus on, you know, A, B and C right now,' so, and then that allows them to help tailor what they're doing. Because in this fiscal climate of decreasing OR resources and time, sometimes [the residents] don't actually get to do the entire case" (Faculty 1, 2015)*

> *"[The tools] make it way easier for you to go over to a faculty and be like, 'Hey, I have never REALLY been evaluated on this step of the procedure—I want to do that today'" (Resident 2, 2015)*

Additionally, residents noted that the assessment tools reminded faculty to let residents try to troubleshoot before jumping in to take over the case.

> *"When a faculty knows you're going to be [doing a procedure], they leave you to do the procedure until you struggle, and then they help guide you through the procedure so you get to think a little bit more critically about the procedure, what do you have to do next, and all the steps of the procedure" (Resident 2, 2016)*

Moving forward, participants suggested that the data generated from the formative assessments could also be considered in aggregate to inform end-of-rotation decisions, shedding light on trainee progression at multiple time points.

"I think that probably the most value is looking at serial assessments of yourself throughout a rotation showing that you have actually made progress. I think that's probably the most useful aspect of the tool" (Resident 1, 2015)

"[The assessments] should be used throughout the whole rotation and you see progress with the tools, and if you don't see progress with the tools it helps you with [the] end-of-rotation decision" (Faculty 1, 2016)

Barriers to Implementation

Despite the positive response to the assessment tools, participants identified a number of barriers to the implementation process. Lack of time was commonly reported as a barrier, as there is rarely free time between cases to have discussions about performance and progress.

"*For any evaluation in surgery, the sooner you do it after the procedure the better it is, and the more useful it is for the learner. So, the drawback is that the evaluations will all be the most accurate [if completed] when the surgeons and the residents tend to be the most busy" (Resident 2, 2015)*

Residents also reported that faculty reluctance to complete evaluations was a major barrier, indicating they were uncomfortable repeatedly asking faculty to complete assessments. At the same time, faculty reported that they felt pressured to allow residents to do more of the case when using the tools.

"I think the biggest weakness of the tool is that you sort of have to force the feedback out of the faculty, and not all the faculty were always happy about it. It's getting better for sure . . . but you have to persistently tell the faculty that you wanted to get them done" (Resident 03, 2016)

"I think sometimes [the tools] made it feel like I'm supposed to let the residents do the entire case, and if there's a point where I want to do a certain part of the case, I'd have to justify why I'm not letting the resident do the case . . . that made me almost a little uncomfortable with the expectation of how much [the residents] are supposed to do and there's no way to explain why I had the feeling" (Faculty 03, 2016)

In addition to these barriers, participants also reported that the tools did not always capture the factors influencing whether a step was assessed or not. For example, a resident may not have been assessed on a certain step due to patient-specific factors (e.g., a complicated case), the OR running late, or the faculty not trusting the resident to attempt the step for some other reason.

"Sometimes the tool doesn't necessarily capture time crunches or things of that nature . . . it is not necessarily [the] inability of residents to do things" (Faculty 1, 2016)

"If [the resident] does something bad, and I have to basically take over the case, there's no way to sort of explain that . . . Did the resident cut the patella tendon? Well, I mean it's game over, they're not touching my patient again if they did, you know what I mean. You don't really get a place to explain why they stopped at a certain point in the case . . . I almost think [the not assessed rating] would make [the resident] appear slightly more competent [than they were]" (Faculty 3, 2016)

Last, participants reported that using paper-based tools was burdensome. The paper copies were not always readily available in the OR, and inputting and collating the data was time- and resource-intensive. Participants wished for a platform that could provide up-to-date summaries to faculty and residents to help track progress and promote targeted studying and teaching.

"I think that it needs to be electronic. I know for a fact other people are using electronic and I think that provides quicker feedback, and is able to be tracked easier so that people know what you have and haven't done . . . it takes the guess work out of it" (Resident 1, 2015)

"I like the digital option . . . there have been times where we've tried to go ahead and do the assessment and they're out of the paper copies" (Resident 3, 2016)

DISCUSSION

With the exception of phase 3, which was a short pilot period, assessment completion rates improved over the 4 years (Table 1). Although this process was neither quick nor easy, valuable lessons were gained on how to successfully implement a new assessment framework. The following discussion highlights some considerations for educators who are creating and/or implementing assessment tools within the context of CBME.

Establishing Buy-In

In order to implement the new assessment framework, gaining support from the program director and division chair was paramount; these individuals acted as

educational leaders, had the experience and ability to convince others of the need for change, and could provide adequate resources to the project. Furthermore, involving faculty in the development process was extremely useful in getting additional buy-in. Faculty who assisted with tool development were much more likely to complete tools. Having faculty discuss the steps of a procedure in focus groups also offered a unique opportunity for faculty development and facilitated conversations on why they choose to use a certain technique over another. Yet, despite this support, tool completion rates were lower than anticipated until the tools were made a formal part of the curriculum and residents were held accountable for having tools completed (Table 1). This highlights the need for some type of administrative follow through; residents and faculty need to know that someone is keeping track of the assessments and is going to hold them accountable as we move forward with CBME.

Psychometric Evidence

Additionally, while the results suggest some evidence of inter-rater reliability and construct validity, the overall purpose of the assessments was to facilitate formative feedback. Thus, it was important to consider other forms of evidence, such as the results of interviews and focus groups with residents and faculty, to determine whether or not the assessment tools had met their intended purpose. It is important for educators to recognize that different assessment goals require different types of evidence to gauge their effectiveness and educational impact. Considering only standard psychometric measures of reliability and validity may not be sufficient when a tool's intended purpose is as a teaching tool rather than a summative evaluation.[18] Furthermore, educators must keep in mind that the validity evidence found in one context may not hold true in another context.[19,20] For example, the results of this study suggest that the assessments generated formative feedback within the context of this program; however, they may not function the same way in other programs. Thus, understanding the *context* in which an assessment tool was developed and the *evidence* used to evaluate the reliability and validity is critical before deciding to include an assessment tool into a training program.

