Point-of-Encounter Assessment: Using Health Belief Model Constructs to Change Grading Behaviors

Susan F McLean[1], Maureen Francis[2], Naomi L Lacy[2] and Andres Alvarado[3]

[1]Department of Surgery, Texas Tech University Health Sciences Center El Paso, El Paso, TX, USA. [2]Department of Medical Education, Texas Tech University Health Sciences Center El Paso, El Paso, TX, USA. [3]Office of Diversity, Inclusion, and Global Health, Texas Tech University Health Sciences Center El Paso, El Paso, TX, USA.

Journal of Medical Education and Curricular Development
Volume 6: 1–11

DOI: 10.1177/2382120519840358
SAGE

ABSTRACT

OBJECTIVE: Fourth-year medical students need feedback to improve. Even during 1-month rotations, there needs to be a formal mid-clerkship feedback session. Better feedback involves multiple surgical evaluators at multiple levels. Constructs from the Health Belief Model of behavior change were used to assess faculty and resident grading behaviors to create a more usable evaluation system. A point-of-encounter (POE) system was created. The objective of this study was to review the efficacy of a POE clinical evaluation card (CEC) system which was initiated to increase evaluator's participation in grading and formative feedback prior to mid-clerkship evaluation.

DESIGN: The study was a 1-year retrospective cohort study reviewing the CECs for level of evaluator, content, and student compliance. A Likert-type scale survey regarding the usage of the clinical cards was also completed by evaluators.

SETTING: Texas Tech University Health Sciences Center at El Paso, during 2 fourth-year medical student rotations, Subinternship (Sub-I), and Surgical Intensive Care Unit (SICU).

PARTICIPANTS: 34 fourth-year medical students and 20 evaluators.

RESULTS: Students turned in a mean of 10 cards, 75% in SICU and 65% in Sub-I turned in all 10 cards. There were significantly greater advanced residents evaluating during Sub-I vs SICU: mean evals by PGY3 were 1.9 vs 0.75 ($p = .01$) and mean evals by PGY5 were 1.4 vs 0.1 ($p < .0001$). There were significantly more faculty completing evaluations during SICU vs Sub-I: 2.5 faculty evals/student vs 1.4 faculty evals/student ($p = .023$). Evaluator ratings were high on a 5-point Likert-type scale, with most responses near the "strongly agree" rating of 4.7 to 4.8.

CONCLUSIONS: Use of POE CECs met goals of having at least 7 CECs turned in by mid-clerkship and 10 at end-clerkships. Formative evaluations by mid-clerkship went from 0 to 7 evaluations. Evaluator surveys highlighted clarity and efficiency as reasons for using CECs.

KEYWORDS: educational feedback, clinical encounter, formative assessment, clinical assessment, competency, surgical clerkships

RECEIVED: February 13, 2019. **ACCEPTED:** March 6, 2019.

TYPE: Original Research

FUNDING: The author(s) received no financial support for the research, authorship, and/or publication of this article.

DECLARATION OF CONFLICTING INTERESTS: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

CORRESPONDING AUTHOR: Susan F McLean, Department of Surgery, Texas Tech University Health Sciences Center El Paso, El Paso, TX 79905, USA.
Email: susan.mclean@ttuhsc.edu

Introduction

Medical students on surgical rotations need frequent feedback and substantive formative assessment. These are not only important for student learning; written assessments form the basis of rotation grading. Furthermore, the Liaison Committee on Medical Education (LCME) Standard 9 requires both narrative and formative feedback as part of a student's assessment.[1] Yet getting enough feedback to form these assessments can be a challenge in an era of competing demands.[2] The nature of surgical rotations during the fourth year further complicates student evaluation. Medical students typically move between settings, including the operating room, emergency department, intensive care unit (ICU), and surgical clinic. Because the student usually spends only a few hours with 1 faculty or resident during a surgical rotation event, comprehensive feedback is difficult.

