

How do I improve the quality of in-training assessment of learners?

Glen Bandiera, MD, MEd

ABSTRACT

Assessing a learner in the course of a hectic emergency department (ED) rotation is a daunting task for both experienced and new supervisors. This is particularly true if the learner is not doing well. In light of numerous impediments provided by the modern ED environment, sticking to basic principles can result in marked improvement in both the process and the outcome of in-training assessment. This article addresses these important principles for assessment as they apply in the clinical realm of the ED, with a focus on matching expectations to both the trainee and the available assessment strategies. It is critical that teachers strive for clarity, consistency, honesty, and adherence to due process in their learner assessments. This article provides an evidence-informed approach to succeeding with such an approach to clinical assessment.

RÉSUMÉ

L'évaluation d'un apprenant au cours d'une rotation frénétique du personnel au service des urgences (SU) est une tâche ardue pour les superviseurs expérimentés comme pour les nouveaux superviseurs. Cela est surtout vrai lorsque l'apprenant ne réussit pas bien. À la lumière de nombreux obstacles que présente l'environnement moderne du SU, le fait de s'en tenir à des principes de base peut mener à une amélioration marquée à la fois du processus et du résultat de l'évaluation de l'apprenant en formation. Cet article examine ces importants principes d'évaluation tels qu'ils s'appliquent dans le domaine clinique du SU, en mettant l'accent sur l'adéquation entre les attentes de l'apprenant en formation et les stratégies d'évaluation disponibles. Il est essentiel que les enseignants s'efforcent d'être clairs, cohérents, honnêtes et qu'ils respectent la procédure régulière lorsqu'ils procèdent à l'évaluation des apprenants. Cet article propose une approche factuelle pour réussir avec une telle approche d'évaluation clinique.

Keywords: assessment, clinical supervision, competency-based education, education, evaluation, feedback, performance

INTRODUCTION

The emergency department (ED) is a particularly complex environment in which to assess learners but one that provides rich opportunities to do it well. Challenges of ED life that specifically complicate learner assessment include varied learner backgrounds and aspirations, unpredictable patient mix and volume, hectic work pace, multiple interruptions, and an inability to plan educational and assessment activities for a given shift. Thankfully, numerous advantages more than mitigate these problems, including a diversity of patients, a diversity of clinical activities such as procedures and resuscitations, the opportunity to observe practice from patient interview through to disposition, the option of seeking input from others working with a learner, and the opportunity to observe unique skills such as time management, teaching, and administrative duties. Additionally, assessment “tools” such as end-of-rotation forms are often designed by central program administrators (e.g., clerkship coordinators, residency directors).

Despite this external constraint and the aforementioned ED pitfalls, there are simple ways individual teachers can improve their assessment of learners. Effective assessment is all about matchmaking, specifically ensuring that the teacher's and the learner's expectations align, assessment tools match the competency

From the Department of Emergency Services, St. Michael's Hospital, Toronto, ON.

Correspondence to: Dr. Glen Bandiera, Department of Emergency Services, St. Michael's Hospital, 30 Bond Street, Suite 1-008e Shuter Wing, Toronto, ON M5B 1W8; bandierag@smh.ca.

Submitted October 4, 2010; Revised January 31, 2011; Accepted February 16, 2011.

© Canadian Association of Emergency Physicians

CJEM 2011;13(4):267-272

DOI 10.2310/8000.2011.110398

being assessed, the teacher's opinion is consistent with available evidence, messages to learners are accurate and consistent, and final summative assessments match prior feedback. This article addresses these principles in turn as they apply in the ED. For a more general review of in-training assessment, readers are directed to two excellent reviews.^{1,2} Although the focus here is on assessment in real-time rotations, teachers should also feel empowered to communicate concerns about the overall system to program leaders.

CLINICAL SCENARIO

You are supervising a first-year internal medicine resident on a 1-month ED rotation in a community hospital. This is a new rotation as a nearby university is expanding to community sites. You have a set of objectives for the rotation, but this is the first time you have supervised a resident. After 1 week, it becomes clear to you that the resident is performing below expectations, particularly in the areas of efficiency and ability to prioritize a differential diagnosis. You feel that this is not compatible with satisfactory completion of the rotation and are unsure what to do.