Finding the "Right Tool for the Right Purpose"

In addition to this, educators must focus on finding the "right tool for the right purpose." A good summative assessment tool may not always be a tool that is good for teaching, and vice versa.[21] There is currently a movement toward generic tools, such as entrustment-based scales that can be used across tasks and specialties[22]; however, depending on the purpose of the tools, this approach might not be optimal. In this study, both residents and evaluators reported that the TSCs were the most useful component of the tools as they facilitated detailed discussions and empowered residents to ask to be evaluated on specific aspects of a procedure. However, trying to consolidate formative feedback from a TSC into a summative, pass/fail decision can be challenging, particularly when the assessments are completed by different raters, in different contexts, and may be focused on different steps of a procedure. In light of this, if the purpose of the assessment is to generate formative feedback, a more detailed TSC will help facilitate that; however, if the purpose is to evaluate entrustment of an entire procedure, then perhaps a global rating that relies on the supervisor's gestalt impression is more appropriate and easier to compare to other supervisors' ratings.

Along these same lines, educators must keep in mind that assessments themselves are neither formative nor summative; rather, assessments are used to make formative or summative *decisions*.[23] The notion of amalgamating "formative" assessments to inform summative decisions remains a common misperception in medical education. For example, participants in this study praised how useful the assessment tools would be to capture progress and inform end-of rotation decisions; however, the purpose of the tools was to generate formative feedback. A theoretical assumption of formative feedback is that it is low stakes and offers a learning opportunity rather than a judgment.[24] Thus, were the assessments in this study used to inform end-of-rotation decisions, they would no longer truly be low stakes, and by extension, no longer be formative. The idea that an assessment cannot be both formative and summative, and that certain assessment strategies are effective for formative feedback while others are effective for making summative decisions on competence, further underscore the idea of using the "right tool for the right purpose." This reflects a major change in how assessment is viewed in medical education and highlights the need for more training on how to build an assessment program.

Program-Specific Needs versus Governing Body Recommendations

Throughout this study, the research team also monitored the decisions being made at the RCPSC level. This was to ensure that the assessments were in alignment with what was happening with respect to CBD implementation in Canada, as one of the goals of this project was to ease this transition for the program while also addressing the program-specific needs around formative feedback. In doing this, a hybrid assessment form was created that included some components being used by the governing

body (for example the O-SCORE anchors), as well as other unique components requested by the faculty and trainees. In alignment with CBD and other frameworks, the program also chose to make collecting assessments the residents' responsibility. While this strategy helped residents feel empowered to ask to participate in surgical cases, residents also felt uncomfortable constantly reminding faculty to complete assessments, particularly when the resident was a junior learner or the faculty member did not seem interested. Furthermore, faculty sometimes felt uncomfortable with residents asking to do more of the case, and felt as though the tools pressured them into giving them more autonomy. Therefore, as programs shift toward CBME, educators must not to lose sight of program-specific needs and keep an eye on the unintended consequences of assessment in this context, including ensuring that relationships between teachers and learners are not damaged.

Looking Beyond Data Collection

In addition to the content of the assessment tool, educators must also consider what modality they want to use to deliver assessments. This study used paper-based tools and while these were easy to use and readily available, collating the assessment data proved to be extremely resource- and time-intensive. This was a major barrier as residents and faculty wanted real-time analytics on resident performance to help direct future learning. While many electronic platforms are currently being developed, few have optimized data visualization to facilitate this type of engagement. Educators and developers should look past just collecting assessment data, and keep in mind how they want to use the data to ensure whatever method (or technology) they are using collects the *right data* in *the right format*.

Last, given that time has always been, and will continue to be, a barrier to assessment completion, educators should be thoughtful about each assessment that is incorporated into their training program. There should be a clear reason for using each assessment, little duplication between forms/tools to avoid evaluation fatigue,[25] a mechanism for faculty and trainees to provide feedback on the assessments, and a team to review the feedback and implement changes. These components are essential to the success of any assessment program, but with the shift towards CBME, will be of the utmost importance.

By sharing our experiences, we hope others will not only gain valuable insights into how assessment can be incorporated into their programs, but also be inspired to share their own experiences and resources with others in similar programs/contexts, be thoughtful and transparent regarding reliability and validity testing, and recognize that to create a good assessment program it is an iterative process that requires gradual refinement.

LIMITATIONS

While this study was conducted at 4 hospital sites, it only represented 1 training program; future work could involve multiple training programs across multiple sites and cities. Inter-rater reliability was only evaluated for a subset of assessment tools (4/24) due to feasibility. Construct validity could not be examined in phases 1 or 3, as assessments had not been submitted across all PGYs. Assessment completion rates were lower than expected; even in phase 4 when the program mandated completion of 1 assessment/week, not all residents participated. Last, as interviews were only conducted during phases 1 and 2, when completion of the tools was voluntary, the interview results may not be representative of those who chose not to participate. Future work should evaluate whether the interview themes changed after the program made the completion of the assessments mandatory for all residents.

CONCLUSION

This article outlines the development, implementation, and evaluation of an assessment framework aimed at improving formative feedback practices at a Canadian orthopaedic postgraduate training program. The 4 phases of this project describe how the development and implementation process, attitudes toward assessment, and the assessment tools themselves evolved over time. Furthermore, as this truly was an iterative process, this article highlights what works, where challenges exist, and what things educators should consider as they seek to create or revamp assessment practices within their own programs. While the assessment tools created in this project are not perfect and may not be applicable or appropriate for all trainees, the approach taken could be broadly applicable to programs as they move forward with CBME. Key take-home points include involving faculty in the development process; ensuring there is accountability for completed assessments; considering the appropriate validity and reliability evidence given the context of assessment; finding the "right tool for the right purpose"; incorporating both program specific needs and governing body recommendations wherever possible to avoid duplication; and looking beyond data collection to ensure there is a system in place to use the data, obtain feedback, and refine assessments in an iterative manner.

ETHICAL APPROVAL

Ethics approval was obtained from the Hamilton Integrated Research Ethics Board (HiREB-0629).

ACKNOWLEDGMENTS

The authors wish to thank Nathan Cupido and Ali Babar for their assistance with data collection and analysis; Portia Kalun, Jacqueline Wilcox, and Jennifer Zering for their help collecting and updating the tools; Brian Chin and Krista Dunn for their help filming procedures; Abbey Payne for facilitating communication with the residents and faculty over the past four years; and the many faculty members who were involved in the tool development process.

