Self-Assessment Workshop

Technical Brief

Overview

Ten teams of innovators—each comprised of surgical practitioners, educators and technologists—from multiple countries are currently competing in the Global Surgical Training Challenge to develop novel surgical training modules for low resource settings. A key element of these training modules is the requirement for a self-assessment component, allowing surgical practitioners to train and assess themselves in obtaining new psychomotor and procedural decision making skills.

A workshop, which took place on 17 June 2021, was designed to help Discovery Award teams improve their understanding of self-assessment principles. Participants learned from experts in surgical education and assessment by working with them in small groups on case studies relevant to their work on the Global Surgical Training Challenge.

The modules being created significantly decrease the need for instructors to train surgical practitioners in the basic and core psychomotor skills. The ability to self-assess and monitor one's own skill acquisition becomes the core source of value to those who will use the surgical training modules. Tying the modules to an assessment is what distinguishes the prototypes from being simply an experience or an exercise. Without an assessment, there is no way to measure what skills the practitioners have obtained. Assessment drives the training process.

The purpose of this technical brief is to provide a high level summary of two of the presentations given to the teams.

The Importance of Self Assessment

Presented by Prof. Carla Pugh, Stanford University

Prof. Carla Pugh is Professor of Surgery at Stanford University School of Medicine and sits on the judging panel for the Global Surgical Training Challenge. She is also the Director of the Technology Enabled Clinical Improvement (T.E.C.I.) Center. In addition to her medical degree, Dr Pugh is the first surgeon in the United States to obtain a PhD in Education. Her goal is to use technology to change the face of medical and surgical education. Her research involves the use

of simulation and advanced engineering technologies to develop new approaches for assessing and defining competency and mastery in clinical procedural skills.

Her presentation focused on describing the purpose and value of self-assessment in surgical training. She set context for the role of self-assessment within the Global Surgical Training Challenge, and she outlined the key principles and frameworks for self-assessment.

Self-assessment minimizes variability.

Self-assessment provides lasting value because it allows surgical trainees and practitioners to **evaluate their performance against an objective standard**. This motivates learners, provides personal goals, and encourages a standard of excellence.

It also minimizes variability in techniques. Learners will be able to detect variations in technique. While there are a variety of factors that influence a patient's outcome, the variability of a surgeon's technique can correlate strongly to outcomes. Self-assessment can identify and correct this variability so they can implement the best techniques.

Feedback is the key principle of self-assessment.

The single most important feature of self-assessment is feedback. In order to be effective, feedback must be:

- Given in a **timely** manner — providing immediate results allows learners to address the questions or issues in their mind before the "teachable moment" is forgotten or passed.
- **Specific** to the learner's individual needs — learners should be able to compare their performance to others. It should also be tailored to the skill level of the learner. Practicing surgeons may be more interested in mastery, which compares them to top performers. Surgical trainees, such as residents, may seek to measure competency against the performance of their peers.
- **Comparative** to an explicit standard — the goals should be made explicit before engaging in the simulation, providing metrics, report card or dashboard.

Self-assessment is distinct within the learning environment.

No matter how realistic the surgical model the learner is training on, without an assessment it is primarily an experience or general activity that does not measure the skills obtained.

Self-assessment is also distinct from teaching. Teaching relies on providing instruction, verbal guidance or directions. Self-assessment is based on immediate feedback on a simulator and

offers the ability of the learner to correct themselves. Students should end their assessment with a report card.

Self-assessment should measure both technical and cognitive skills.

Surgical training modules should be holistic in their approach. Self-assessment should cover both surgical technique, or psycho-motor skills, and the surgical practioner's decision-making skills.

Cognitive skills are based on the ability to plan a procedure based on diagnostic information. Intraoperative decision making includes responses in an unexpected event, a crisis, or emergency.

There are different types of assessment.