We discovered that our system was not giving us the needed information. Faculty and residents frequently did not fill out formative assessments. These assessments were online, which required logging into the system and covered multiple aspects of all of the student competencies. Because the assessment forms were not completed, there was a paucity of information with which to advise the student on progress and form the final grade. Effectively, the clerkship director had to rely on verbal reports when compiling assessments.

Our intervention goal was to improve formative feedback prior to a mid-clerkship meeting, and also throughout the rotation. A secondary goal of the study was to have more evaluators involved in student evaluation.

Intervention

Setting

At our institution, sub-internships (Sub-I) and critical care required 1-month rotations during the fourth year. Students

interested in surgery take these as the Sub-I and the Surgical Intensive Care Unit (SICU) rotation. There are 2 to 3 students at a time on either the Sub-I or SICU rotation.

During the Sub-I, students are required to round with the team in the morning, and scrub on cases and attend surgical clinics. Students may also go on consults with residents, see patients in the emergency department with residents, and also participate in minor procedures. Team members and evaluators included 3 to 4 faculty per team, and 5 residents per team, consisting of 1 chief resident, 1 third-year resident, 1 second-year resident, and 2 interns, from surgery and occasionally from Family Medicine or Orthopedic Surgery.

The student on SICU pre-rounds on up to 3 patients and reports daily findings on rounds. The student may also assist or be proctored in performing a number of procedures in ICU and may also attend cases in an operating room. The student is required to create a short PowerPoint presentation by the end of the clerkship. There are 3 main faculty who rounded in SICU 1 week at a time, and a team of 4 residents, 1 second- or third-year surgical resident and 3 interns, from general surgery, Ob-Gyn, or emergency medicine.

Understanding the problem

Prior to developing an intervention, faculty and residents were queried in a focus group about why they did not participate in the assessment system. The top reasons faculty and residents gave was that first they did not feel that they could give a fair comprehensive evaluation ("I don't really know the student well enough . . .") and second they felt they did not have enough time to fill out the forms ("it takes too long to fill out the forms"). There were a recognition that the number and scope of formative assessments were inadequate and concerns that the system might result in inaccurate assessments of the student's overall performance. Some faculty had multiple student educational encounters, but using only those faculty would mean that students would only get graded by those faculty and miss out on having performance assessment by all persons who worked with the students.

Framing the problem

The problem was framed by the intervention team using the constructs of the Health Belief Model (HBM[3]; Figure 1). Although the HBM was initially created to explain lack of compliance with tuberculosis screening, the constructs can be adapted to explain and create change in grading behaviors. The model explains motivation to do a health-promoting behavior regarding perceptions and beliefs of the person who is contemplating or needs to do the behaviors. It frequently is applied to health-preserving behaviors, for example, screening or wearing protective equipment, which must be performed to avoid a more distant negative consequence. The HBM is especially good for predicting behaviors in the face of barriers.[3] The HBM constructs to explain a problem are as follows: perceived susceptibility to the problem (or health consequence), perceived severity of the problems (or health consequence), perceived benefits of complying with the new or proposed behavior, and perceived barriers to performing the behavior. Another construct is self-efficacy, which is the belief in oneself that one can perform the behavior, especially in the face of barriers. "Cues to action" is a construct that physical or verbal cues can aid in enhancing motivation to complete the behavior. Although HBM is usually used for health behavior compliance motivation, in this setting we are using these constructs to try to improve the behavior of evaluators performing the behavior of evaluating fourth-year medical students during a clerkship. These constructs fit this problem well.

The "primary problem" was the compliance with grading behaviors. Faculty and residents, the primary evaluators, would not fill out the existing online system, which was designed as a comprehensive evaluation. This evaluation required a computer. The evaluation had 15 questions with mostly Likert-type scale responses and required explanations if a student did well or performed below average. In addition, comments were required at the end of the evaluation. This evaluation took approximately 10 minutes to complete. The HBM constructs which can explain this lack of motivation to comply with grading are explained below.[3]

Lack of perceived risk and lack of perceived severity. These constructs represent the outcomes individuals perceive likely if they do not perform the behavior. Both constructs were issues with grading. Faculty and residents did not expect any personal repercussions if they did not grade the students at all and did not think any repercussions would be severe. They found it easier to have the clerkship director do all the grading, even though they were worried about inaccurate grading. In our review of the system, we decided that this perception was accurate as there were no consequences for not submitting online assessments. One perceived risk which was voiced was that some faculty felt that grading a student negatively could lead to student complaints, more meetings, and more lost time.