REFERENCES

1. Frank JR, Snell LS, Cate OT, et al. Competency-based medical education: theory to practice. *Med Teach*. 2010;32:638–645.
2. Ferris H, Flynn DO. Assessment in medical education; what are we trying to achieve. *Int J Higher Educ*. 2015;4:139–144.
3. Patel R, Drover A, Chafe R. Pediatric faculty and residents perspectives on In Training Evaluation Reports (ITERs). *Can Med Educ J*. 2015;6:e41–e53.
4. Sadler DR. Formative assessment and the design of instructional systems. *Instr Sci*. 1989;18:119–144.
5. Hattie J, Timperley H. The power of feedback. *Rev Educ Res*. 2007;77:81–112.
6. Sonnadara R, McQueen S, Mironova P, et al. Reflections on current methods evaluating skills during joint replacement surgery. *Bone Joint J*. 2013;95-B:1445–1449.
7. Ahmed M, Sevdalis N, Vincent C, Arora S. Actual vs perceived performance debriefing in surgery: practice far from perfect. *Am J Surg*. 2013;205:434–440.
8. Dudek N, Marks MB, Regehr G. Failure to fail: The perspectives of clinical supervisors. *Acad Med*. 2005;80:S84–S87.
9. McQueen SA, Petrisor B, Bhandari M, Fahim C, McKinnon V, Sonnadara RR. Examining the barriers to meaningful assessment and feedback in medical training. *Am J Surg*. 2016;211:464–475.
10. Frank JR, Snell L, Sherbino J. CanMEDS 2015 Physician Competency Framework. 2015.
11. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol*. 2006;3:77–101.
12. Martin JA, Regehr G, Reznick R, et al. Objective structured assessment of technical skill (OSATS) for surgical residents. *Br J Surg*. 1997;84:273–278.
13. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15:155–163.
14. Hauer KE, Ten Cate O, Boscardin C, Irby DM, Iobst W, O'Sullivan PS. Understanding trust as an essential element of trainee supervision and learning in the workplace. *Adv Health Sci Educ Theory Pract*. 2014;19:435–456.
15. Rekman J, Gofton W, Dudek N, Gofton T, Hamstra SJ. Entrustability scales: outlining their usefulness for competency-based clinical assessment. *Acad Med*. 2016;91:186–190.
16. Gofton WT, Dudek NL, Wood TJ, Balaa F, Hamstra SJ. The Ottawa Surgical Competency Operating Room Evaluation (O-SCORE): a tool to assess surgical competence. *Acad Med*. 2012;87:1401–1407.
17. RCPSC. Understanding workplace based assessment and CBD. In: 2017.
18. Cook DA, Brydges R, Ginsburg S, Hatala R. A contemporary approach to validity arguments: a practical guide to Kane's framework. *Med Educ*. 2015;49:560–575.
19. Messick S. Validity of psychological assessment: validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *Educ Test Serv*. 1994: 1–28.
20. Kane MT. Validating the interpretations and uses of test scores. *J Educ Meas*. 2013;50:1–73.
21. Norcini J, Anderson MB, VBollela V, et al. Consensus framework for good assessment. *Med Teach*. 2018. Epub ahead of print.
22. Halman S, Rekman J, Wood T, Baird A, Gofton W, Dudek N. Avoid reinventing the wheel: implementation of the Ottawa Clinic Assessment Tool (OCAT) in Internal Medicine. *BMC Med Educ*. 2018;18:218.
23. Watling CJ, Ginsburg S. Assessment, feedback and the alchemy of learning. *Med Educ*. 2018.

24. Schut S, Driessen E, van Tartwijk J, van der Vleuten C, Heeneman S. Stakes in the eye of the beholder: an international study of learners' perceptions within programmatic assessment. *Med Educ*. 2018;52:654–663.

25. Fahim C, Bhandari M, Yang I, Sonnadara R. Development and early piloting of a CanMEDS competency-based feedback tool for surgical grand rounds. *J Surg Educ*. 2016;73:409–415.

SUPPLEMENTARY INFORMATION

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.jsurg.2019.03.015.

APPENDIX A. PHASE ONE EXAMPLE TOOL

Primary Total Knee Arthroplasty		
Trainee:	**PGY Level:**	**Date:**
Assessor:	**Location:**	
☐ **Procedure more difficult than usual? If yes, explain:**		
☐ **Procedure easier than usual? If yes, explain:**		

I = In Training, C = Competent, N = Not Assessed; G = General, PS = Procedure-specific

	Competency	In Training (I)	Competent (C)
Preoperative Planning			
G	Appropriate Surgical Indication		
G	Correct Side Identification		
G	Appropriate Pre-operative Planning		
PS	Appropriate Pre-operative X-ray Templating		
PS	Appropriate Implant Selection		
Operative Technique			
G	Correct Patient Positioning	• Patient not placed supine on bed, no tourniquet	• Patient placed supine on bed with tourniquet
G	Proper Gowning & Gloving	• Contamination	• Done appropriately
G	Proper Prepping & Draping	• Contamination	• Done appropriately
Surgical Approach			
G	Anatomical Landmarks Located & Identified Appropriately	• Landmarks not located or identified appropriately	• Landmarks appropriately located and identified
PS	Midline Incision Made Appropriately	• Incision not midline, not anterior to knee • Medial parapatellar approach not used • Inappropriate use of extensile exposure	• Midline, anterior to knee • Medial parapatellar approach used • Appropriate use of extensile exposure if necessary
PS	Proper Medial Release for Exposure	• MCL cut or damaged	• Preservation of MCL
PS	Appropriately Exposes Distal Femur	• Distal femur not exposed appropriately	• Evert/sublux the patella appropriately
PS	Proper Preservation of Patellar Tendon	• Excessive stress placed on patellar tendon or ligament	• Proper preservation
PS	Proper Removal of Patellar Fat Pad	• Patellar tendon cut • Inadequate resection	• Preservation of patellar tendon • Adequate resection