MATCHING EXPECTATIONS

Good teachers go beyond generic rotation objectives and tailor explicit goals and expectations for each learner.^{3,4} It is hard to predict learners' objectives based simply on their program or specialty because learners self-select objectives based also on their perceived personal needs, previous experience, and what they feel the rotation can offer.^{5,6} Teachers must be active participants in setting objectives and establish their personal expectations related to their work and teaching style. For example, some teachers prefer that residents present a brief synopsis of each case starting with their impression and plan, whereas others prefer a systematic, comprehensive case presentation starting with the history and system review.³ Learners should not need to guess about a supervisor's preferred work style. This can be a complicated step, and developing effective learning objectives is the focus of another article in this series. Once established, teachers should relate all subsequent assessment (especially pass/fail and other formal high-stakes decisions) on mutually agreed on objectives and expectations. In the above scenario, residents may be unaware that the expected

ED approach is more concise and goal driven than in an internal medicine clinic. They may fear appearing "dumb" if they have not researched on the Web every case encounter before presenting it. These problems (and related misperceptions) could be avoided with an explicit discussion of expectations.

MATCHING OBJECTIVES WITH EVIDENCE

Most teachers rely on observation in real time to gather data that eventually forms the basis for an assessment. This sounds easy, but good teachers work hard to match their assessments with appropriate objective evidence. For example, one can be sure that a learner knows certain facts only by direct questioning or formal tests of knowledge. Teachers should not infer that learners know a critical fact because they happened to demonstrate a behaviour consistent with this knowledge (maybe they got lucky!). Thus, questioning and brief oral or written examinations are appropriate for assessing knowledge. To determine if learners can perform a task, they should be observed performing the task. One should not assume that a learner is adept at all aspects of a procedure just because the immediate outcome seems fine (did they really use appropriate sterile technique for that chest tube and did they really discuss all aspects of informed consent?). The mechanics of procedures can also be observed using simulated models. Communication skills can be observed or assessed through role modeling ("You haven't had a chance to give bad news yet. Why don't you treat me like a parent whose child has just died and walk me through how you would approach this conversation?"). Teachers might want to involve other health care professionals in their assessments. Nurses and social workers, for example, can have insightful comments about a resident's collaborative abilities; students can comment on teaching. This resembles, but is not the same as, a formal 360° assessment or a multisource feedback (these processes require appropriate rater training and validated tools). Rather, teachers should either personally validate the opinions of others or acknowledge that they are passing on others' perceptions as part of the assessment when communicating with the learner. Teachers should keep track of specific examples of performance (cases or discussions/questions) to help recognize patterns, justify an ensuing assessment, and help the learner establish context.

Studies have demonstrated a general lack of direct observation of learners during care provision.⁷ This is a problem. Although it is clearly unrealistic and inappropriate to observe all learner activities, selected observations are important to underpin eventual assessment. Evidence suggests that at least seven observations are required for a reliable assessment of each skill.⁸ For practicality, teachers should reserve such intense focus for critical skills or for those they have reason to believe the learner is deficient in. For example, clinical clerks should be observed doing several patient assessments. Senior learners should not require such observation unless their case presentations betray possible deficiencies.

Many assessment tools and processes have been developed to help assessors in their work, and a thorough review of their strengths and weaknesses is beyond the scope of this chapter (see Bandiera and colleagues in the Annotated Bibliography section). Some basic tools deserve special mention, however, for their ability to facilitate performance-based assessment in the ED environment. Daily encounter cards (DECs) are brief comment cards deployed at the end of a defined experience on which a teacher documents his or her impression of the learner's performance. The design of the cards varies widely and can include numerical ratings, qualitative comments, or a combination of the two. Ideally, the cards would be designed around understood objectives and/or an accepted educational competency-based model such as CanMEDS. The strength of this tool is that it encourages the provision of regular, documented, formalized assessment to the learner that is timely, contextual, and based on recent behaviours. Properly deployed, DECs can alert rotation supervisors to developing patterns of behaviour and provide backup documentation to remediation or other educational interventions undertaken to improve learner performance.