Perceived barriers. These were time constraints and need for using a regular (not mobile) computer, which was often not near where the evaluators were working. Time constraints were listed as the length of time to get to a computer, long on, and fill out an online form.

Self-efficacy. We framed the sense that they could not fill out a comprehensive evaluation form because they had not observed all the student performance areas as partly a lack of self-efficacy. Lack of familiarity with students was the feeling by evaluators that they did not know the student well enough to fill out an evaluation by mid-clerkship, at 2 weeks into the rotation.

Cues to action. Before the intervention, cues to action were emails, which again required the evaluators to be at a computer logged on to the university email system. As these cues only

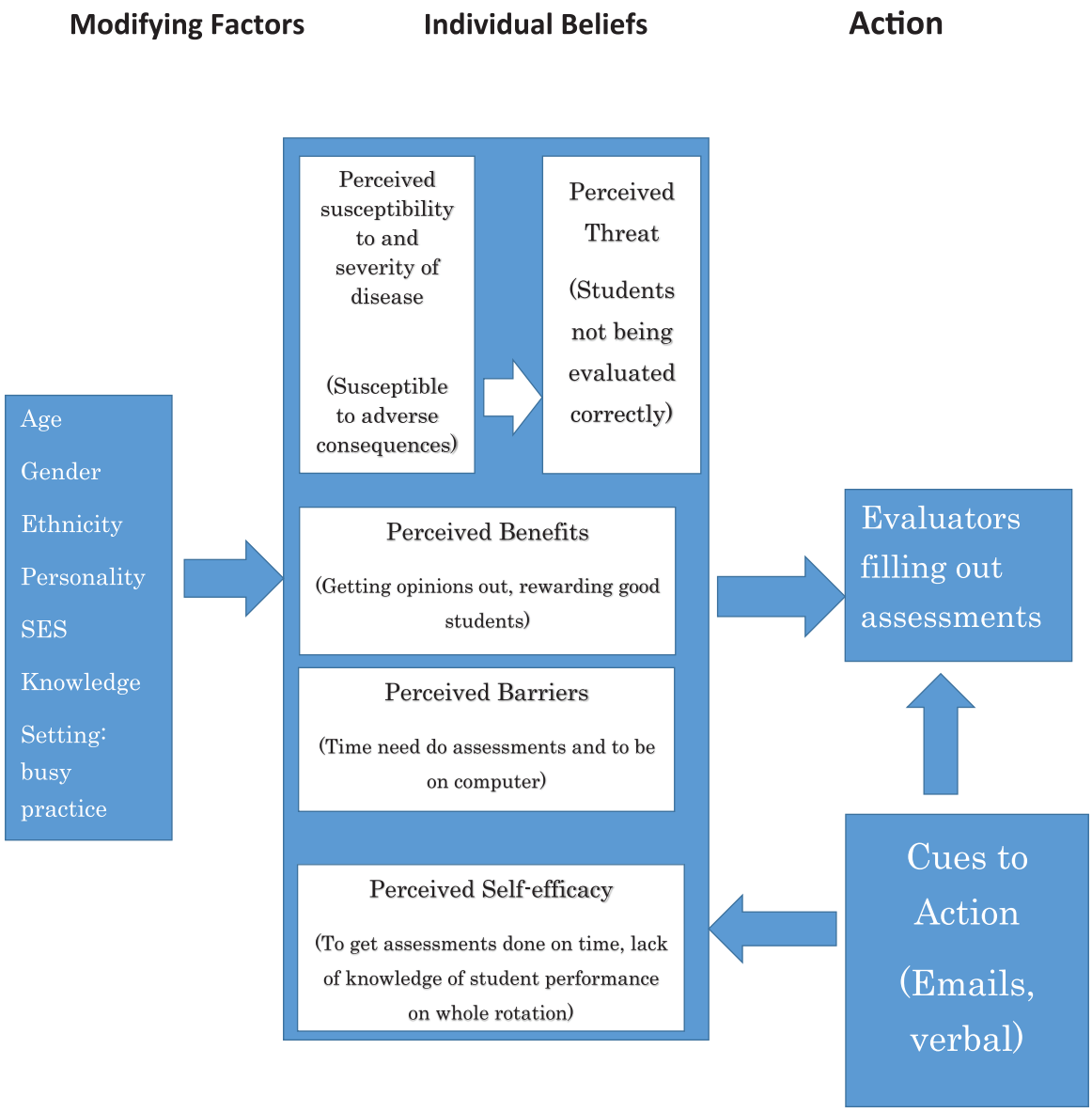


Figure 1. Health Belief Model: Adapted from Skinner et al. "Health Belief Model" in Glanz, Rimer, Viswanath: "Health Behavior, Theory, Research and Practice." (3). Adapted from and Used with permission from Wiley and Sons, Inc.

appeared when the evaluators were on the email system, they were not frequent.

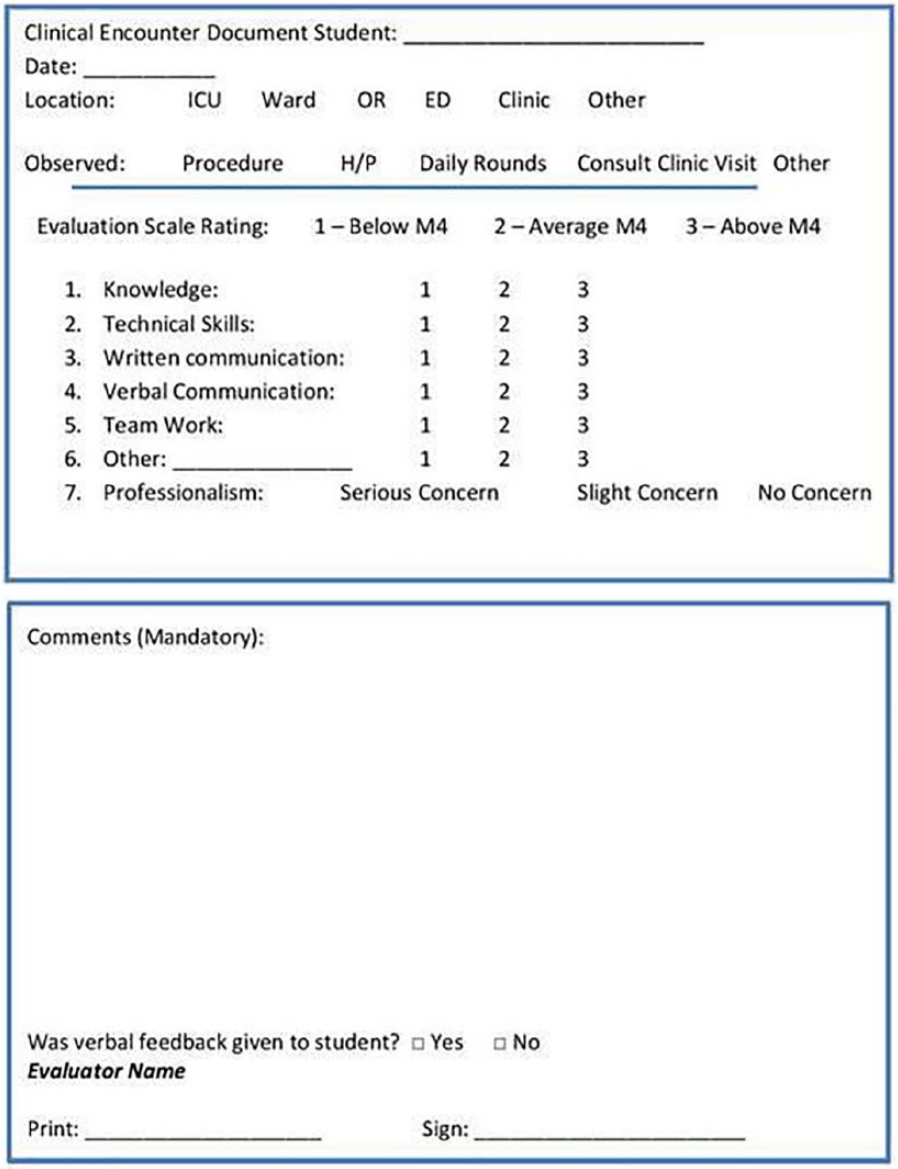
The advantage of this model was that it allowed us to identify the main constructs to address and gave us an additional construct to assist in developing an intervention, "Cues to Action." These are verbal, visual, or other cues which act as reminders to perform a certain behavior.