PS	Correctly Identifies Appropriate Start Point for Intramedullary Guide	• Site not identified correctly	• Correct site identified
Femoral Preparation			
PS	Adequate Selection, Placement & Alignment of Femoral Cutting Guide	• Inappropriate size of guide selected • Inappropriate technique used during pinning • Guide inappropriately placed, not checked • MCL and LCL not protected • Saw not safely or appropriately used	• Appropriate size of guide selected • Appropriate technique used during pinning • Guide appropriately placed, checked • MCL and LCL protected during cutting stage • Saw safely and appropriately used
PS	Adequate Cutting of Distal Femur	• Cut is not made safely • Soft tissue and ligaments not protected • Inappropriate rotation • External rotation is not checked (most common error)	• Cut is made safely • Soft tissue and ligaments protected • Appropriate rotation • External rotation is checked
Tibial Preparation			
PS	Safely Dislocates Knee	• Not done safely	• Done safely
PS	Appropriately Places Tibial Cutting Guide	• Cut not performed safely • Patellar tendon not protected	• Cut performed safely with care to protect patellar tendon
PS	Appropriately Prepares & Sizes Tibia	• Osteophytes not removed • Soft tissues not protected • Sizing template not well selected	• Osteophytes removed • Soft tissues protected • Sizing template well selected
Balancing the Knee			
PS	Proper Soft Tissue Balance	• Osteophyte removal not done appropriately • Ligaments and soft tissue release not done appropriately	• Osteophyte removal done appropriately • Ligaments and soft tissue release done appropriately
PS	Appropriate Assessment of Flexion & Extension Gap	• Appropriate release not identified	• Appropriate release identified
PS	Appropriate Trialing	• Inappropriate trial selected, placed incorrectly or not checked • Rotation not checked • Does not mark template locations on tibia with cautery or other technique • Intramedullary cut made inappropriately • Does not considers drilling tibia for interdigitation of cement • Failure to recognize instability	• Appropriate trial selected, placed and checked • Rotation checked • If range is acceptable, marks template locations on tibia with cautery or other technique • Intramedullary cut made appropriately • Considers drilling tibia for interdigitation of cement
PS	Appropriately Confirms Position, Size & Rotation		
PS	Correctly Sizes Patella with Cutting Guide; Cuts, Sizes & Drills Patella	• Failure to measure width/height	• Done appropriately
PS	Appropriately Prepares Bones for Cement	• Not done appropriately	• Done appropriately
PS	Mixes Cement Appropriately	• Inappropriate viscosity	• Appropriate viscosity
PS	Appropriately Cements Prosthesis	• Not enough/too much cement	• Cement appropriately placed with

		used, not well placed • Does not remove excess cement • Does not ensure proper alignment of hardware	proper amount • Removes excess cement • Ensures proper alignment of hardware
PS	Adequately Removes Remaining Debris	• Does not remove excess, wash or suction appropriately • Does not remove debris appropriately	• Remove excess, wash and suctions appropriately • Removes debris appropriately
PS	Appropriate Wound Closure	• Does not range knee to ensure proper alignment before final closure • Contamination • Skin is approximated too tightly • Layered closure not appropriately done • Does not cover with sterile dressing	• Ranges knee prior to final closure. • Appropriate skin tension • Appropriate layered closures • Covers with sterile dressing
G	Sufficient Knowledge of Procedure	• Not appropriate for level of training	• Appropriate for level of training

Global Rating of Performance					
Medical Expert	**1** - **Well below expectations for level of training** - Unable to identify anatomy - Little knowledge of procedure - Inadequate knowledge of surgical approaches - Poor technical skills - Poor use of instruments - Poor handling of soft tissues	**2**	**3** **- Appropriate for level of training** - Identifies important anatomy -Understands important steps of procedure & surgical approaches - Adequate technical skills - Adequate use of instruments & handling of soft tissues	**4**	**5** - **Well above expectations for level of training** - Complete understanding of anatomy - Extensive understanding of procedure & surgical approaches - Manages and minimizes complications - Excellent technical skills, use of instruments & handling of soft tissues
Communicator	**1** - **Well below expectations for level of training** - Difficulties with verbalizing procedural information - Inadequately orders postoperative X-rays - Poor communication with other members of the healthcare team	**2**	**3** **- Appropriate for level of training** - Adequately verbalizes procedural information - Adequately orders postoperative X-rays - Adequate communication with other members of the healthcare team	**4**	**5** - **Well above expectations for level of training** - Clearly verbalizes all procedural information - All orders including postoperative X-rays completed correctly - Excellent communication with other members of the healthcare team
Collaborator	**1** - **Well below expectations for level of training** - Poor use of assistants - Does not recognize limitations - Does not know when to ask for assistance	**2**	**3** **- Appropriate for level of training** - Adequate use of assistants - Sometimes recognizes limitations - Knows when to ask for assistance	**4**	**5** - **Well above expectations for level of training** - Efficient and effective use of assistants - Recognizes limitations -Goes beyond expectations to assist other members of the healthcare team effectively

Leader	**1** - **Well below expectations for level of training** - Does not appropriately lead the healthcare team -Does not know when to follow	**2**	**3** **- Appropriate for level of training** - Appropriately leads the healthcare team - Sometimes follows appropriately	**4**	**5** - **Well above expectations for level of training** - Effectively and efficiently leads the healthcare team - Knows when to follow - Has a sense of vision, direction and purpose
Professional	**1** - **Well below expectations for level of training** - Lacks common courtesy and respect	**2**	**3** **- Appropriate for level of training** - Often demonstrates respect for others	**4**	**5** - **Well above expectations for level of training** - Professional & respectful towards others at all times

Overall Level of Performance				
1	**2**	**3**	**4**	**5**

Please provide one specific piece of feedback:

Final Comments:

APPENDIX B. PHASE TWO EXAMPLE TOOL

Primary Total Knee Arthroplasty Guide		
Trainee:	**PGY Level:**	**Date:**
Assessor:	**Location:**	
☐ **Procedure more difficult than usual? If yes, explain:**		
☐ **Procedure easier than usual? If yes, explain:**		

T = In Training, C = Competent, N = Not Assessed; G = General, PS = Procedure-specific