The Mini Clinical Evaluative Examination (Mini-CEX) is a specific clinical experience that is defined in advance and observed as an explicit assessment opportunity. It is usually embedded in a clinical experience. Learners are observed undertaking a specific task related to a specific area of competence (such as the counseling of a patient around postexposure prophylaxis or, in an example pertinent to the clinical scenario that opened this chapter, the acquiring of a history from a talkative patient). The assessment relates to a defined competency area, occurs at a

defined period of time, and is assessed with the help of a validated instrument developed for this purpose. The Mini-CEX is typically deployed for all learners as a formal assessment strategy integral to a rotation by a program director or rotation coordinator but can be used as an ad hoc tool for a given learner to further assess a specific area of concern.

A newer tool that resembles the Mini-CEX is the Standardized Assessment of Clinical (STACER). This tool is typically designed around very specific competencies and, owing to the focus on standardization (to enable high-stakes decisions to be made), often makes use of a simulated, well-scripted scenario. This does not preclude the use of a STACER-like tool in the clinical ED, however; certain scenarios are frequent and standard enough to lend themselves to this modality (e.g., simple sutures), or a simulated patient encounter can be embedded into an ED shift to address a certain learner assessment need. As with the Mini-CEX, assessment of the performance must be done with the assistance of a validated data recording instrument such as a numerical scale or global assessment form.

Finally, in a variation on the DEC, input from various nonphysicians can form the basis of an abbreviated multisource assessment. Soliciting input from trained members of the health care team who are well-versed in what to look for and what expected levels of performance are for a given level of trainee can shed significant insights into learner performance beyond what a supervisor might be able to observe. It is important that those called on to provide input are seasoned, well trained, and in a position to observe and comment on the behaviour of interest (e.g., a nurse might be called on to comment on a resident's collaboration or communication skills but should not be asked to assess procedural competence on a complex wound repair).

In our scenario, the ability to develop a differential diagnosis depends on both knowledge and problem solving. The former is easily tested through a brief written examination ("Write down the differential of sudden painless loss of vision"), whereas the other requires some sophisticated questioning ("Why did you include or not include that item? Which do you think are most likely and why?").

MATCHING EVIDENCE TO AN OPINION

When based on direct observation and related to established expectations, matching evidence to an

eventual opinion is relatively easy. Teachers must decide what aspects (both good and bad) of the learner's performance should be addressed and clarify for themselves why they believe this to be the case. They must then relate their impression to external criteria (what the objectives say a learner of this type should be able to accomplish), the standard of the learner's peer group ("this person is better, the same, or worse in this skill than most other similar individuals I have supervised"), and the teacher's own norms and expectations. Assessing a learner's performance against an external objective measure is referred to as criterion referencing. All learners who attain a level of performance commensurate with the external measure are given credit.

Assessing a learner's performance against his or her peer group is referred to as norm referencing. Teachers should be careful to distinguish the two because even though a learner is significantly "behind" his or her peers, he or she may still demonstrate adequate performance to meet external objectives for the rotation. Conversely, even though a learner is performing better than his or her peers, this is no guarantee that he or she (or any member of their cohort) is actually meeting expectations. Furthermore, teachers should always be attentive to the tendency to shift their expectations over time based on their cumulative experience with learners. A particularly strong cohort of learners or a sudden influx of learners with a different background can result in a significant shift of the teacher's expectations for subsequent learners even though the actual objectives and performance expectations for the rotation have not changed.

As mentioned above, a brief review of expectations at the start of a clinical experience can help both teacher and learner remain focused on the actual intended and mutually agreed on expectations for the experience. Teachers should do their best to personally corroborate evidence communicated by others. Discussion of a learner's performance with other supervising staff can help the teacher fully understand observed behaviours, obtain further examples, and provide external validation of his or her opinion. In our scenario, perceived inefficiency could result from prolonged or unfocused patient assessments, leisurely pace, inability to delegate, excessive extraneous chatter, and excessive reading around cases. The teacher, having identified the apparent "symptom" (inefficiency), must now do some "diagnosis" to determine how to help the resident improve.