Literature used in designing the intervention

A literature search on improving feedback and assessment narrowed down to clinical evaluation card (CEC) and point-of-encounter (POE) systems as these seem to address the issues from our HBM model. CECs have been shown to increase feedback[2,4,5] and student's perception that they received feedback.[6] One CEC study noted that students who received written feedback on cards had higher scores on a clinical evaluation test similar to the Observed Skills Clinical Examination (OSCE).[6] Paukert et al[2] noted that they had both resident and faculty evaluators, and noted that as the postgraduate year (PGY) increased, the number of evaluations decreased. POE systems using CECs have been described[4,6-8] with both the benefits seen of the CEC and immediacy of feedback being more useful to students.[7,9]

Designing the intervention

A meeting with the clerkship director for year 4 surgery rotations, the Director for Assessment & Evaluation, the Associate Dean for Student Affairs, and the Assistant Dean for Year 3-4 medical education was held to discuss using CECs to improve formative feedback for year 4 students in the Sub-I and the SICU rotations. The group discussed card design, process of obtaining assessments, total assessments required, and who


Clinical Encounter Document Student: _____
Date: _____
Location: ICU Ward OR ED Clinic Other

Observed: Procedure H/P Daily Rounds Consult Clinic Visit Other

Evaluation Scale Rating: 1 – Below M4 2 – Average M4 3 – Above M4

1. Knowledge: 1 2 3
2. Technical Skills: 1 2 3
3. Written communication: 1 2 3
4. Verbal Communication: 1 2 3
5. Team Work: 1 2 3
6. Other: _____ 1 2 3
7. Professionalism: Serious Concern Slight Concern No Concern

Comments (Mandatory):

Was verbal feedback given to student? □ Yes □ No
Evaluator Name

Print: _____ Sign: _____

Figure 2. CEC Example: CEC, Clinical evaluation card; ED, emergency department; ICU, intensive care unit.

should do assessments. The literature review aided card design. After the meeting, a sample of the proposed CEC was sent to all stakeholders. After several iterations of the CEC, the group decided on the final card design (see Figure 2).

The design took into account both the elements from our HBM model and a set of assumptions about how skills and feedback needs varied from third-year students. To overcome the time barrier, we designed a clinical encounter card which had a front face consisting of items to circle and a back face for comments. Some comment abbreviation lists were also handed out, just to make filling out the cards faster. As these were index cards, no evaluator had to log on to a computer, thus saving more time—the time to travel to a computer, log in, and then log into the grading system. The estimated time to fill out the cards was approximately 2 minutes. To create "cues to action," we instructed the students to give the cards to evaluators and ask them to fill out the cards. To enhance "perceived benefits" to filling out the cards, the entire department was instructed on CEC system use and the decreased time and effort required was stressed. For the "perceived susceptibility" to not doing accurate evaluations, a discussion about the necessity of accurate grading was held at the weekly educational conference of residents and staff. Finally, fourth-year medical students are expected to be able to perform a history and physical, but need increasing amounts of verbal and written communication, multidisciplinary teamwork, and exposure to procedures. The team felt that having the entire back face for comments would increase communication in this area.