		Competency	In Training (T)	Competent (C)
Pre-Operative Planning				
1	G	Appropriate surgical indication	Absent	Present
2	G	Appropriate pre-operative planning	Absent	Present
3	G	Patient reviewed and appropriate side initialed	Did not review and/or initial on appropriate side	Reviewed and initialed on appropriate side
4	PS	Appropriate pre-operative X-ray templating	Did not complete	Completed
5	PS	Appropriate implant selection	Incorrect selection	Correct selection
6	G	Confirms equipment	Did not confirm that all appropriate is present	Confirmed that all appropriate equipment is present
7	G	Surgical pre-operative checklist	Did not complete	Completed
8	G	Appropriate gowning & gloving	Contaminated	Sterile
9	G	Appropriate prepping & draping	Contaminated	Sterile
10	PS	Correct patient positioning	Patient not placed supine on bed, no tourniquet	Patient placed supine on bed with tourniquet
Operative Technique				
11	PS	Anatomical landmarks located & identified appropriately	Landmarks not located or identified appropriately	Landmarks appropriately located and identified
12	PS	Midline incision made appropriately	Incision not midline, not anterior to knee, medial parapatellar approach not used, inappropriate use of extensile exposure	Midline, anterior to knee, medial parapatellar approach used, appropriate use of extensile exposure if necessary

13	PS	Proper medial release for exposure	MCL cut or damaged	Preservation of MCL
14	PS	Appropriately exposes distal femur	Distal femur not exposed appropriately	Evert/sublux the patella appropriately
15	PS	Appropriate preservation of patellar tendon	Excessive stress placed on patellar tendon or ligament	Proper preservation
16	PS	Appropriate removal of patellar fat pad	Patellar tendon cut, inadequate resection	Preservation of patellar tendon, adequate resection
17	PS	Appropriate & safe use of retractors	Not used safely or appropriately	Safe and appropriate use
18	PS	Correct identification & removal of osteophytes	Not correctly identified	Correctly identified
19	PS	Appropriate synovectomy	Does not do any synovectomy, does not do complete synovectomy, takes too much tissue beyond synovium	Performs complete synovectomy, proper resection
20	PS	Appropriate removal of medial & lateral menisci	Does not do resection, incomplete resection, resection of too much tissue	Adequate meniscal resection, proper care of soft tissues around meniscus
Femoral Preparation				
21	PS	Correctly identifies appropriate start point for intramedullary guide	Site not identified correctly	Correct site identified
22	PS	Adequate selection, placement & alignment of femoral cutting guide	Inappropriate size of guide selected, inappropriate technique used during pinning, guide inappropriately placed, not checked, MCL and LCL not protected, saw not safely or appropriately used	Appropriate size of guide selected, appropriate technique used during pinning, guide appropriately placed, checked, MCL and LCL protected, saw safely and appropriately used
23	PS	Adequate cutting of distal femur	Cut is not made safely, soft tissue and ligaments not protected, inappropriate rotation, external rotation is not checked (most common error)	Cut is made safely, soft tissue and ligaments protected, appropriate rotation, external rotation is checked
Tibial Preparation				
24	PS	Safely dislocates knee	Not done safely	Done safely
25	PS	Appropriately places tibial cutting guide	Cut not performed safely, patellar tendon not protected	Cut performed safely with care to protect patellar tendon
26	PS	Appropriately prepares & sizes tibia	Osteophytes not removed, soft tissues not protected, sizing template not well selected	Osteophytes removed, soft tissues protected, sizing template well selected
Balancing the Knee				
27	PS	Proper soft tissue balance	Osteophyte removal not done appropriately, ligaments and	Osteophyte removal done appropriately, ligaments and soft tissue release done

			soft tissue release not done appropriately	appropriately
28	**PS**	Appropriate assessment of flexion & extension gap	Appropriate release not identified	Appropriate release identified
29	**PS**	Appropriate trialing	Inappropriate trial selected, placed incorrectly or not checked, rotation not checked, does not mark template locations on tibia with cautery or other technique, intramedullary cut made inappropriately, does not considers drilling tibia for interdigitation of cement, failure to recognize instability	Appropriate trial selected, placed and checked, rotation checked, if range is acceptable, marks template locations on tibia with cautery or other technique, intramedullary cut made appropriately, considers drilling tibia for interdigitation of cement
30	**PS**	Appropriately confirms position, size & rotation	Not done appropriately	Done appropriately
31	**PS**	Appropriately assesses patellar tracking	Fails to complete, does not approximate the medial parapatellar arthrotomy temporarily while tracking	Does tracking with appropriate approximation of arthrotomy, does tracking with components in appropriate rotation
32	**PS**	Correctly sizes patella with cutting guide; cuts, sizes & drills patella	Failure to measure width/height	Done appropriately
33	**PS**	Appropriately prepares bones for cement	Not done appropriately	Done appropriately
34	**PS**	Mixes cement appropriately	Inappropriate viscosity	Appropriate viscosity
35	**PS**	Appropriately cements prosthesis	Not enough/too much cement used, not well placed, does not remove excess cement, does not ensure proper alignment of hardware	Cement appropriately placed with proper amount, removes excess cement, ensures proper alignment of hardware
36	**PS**	Adequately removes remaining debris	Does not remove excess, wash or suction appropriately, does not remove debris appropriately	Removes excess, wash and suctions appropriately, removes debris appropriately
Wound Closure				
37	**PS**	Appropriate wound closure	Does not range knee to ensure proper alignment before final closure, contamination, skin is approximated too tightly, layered closure not appropriately done, does not cover with sterile dressing	Ranges knee prior to final closure, appropriate skin tension, appropriate layered closures, covers with sterile dressing

Global Ratings of Performance					
Surgical Independence					
	Able to perform some steps with close supervision		Able to perform some steps independently		Able to perform entire procedure independently
Technical Skill	**1** - Poor handling of instruments - Poor handling of soft tissues - Improper technique	2	**3** - Adequate handling of instruments - Adequate handling of soft tissues - Adequate technique	4	**5** - Competent handling of instruments - Competent handling of soft tissues - Proper technique
Surgical Judgment	**1** - Cannot anticipate errors - Cannot correct errors	2	**3** - Some anticipation of errors - Can correct some errors	4	**5** - Anticipates errors - Corrects errors with ease
Operative Flow	**1** - Many unnecessary moves - Inefficient - Frequent stops - Often unsure of next move	2	**3** - Some unnecessary moves - Some efficiency - Few stops - Some forward planning	4	**5** - No unnecessary moves - Maximum efficiency - Effortless flow - No hesitation
Overall Performance	**1**	**2**	**3**	**4**	**5**