COMMUNICATING A CONSISTENT MESSAGE

Assessment is an ongoing process, and opinions must be regularly communicated to the learner. Such communication, or feedback, can be viewed as either formative or summative assessment. Teachers must distinguish between the two. Formative assessment is information provided to learners based on their performance that is intended to guide them toward better practice. It is intended to be part of an ongoing information exchange that moulds new behaviour. Ideally, formative assessment and subsequent feedback will be given during a shift, at the end of a shift, and more formally at a fixed point during the rotation. Formative assessment should identify strengths and weaknesses. It should be explicit, timely, and based on actual behaviours. It should include examples and provide the recipient with an opportunity to reflect on, comment on, and commit to a course of change. Appropriate feedback is important to ensure ongoing progress, that goals and expectations are clear, and that learners have every opportunity to adjust their performance as required. A learner should never be surprised by a subsequent summative assessment if feedback based on formative assessment has been adeptly given along the way. This is particularly important for learners in difficulty because they need to be provided with explicit concerns and constructive assistance in addressing these. Furthermore, clearly communicated and documented assessment is required as due process should a learner not meet the requirements for rotation completion. Summative assessment is usually more formal, includes some judgment of the performance based on some external criteria or structure (such as a centrally derived assessment form), provides a summary perspective on a defined period of experience, and includes some provision for rating the learner.

Finally, one of the most self-evident, yet difficult, principles of assessment to implement is honesty! Nobody wants to be the bearer of bad news, but avoiding provision of an honest opinion does not serve the learner (or the program director or the public) well and is an abdication of responsibility. There are numerous reasons why teachers are reluctant to provide honest (particularly if negative) assessments, including fear of reprisal, lack of confidence in their ability to provide feedback properly, fear of further work, and social desirability bias toward being "nice." In our scenario, the resident must be provided with the concerns (efficiency), the supporting

Table 1. Common pitfalls in learner assessment

Step	Pitfall	Potential solution(s)
Matching expectations	Assuming agreement on objectives and performance	Insist on review of objectives with each learner Verify understanding Use previous experience to anticipate common misunderstandings Ask learners about their experiences/expectations Aim to help each learner to their next level
	Treating all learners the same	
Matching objectives with evidence	Using the wrong tools	Assess knowledge with questions (oral or written) Observe procedures and patient interactions as much as possible Assess behaviours rather than inferred personality/attitudinal traits
	Making assumptions	
Matching evidence to an opinion	Poor feedback practices (communication of formative assessment)	Give feedback as soon as possible Base comments on observed behaviour Avoid making assumptions about character or motives Balance "good" and "bad" comments Give examples; refer to actual events Allow time for reflection (but not necessarily rebuttal) Get a commitment to action (e.g., change, follow-up)
	Misunderstanding of references	For criterion-referenced systems, regularly review the meaning of external measures (e.g., what constitutes a "fail") For norm referencing, base comparisons on reasonable peer groups
Communicating a consistent message	Being a poor role model	Ensure you demonstrate through actions and words what you wish learners to learn
	Inconsistent expectations	Hold all trainees at a similar level to a similar standard
Consistent summative assessments	End-of-rotation "surprises"	Include nothing new in summative assessment that was not previously discussed with an opportunity for the learner to address

evidence ("You saw half the number of cases your colleagues did over the last week, and your average time per case is 60 minutes"), and supportive commentary ("I would like to see you better focus your review of systems for straightforward cases").