Card system

Medical students had a goal of handing out 1 card per day. The minimum requirements were for the students to hand in 7 cards by mid-clerkship and 10 cards by the end of the rotation. Students did not have to turn in every card, but they did

anyway. Students had an option to meet to hand out cards to the same evaluator at the beginning and end of the clerkship. This would also allow the clerkship director to assess progress during the rotation. To ensure that a broad array of individuals participated, not just those students perceived to be easy graders, the clerkship director gave CECs to the faculty.

Evaluators completed cards promptly when asked and wrote a comment on the back of each card, then handed cards back to students. Faculty evaluators had a list of abbreviations to assist with evaluations and save time. To obviate any anxiety that the students may have had about giving an evaluation card to the faculty, each faculty had several cards to fill out and turn in to the clerkship coordinator directly.

The clerkship coordinator compiled the cards for each student for the clerkship director for the mid-clerkship as well as final evaluations. Other activities evaluated included the OPLOG, or a list of procedures/patients/activities student did during the rotation and a student PowerPoint short topic presentation. The number of numerical grades on the front of the card as well as the trajectory of evaluations during the clerkship and also the written comments determined final grades. The final grade report on an online form included the cards' written comments.

Methods

The intervention was implemented during the 2015-2016 academic year and evaluated at the end of the first year. The CEC system was evaluated on several outcomes: the numbers of cards, level of evaluators, locations, evaluation by faculty vs resident, and positive vs negative evaluation.

Evaluators were educated about the new CEC system by having an explanation at the weekly educational session on Thursday mornings. Card samples were handed out, an explanation that students would be handing faculty and residents cards to fill out after brief encounters, eg, surgical cases, clinic, rounds, consults, or ICU rounds, and the expectation that each card would could be completed in 2 minutes or less. Evaluators were told that this replaced the computer evaluations for them. Evaluators were told that comments on the back were mandatory. In addition, the clerkship director met with each critical care attending, showed them the cards, and gave them a list of accepted abbreviations to use on the cards. Evaluators were told that the clerkship director was requiring the students to turn in 10 cards for each rotation.

To evaluate the acceptability of the intervention, both students and faculty were asked to complete a short survey about the cards. The stakeholders who were present in the initial planning meeting reviewed and approved the survey design. The evaluator survey was a series of questions with Likert-type scale 1-5, from strongly disagree to agree, with 3 being neutral.

One set of items was expected to act as a scale relative to the barrier identified as Time.

A Comfort Scale was constructed bringing together variables related to the evaluator's comfort with grading, because a perceived problem before the program was that the negative outcome resulting from having to fill out a long form was that each evaluator felt that they could not accurately evaluate the student on every item. The Comfort Scale was composed of "I found it easier to point out areas for improvement," "I found out it easier to give positive reinforcement," and "When a student gave me a card to complete, I felt comfortable giving feedback to the student at the same time."

The survey was distributed to evaluators at a surgery department meeting.

Statistical analysis

Students who rotated during both rotations were considered separate students for analysis.

Continuous variables were summarized using mean, SD, median, minimum, and maximum, whereas categorical variables were summarized using frequency and percentages. The variable "Type of comment" was created based on 2 variables: "Number of positive comments" and "Received a needs improvement comment." If a student received any number of a needs improvement comment, then they were marked as "yes" for the needs improvement variable. If they received positive comment and no "needs improvement" comment, then they were considered as "received only positive comments." All variables of interest were compared with the type of rotation and the type of comment received. A Poisson regression analysis was used to determine the relationship of all the continuous variables, whereas Fisher exact test was used to determine the association of categorical variables.

Two scale scores were created from the evaluator survey: the "Time Scale" and the "Comfort Grade Scale." The Time Scale was created by adding the variables "I found the CECs easy to read," "I found the CECs easy to fill out," and "The time required to fill out the CEC was reasonable" and dividing that sum by 3. The scales were evaluated for reliability using SPSS v 25.

Values were considered significant at the 5% level of significance. All other analyses were conducted using SPSS v 25.