Global Rating of Performance					
CanMEDS Competencies					
	Well below expectations for level of training		Appropriate for level of training		Well above expectations for level of training
Medical Expert	**1** - Unable to identify anatomy - Little knowledge of procedure - Inadequate knowledge of surgical approaches - Poor technical skills - Poor use of instruments - Poor handling of soft tissues	**2**	**3** - Identifies important anatomy -Understands important steps of procedure & surgical approaches - Adequate technical skills - Adequate use of instruments & handling of soft tissues	**4**	**5** - Complete understanding of anatomy - Extensive understanding of procedure & surgical approaches - Manages and minimizes complications - Excellent technical skills, use of instruments & handling of soft tissues
Communicator	**1** - Difficulties with verbalizing procedural information - Inadequately orders postoperative X-rays - Poor communication with other members of the healthcare team	**2**	**3** - Adequately verbalizes procedural information - Adequately orders postoperative X-rays - Adequate communication with other members of the healthcare team	**4**	**5** - Clearly verbalizes all procedural information - All orders including postoperative X-rays completed correctly - Excellent communication with other members of the healthcare team

Collaborator	**1** - Poor use of assistants - Does not recognize limitations - Does not know when to ask for assistance	**2**	**3** - Adequate use of assistants - Sometimes recognizes limitations - Knows when to ask for assistance	**4**	**5** - Efficient and effective use of assistants - Recognizes limitations -Goes beyond expectations to assist other members of the healthcare team effectively
Leader	**1** - Does not appropriately lead other members of the healthcare team -Does not know when to follow	**2**	**3** - Appropriately leads other members of the healthcare team - Sometimes follows appropriately	**4**	**5** - Effectively and efficiently leads other members of the healthcare team - Knows when to follow - Has a sense of vision, direction and purpose
Professional	**1** - Lacks common courtesy and respect	**2**	**3** - Often demonstrates respect for others	**4**	**5** - Professional & respectful towards others at all times

Comments:

APPENDIX C. PHASE THREE EXAMPLE TOOL

Primary Total Knee Arthroplasty Assessment Tool		
Trainee:	**PGY Level:**	**Date:**
Assessor:	**Location:**	

☐ **Procedure more difficult than usual? If yes, explain:**

☐ **Procedure easier than usual? If yes, explain:**

PLEASE CIRCLE THE APPROPRIATE BOX

Legend: G = General, **PS** = Procedure Specific

		Competency	In Training (T)	Competent (C)	Not Assessed (N/A)
Pre-Operative Planning					
1	**G**	Identified appropriate surgical indication	Absent	Present	Not assessed
2	**G**	Demonstrated appropriate pre-operative planning	Absent	Present	Not assessed
3	**G**	Reviewed patient and initialed appropriate side	Did not review and/or initial on appropriate side	Reviewed and initialed on appropriate side	Not assessed
4	**PS**	Demonstrated appropriate pre-operative templating	Did not complete	Completed	Not assessed
5	**PS**	Selected appropriate implant	Incorrect selection	Correct selection	Not assessed
6	**G**	Confirmed all necessary equipment	Did not confirm all appropriate equipment was ready	Confirmed all appropriate equipment was ready	Not assessed
7	**G**	Completed pre-operative checklist	Did not complete or incorrectly completed checklist	Correctly completed checklist	Not assessed
8	**G**	Demonstrated appropriate sterile technique (E.g. gowning, gloving, prepping, & draping)	Did not use appropriate sterile technique and/or contaminated the field	Sterility maintained throughout setup	Not assessed
9	**PS**	Demonstrated appropriate patient positioning	Patient incorrectly positioned	Patient placed in a supine position with tourniquet	Not assessed
Operative Technique					

10	PS	Identified the appropriate anatomical landmarks	Landmarks not located or identified appropriately	Landmarks appropriately located and identified	Not assessed
11	PS	Completed an appropriate midline incision	Incision not midline, did not use medial parapatellar approach, or inappropriate use of extensile exposure	Used the medial parapatellar approach to create a midline incision, anterior to the knee, with appropriate extensile exposure if necessary	Not assessed
		Competency	**In Training (T)**	**Competent (C)**	**Not Assessed (N/A)**
12	PS	Demonstrated a proper medial release for exposure	MCL cut or damaged	MCL preserved	Not assessed
13	PS	Exposed distal femur	Distal femur not appropriately exposed	Everted/subluxed the patella appropriately	Not assessed
14	PS	Performed arthrotomy	Damage to patellar tendon, damage to tendon insertion	Properly preserved patellar tendon and ligament	Not assessed
15	PS	Removed patellar fat pad	Patellar tendon cut, inadequate resection	Patellar tendon preserved, adequate resection	Not assessed
16	PS	Demonstrated appropriate & safe use of retractors	Retractors not used safely or appropriately	Retractors used safely and appropriately	Not assessed
17	PS	Identified and removed osteophytes	Osteophytes not correctly identified and/or removed	Osteophytes correctly identified and removed	Not assessed
18	PS	Demonstrated appropriate synovectomy	Did not correctly complete synovectomy	Performed complete synovectomy, proper resection	Not assessed
19	PS	Removed medial & lateral menisci	Did not do resection, incomplete resection, resected too much tissue	Adequate meniscal resection, other soft tissue preserved	Not assessed
Femoral Preparation					
20	PS	Correctly identified appropriate start point for intramedullary guide	Site not identified correctly	Site identified correctly with appropriate landmarks	Not assessed
21	PS	Selected, placed, and aligned femoral cutting guide	Selected an inappropriate guide, incorrectly placed guide, did not check alignment, did not protect ligaments	Selected appropriate size of guide, appropriately placed guide, checked alignment, protected ligaments	Not assessed
22	PS	Demonstrated appropriate distal femur cutting	Cut was not made safely, used inappropriate rotation, did not check external rotation	Soft tissue protected, used appropriate rotation, checked external rotation	Not assessed
23	PS	Completed posterior capsule release and osteophyte removal, if necessary	Release not completed, osteophytes not removed, other soft tissue compromised	Appropriate posterior release completed with osteophyte removal	Not assessed