CONSISTENT SUMMATIVE ASSESSMENTS

If previous steps are carried out well, summative assessments are straightforward. These are usually completed on a predetermined form, and teachers should adhere to proven principles for effective form completion. Evidence again suggests that documented final assessment is not always consistent with the assessor's judgment of the performance, especially for the poorly performing resident.^{10,11} Often this is simply due to the lack of knowledge about how to complete the form.⁹ Recent research has demonstrated that appropriate form completion (regardless of the design of the form) includes a balance of strengths and areas for improvement with sufficient variability in numerical ratings to clearly demonstrate the two, a description of the learner's response to feedback, comments to justify or explain the numerical ratings, clear examples of both strengths and weaknesses, recommendations for improvement, comments that are supportive and relevant to objectives, and provision of enough detail

for an independent reader to understand the learner's performance.¹² In the case scenario, if the resident still had difficulty at the end of the rotation, they might get a "below expectation" rating on "efficiency" and a comment such as "Resident improved during rotation but still needs to better adapt system review to presenting problem to improve efficiency."

CONCLUSION

To aid in the application of the aforementioned principles, Table 1 contains a summary of common pitfalls and strategies for success. Teachers can optimize their learner assessments regardless of the system in which they work by adhering to simple principles. Proper matchmaking is the key, and satisfying results can emerge with a little focus on consistency and objectivity.

Competing interests: None declared.

REFERENCES

1. Turnbull J, Van Barneveld C. Assessment of clinical performance: in-training evaluation. In: Norman GR, Van der Vleuten CPM, Newble DI, editors. *International handbook of research in medical education*. London: Kluwer Academic Publishers; 2002. p. 793-810.

2. Turnbull J, Gray J, MacFadyen J. Improving in-training evaluation programs. *J Gen Intern Med* 1998;13:317-23.
3. Bandiera GW, Lee S, Tiberius R. Effective teaching in the emergency department: how effective teachers get it done. *Ann Emerg Med* 2005;43:253-61.
4. Thurgur L, Bandiera G, Lee S, et al. What emergency medicine learners wish their teachers knew. *Acad Emerg Med* 2005;12:856-61.
5. Tenn-Lyn NA, Leblanc VA, Bandiera G. Can we predict what objectives off-service residents have for their emergency medicine rotations? *Ann Emerg Med* 2008;51:516.
6. Tenn-Lyn NA, Bandiera G, Hodges B, et al. Factors influencing self-directed objective setting by off-service residents in emergency medicine. *Ann Emerg Med* 2008;51:504.
7. Pulito AR, Donnelly MB, Plymale M, et al. What do faculty observe of medical students' clinical performance? *Teach Learn Med* 2006;18:99-104.
8. Carline JD, Paauw DS, Thiede KW, et al. Factors affecting the reliability of ratings of students' clinical skills in a medicine clerkship. *J Gen Intern Med* 1992;7:506-10.
9. Dudek NL, Marks MB, Regehr G. Failure to fail—the perspectives of clinical supervisors. *Acad Med* 2005;80(10 Suppl):S84-7.
10. Cohen GS, Blumberg P, Ryan NC, et al. Do final grades reflect written qualitative evaluations of student performance? *Teach Learn Med* 1993;5:10-5.
11. Hatala R, Norman GR. In-training evaluation during an internal medicine clerkship. *Acad Med* 1999;74(10 Suppl):S118-20.
12. Dudek N, Marks M, Lee C, et al. Assessing the quality of supervisors' completed clinical evaluation reports. *Med Educ* 2008;42:816-22.

ANNOTATED BIBLIOGRAPHY

Bandiera GW, Frank J, Sherbino J, editors. *The assessment tools handbook*. Ottawa: Royal College of Physicians and Surgeons of Canada; 2005.

This concise handbook provides a review of learner assessment methods beyond global assessment forms, including selection of an appropriate method, effective deployment of the method, and advantages and disadvantages. It focuses on the CanMEDS set of physician competencies that underpin residency (and increasingly more commonly undergraduate) education in Canada and how to match the competency of interest with an appropriate assessment method.

Epstein RM. Assessment in medical education. *N Engl J Med* 2007;356:387-96.

This review article provides a general discussion of issues related to learner assessment independent of specialty area. The article includes a discussion of the strengths and limitations of specific techniques and how assessment in a clinical rotation fits into the overall assessment of medical learners.

Gray JD. Global rating scales in residency education. *Acad Med* 1996;71(1 Suppl):S55-63.

This review article reviews the evidence and principles behind the design and use of the most common and important document that teachers complete: the end-of-rotation assessment form.