- **Formative —** done as the learner is learning (for example, midway through a course) — that guides the learning. The goal is to improve student achievement and guide learning.
- **Summative** — done as a final exam to set minimum performance criteria. It is a way to measure preparedness for advancing to the next level of training.

There is a difference between **OSATS** (objective structured assessment of technical skills) vs. **Procedure-Specific.** You should do both. However, OSATS may have shortcomings. In studies that Prof. Pugh and her team conducted, students who had high OSATS scores were often unable to complete a procedure or had an incorrect result in the procedure-specific assessment. In other words, **traditional checklist metrics for surgical performance can miss key intraoperative decisions that impact procedural outcomes**. Being holistic in the assessment is strongly recommended.

Self-assessment can be based on different theoretical frameworks.

- **Cognitive Task Analysis** yields a description of the learning objectives, equipment, conceptual and procedural knowledge, and performance standards used by experts.
- **Naturalistic Decision Theory** focuses on how people make decisions and perform cognitively complex functions in dynamic environments.
- **Motor theory learning** measures every skill and requires cognitive and motor input.

Behaviors can be measured based on skills, rules, or knowledge. Looking at behavior from these different perspectives can improve the quality of the assessment.

Creating Effective Assessment Items

Presented by Prof. Richard Arnett, Director of Psychometrics, Royal College of Surgeons in Ireland

Professor Arnett shared a series of guidelines for developing multiple choice questions, or assessment items, that can increase the validity of assessment against the goals of the training module. During the presentation, he emphasized the importance of timely feedback and creating a mechanism for learners to move from one level of learning to the next.

Define the purpose of the assessment

Before developing the assessment items, teams should have a clear consensus on the purpose of the assessment. It is helpful to document responses to the following questions to ensure a common understanding. This documentation facilitates the development of the assessment.

- What are you testing?
    - Bloom's Taxonomy refers to six major categories for setting educational goals: knowledge, understanding, application, analysis, synthesis, evaluation. Testing knowledge is easier, but it becomes progressively more difficult as you progress through the categories.
- Why is it important?
    - Teams should have a clear understanding of why the test objectives are important.
- How and when do you assess?
    - The team also needs to document how and when the assessment will occur. Sometimes it is helpful to have shorter assessments dispersed throughout the module to promote ongoing learning rather than one major assessment at the end of the learning.
- Who will be using the module?
    - Understanding your intended audience will help you target the level at which the assessment should occur.

Avoid common mistakes in developing self assessment items

Best practices for crafting good questions and options in multiple choice questions (MCQs) are based on both context-setting and the actual framing of the question.

Provide context
Questions should be framed within an appropriate context. For example, an anatomical question might be framed in the context of a case study that presents a patient. It is important to provide sufficient relevant information for an informed response. This also creates a “real world” scenario that helps assess how the learner might apply the learning.

Frame questions appropriately
Good questions have common structures. For example, they are phrased as questions rather than statements.

They are positive rather than deductive. In other words, they do not ask “which of the following are NOT….”).

Offer clear options for responses
If the learner is selecting an answer from a series of options, the choices must be very clear and succinct.

The correct answer should be unambiguously correct, or stand out as the best option out of several possible responses. This is particularly important in ambiguous settings, such as medicine, when there may be several possibilities for a diagnosis.

Also, avoid using vague terminology, such as “frequently” or “rarely.” Be specific. Likewise, avoid absolute terms, such as “always” or “never.”

Avoid clues
Learners should be able to get the correct answer based solely on their knowledge rather than any clues you leave.

Review your slate of optional answers and remove any that are obviously unrealistic choices. For example, if the question refers to an “oral administration” of a particular medication, be sure that all the medications or treatments listed are available in oral formulation.

Ensure that the options are generally the same length — longer answers tend to be the correct ones due to their clarity.

Gauge the responses
Particularly during the piloting phase, look for responses and evaluate the items that either no one is getting correct or everyone is getting correct. This may provide important feedback on the structure of the question or phrasing of the responses.