Tibial Preparation					
24	**PS**	Demonstrated safe knee dislocation	Did not perform safe dislocation	Performed safe dislocation	Not assessed
25	**PS**	Placed tibial cutting guide	Knee not maximally flexed, alignment not correct	Guide placed in appropriate alignment, with knee maximally flexed	Not assessed
26	**PS**	Performed tibial cuts	Did not perform cut safely, did not protect patellar tendon	Performed cut safely and protected patellar tendon	Not assessed
		Competency	**In Training (T)**	**Competent (C)**	**Not Assessed (N/A)**
27	**PS**	Prepared and sized tibia	Did not remove osteophytes, protect soft tissues, or select appropriate template	Removed osteophytes protected soft tissues, selected appropriate template	Not assessed
Balancing the Knee					
28	**PS**	Balanced soft tissue	Did not remove osteophytes or release ligaments and soft tissues appropriately	Removed osteophytes and released ligaments and soft tissues appropriately	Not assessed
29	**PS**	Assessed flexion and extension gap	Did not assess, or incorrectly assessed release	Correctly assessed and identified appropriate release	Not assessed
30	**PS**	Demonstrated appropriate trialing	Selected an inappropriate trial, placed trial incorrectly, did not mark template locations on tibia, made inappropriate intramedullary cut, did not consider drilling tibia for interdigitation of cement, did not check or recognize instability	Selected an appropriate trial, placed trial correctly, marked template locations on tibia, made an appropriate intramedullary cut, considered drilling tibia for interdigitation of cement, checked for instability	Not assessed
31	**PS**	Confirmed position, size and rotation	Did not confirm or incorrectly confirmed	Correctly confirmed position, size and rotation	Not assessed
32	**PS**	Assessed patellar tracking	Did not assess, did not approximate the medial parapatellar arthrotomy temporarily while tracking	Assessed patellar tracking with appropriate rotation and approximation of arthrotomy	Not assessed
33	**PS**	Sized patella with cutting guide; cut and drilled patella	Did not measure width/height, osteophytes not removed	Correctly sized, cut & drilled patella, osteophytes removed	Not assessed
34	**PS**	Prepared bones for cement	Did not prepare bones, incorrect preparation	Correctly prepared bones for cement	Not assessed

35	**PS**	Mixed cement appropriately	Mixed cement incorrectly	Correctly mixed cement to the appropriate viscosity	Not assessed
36	**PS**	Cemented prosthesis appropriately	Used an inappropriate amount of cement, did not remove excess cement, did not ensure proper alignment of hardware	Placed appropriate amount of cement, removed excess cement, ensured proper alignment of hardware	Not assessed
Wound Closure					
37	**PS**	Demonstrated appropriate wound closure	Did not ensure proper alignment before final closure, approximated skin too tightly, did not correctly complete layered closure, did not cover with sterile dressing	Ranged knee prior to final closure, used appropriate skin tension, completed appropriate layered closures, covered with sterile dressing	Not assessed

Global Ratings of Performance

Technical Skill	**1** - Handled instruments and/or soft tissues poorly - Demonstrated improper technique	2	**3** - Handled instruments and/or soft tissues adequately - Demonstrated adequate technique	4	**5** - Handled instruments and/or soft tissues competently - Demonstrated proper technique
Surgical Judgment	**1** - Unable to anticipate errors - Did not correct errors	2	**3** - Anticipated some errors - Able to correct some errors	4	**5** - Anticipated errors - Corrected errors with ease
Operative Flow	**1** - Many unnecessary movements - Inefficient - Stopped frequently - Often unsure of next move	2	**3** - Some unnecessary movements - Some efficiency - Stopped sometimes - Demonstrated some forward planning	4	**5** - No unnecessary movements - Highly efficient - Effortless flow - No hesitation
Professionalism	**1** - Arrived late - Unprepared - Did not display respect for OR personnel -Some inappropriate behaviour	2	**3** - Demonstrated some preparation - Displayed respect for OR personnel most of the time - Appropriate behaviour	4	**5** - Arrived on time - Clearly prepared for the case - Displayed respect for all OR personnel throughout case - Exemplary behaviour
Operative Awareness	**1** - Unaware of operative surroundings and team	2	**3** - Somewhat aware of operative surroundings and team	4	**5** - Fully aware of operative surroundings and team
Overall Entrustment	**1** - Required extensive guidance - Was not able to actively participate	2	**3** - Demonstrated some independence - Required some direction	4	**5** - Understood risks and performed procedure safely - Ready for independent practice

Comments *(e.g. one thing to improve):*

APPENDIX D. PHASE FOUR EXAMPLE TOOL

Procedure/Task: **Primary Total Knee Arthroplasty**	
Resident Name:	Stage of training (PGY):
Date:	Assessor:
Role (circle one): Primary / 1st Assist	

COMPLEXITY OF CASE:

Simple Moderate High

NOTES:

EVALUATION OF SURGICAL INDEPENDENCE:

	Not Assessed	In Training				Competent
PREOPERATIVE PLANNING						
1. Identified appropriate surgical indication	N/A	I had to do	I had to talk them through	I needed to prompt	I needed to be there just in case	I didn't need to be there
2. Appropriate preoperative planning (includes review of appropriate films)	N/A	I had to do	I had to talk them through	I needed to prompt	I needed to be there just in case	I didn't need to be there
3. Reviewed patient and initialed appropriate side	N/A	I had to do	I had to talk them through	I needed to prompt	I needed to be there just in case	I didn't need to be there
4. Demonstrated appropriate preoperative templating	N/A	I had to do	I had to talk them through	I needed to prompt	I needed to be there just in case	I didn't need to be there
5. Selected appropriate implant	N/A	I had to do	I had to talk them through	I needed to prompt	I needed to be there just in case	I didn't need to be there
6. Confirmed all necessary equipment	N/A	I had to do	I had to talk them through	I needed to prompt	I needed to be there just in case	I didn't need to be there
	Not Assessed	**In Training**				**Competent**
7. Completed preoperative checklist	N/A	I had to do	I had to talk them through	I needed to prompt	I needed to be there just in case	I didn't need to be there
8. Maintained appropriate sterile technique throughout procedure (i.e. gowning, gloving, prepping & draping)	N/A	I had to do	I had to talk them through	I needed to prompt	I needed to be there just in case	I didn't need to be there
9. Demonstrated appropriate patient positioning	N/A	I had to do	I had to talk them through	I needed to prompt	I needed to be there just in case	I didn't need to be there
OPERATIVE TECHNIQUE						
10. Appropriate anatomical landmark identification	N/A	I had to do	I had to talk them through	I needed to prompt	I needed to be there just in case	I didn't need to be there
11. Completed an appropriate midline incision (i.e. used the medial parapatellar approach to create a midline incision, anterior to the knee, with appropriate extensile exposure if necessary)	N/A	I had to do	I had to talk them through	I needed to prompt	I needed to be there just in case	I didn't need to be there
12. Demonstrated a proper medial release for exposure, MCL preserved	N/A	I had to do	I had to talk them through	I needed to prompt	I needed to be there just in case	I didn't need to be there
13. Exposed distal femur, everted/subluxed the patella appropriately	N/A	I had to do	I had to talk them through	I needed to prompt	I needed to be there just in case	I didn't need to be there
14. Performed arthrotomy, properly preserved patellar tendon and ligament	N/A	I had to do	I had to talk them through	I needed to prompt	I needed to be there just in case	I didn't need to be there
15. Removed patellar fat pad, patellar tendon preserved with adequate resection	N/A	I had to do	I had to talk them through	I needed to prompt	I needed to be there just in case	I didn't need to be there
16. Demonstrated appropriate & safe use of retractors	N/A	I had to do	I had to talk them through	I needed to prompt	I needed to be there just in case	I didn't need to be there

	Not Assessed	In Training				Competent
17. Correctly identified and removed osteophytes	N/A	I had to do	I had to talk them through	I needed to prompt	I needed to be there just in case	I didn't need to be there
18. Appropriately performed synovectomy	N/A	I had to do	I had to talk them through	I needed to prompt	I needed to be there just in case	I didn't need to be there
19. Removed medial & lateral menisci	N/A	I had to do	I had to talk them through	I needed to prompt	I needed to be there just in case	I didn't need to be there
Femoral Preparation						
20. Correctly identified appropriate start point for intramedullary guide	N/A	I had to do	I had to talk them through	I needed to prompt	I needed to be there just in case	I didn't need to be there
21. Appropriately selected, placed, and aligned femoral cutting guide	N/A	I had to do	I had to talk them through	I needed to prompt	I needed to be there just in case	I didn't need to be there
22. Demonstrated appropriate distal femur cutting (i.e. soft tissue protected, used appropriate rotation, checked external rotation)	N/A	I had to do	I had to talk them through	I needed to prompt	I needed to be there just in case	I didn't need to be there
23. Completed posterior capsule release and osteophyte removal (if necessary)	N/A	I had to do	I had to talk them through	I needed to prompt	I needed to be there just in case	I didn't need to be there
Tibial Preparation						
24. Demonstrated safe knee dislocation	N/A	I had to do	I had to talk them through	I needed to prompt	I needed to be there just in case	I didn't need to be there
25. Appropriately selected, placed, and aligned tibial cutting guide with knee maximally flexed	N/A	I had to do	I had to talk them through	I needed to prompt	I needed to be there just in case	I didn't need to be there
26. Demonstrated appropriate distal tibial cutting (i.e. protected patellar tendon)	N/A	I had to do	I had to talk them through	I needed to prompt	I needed to be there just in case	I didn't need to be there
27. Prepared and sized tibia	N/A	I had to do	I had to talk them through	I needed to prompt	I needed to be there just in case	I didn't need to be there

	Not Assessed	In Training				Competent
Balancing the Knee						
28. Balanced soft tissue (i.e. removed osteophytes and released ligaments and soft tissues appropriately)	N/A	I had to do	I had to talk them through	I needed to prompt	I needed to be there just in case	I didn't need to be there
29. Correctly assessed flexion and extension gap	N/A	I had to do	I had to talk them through	I needed to prompt	I needed to be there just in case	I didn't need to be there
30. Demonstrated appropriate trialing (i.e. selected an appropriate trial, placed trial correctly, marked template locations on tibia, made an appropriate intramedullary cut, considered drilling tibia for interdigitation of cement, checked for instability)	N/A	I had to do	I had to talk them through	I needed to prompt	I needed to be there just in case	I didn't need to be there
31. Correctly confirmed position, size and rotation	N/A	I had to do	I had to talk them through	I needed to prompt	I needed to be there just in case	I didn't need to be there
32. Assessed patellar tracking with appropriate rotation and approximation of arthrotomy	N/A	I had to do	I had to talk them through	I needed to prompt	I needed to be there just in case	I didn't need to be there
33. Correctly sized, cut & drilled patella, osteophytes removed	N/A	I had to do	I had to talk them through	I needed to prompt	I needed to be there just in case	I didn't need to be there
34. Correctly prepared bones for cement	N/A	I had to do	I had to talk them through	I needed to prompt	I needed to be there just in case	I didn't need to be there
35. Correctly mixed cement to the appropriate viscosity	N/A	I had to do	I had to talk them through	I needed to prompt	I needed to be there just in case	I didn't need to be there
36. Appropriately cemented prosthesis (i.e. placed appropriate amount of cement, removed excess cement, ensured proper alignment of hardware)	N/A	I had to do	I had to talk them through	I needed to prompt	I needed to be there just in case	I didn't need to be there

	Not Assessed	In Training				Competent
Wound Closure						
37. Demonstrated appropriate wound closure (i.e. angled knee prior to final closure, used appropriate skin tension, completed appropriate layered closures, covered with sterile dressing)	N/A	I had to do	I had to talk them through	I needed to prompt	I needed to be there just in case	I didn't need to be there
POSTOPERATIVE CARE						
38. Appropriate postoperative orders completed and documented (i.e. medications, activity levels, and postoperative plan)	N/A	I had to do	I had to talk them through	I needed to prompt	I needed to be there just in case	I didn't need to be there
39. Assessed and documented neurovascular function	N/A	I had to do	I had to talk them through	I needed to prompt	I needed to be there just in case	I didn't need to be there

NARRATIVE FEEDBACK: Give at least 1 specific suggestion the resident should work on improving

OVERALL PERFORMANCE: In regards to the assessed activity, the resident (circle one)

GLOBAL RATING OF ENTRUSTMENT				
1 Unable to perform task; supervisor had to take over	**2** Required constant supervision/ prompting to complete task	**3** Required some supervision/ prompting to complete task	**4** Required minimal supervision/ prompting to complete task	**5** Ready to perform task independently

Expectations:

ONE evaluation must be submitted by each resident EVERY WEEK
Please deposit forms into a locked evaluation box located in one of the following areas:
MUMC – Fracture Clinic
HGH – OR Lounge
SJHH – OR Lounge
JURV – OR Lounge