Results

Participants

A total of 20 students rotated during SICU and 14 during Sub-I in the first year of program implementation. There were 15 surgical residents, 5 surgical faculty, and 1 critical care nurse practitioner (who left part way through). Rotating interns also could assess the student.

Cards

The students submitted a total of 339 CECs. Median number of evaluations was 10, with median 2 faculty and 7.5 residents. Mean was 9.975 or 10. The minimum number of evaluations was 3 (student dropped the rotation) and the maximum was 17 CECs. The mean number of faculty evaluations was 2 and that of resident evaluations was 7.6. There were a total of 7 nurse practitioner evaluations.

Item	Rotation	N	Mean	SD	P
Number of assessments - TOTAL	SICU	20	10.1	3.0	.78+
	Sub_I	14	9.8	2.7	
Number of assessments - faculty	SICU	20	2.5	1.6	.023+
	Sub_I	14	1.4	1.5	
Number of assessments - Resident	SICU	20	7.3	2.8	.35+
	Sub_I	14	8.1	3.1	
Number of assessments - PGY-1	SICU	14	5.6	2.4	.008+
	Sub_I	20	3.6	1.5	
Number of assessments - PGY-2	SICU	20	0.85	0.87	0.16+
	SUB_I	14	1.36	1.44	
Number of assessments - PGY-3	SICU	20	0.75	.85	.01*
	SUB_I	14	1.9	1.0	
Number of assessments - PGY-4	SICU	20	.00	.00	.4*
	SUB_I	14	.07	.26	
Number of assessments - PGY-5	SICU	20	0.1	0.31	<.0001*
	SUB__I	14	1.4	.93	
+ Poisson Regression *Fisher's Exact test					

Figure 3. CEC assessment total by location and assessor level.
CEC, clinical evaluation card; PGY, postgraduate year; SICU, Surgical Intensive Care Unit; Sub-I, Subinternship.

Per rotation, the mean numbers of CECs were 10.1 for SICU and 9.8 for Sub-I, again because 1 student dropped without a formal CEC for faculty. There were differences by level of evaluator (Figure 3); SICU rotation had significantly more CECs from faculty (mean 2.5 vs 1.4, $p = .048$; Figure 2). In the rest of the evaluations, for residents, there was a significant difference in the numbers of PGY1, PGY3, and PGY5 evaluations. The numbers of PGY3 and PGY5 were different with students in the Sub-I receiving more evaluations from PGY3 and PGY5 residents (Figure 3). Although we met our goals with mean number of evaluations, there were some differences among students. In total, 75% of students in SICU had 10 or greater cards, whereas 65% of students in Sub-I had 10 or greater cards. All students except 1 had 7 or more cards.

The activities during the 2 rotations during which evaluations were completed varied between the 2 rotations. Overall, there were 237 evaluations during daily rounds, 127 evaluations after procedures, 46 after history and physical, 23 during "other" which includes PowerPoint presentations, 6 during clinic, and 2 during consults. By rotation, there were differences, with SICU having more rounds and Sub-I having more evaluations during procedures, and these were significantly different (Figure 4).

Qualitative analysis. The research team reviewed cards for positive and negative comments. In total, 300 cards had positive comments and 29 had needs improvement comments. The researchers assessed whether or not a comment was positive if an average person would consider it positive, the wording expressed that the student had done something correctly, in an above-average manner, or was especially helpful. A "needs improvement" comment was assessed to be any comment which stated that the student either did not do something expected, needed improvement in skills, or was commented to be below expected levels in an activity. These comments were assessed by reading every comment and assigning them to "positive" or "needs improvements" values. An example of a positive comment was "[student] is clearly dedicated to surgical work and is proactive to change dressings and always offers to pull drains." An example of a needed improvement comment would be: "Did not participate in morning rounds for unknown reasons at times." The 10 other cards either did not have a comment or the comment was neutral (Figures 5 and 6). Professionalism evaluations were almost 100% "No concerns." There was 1 student who had "slight concern" after 3 evaluations. That student met for discussion and ultimately dropped the rotation.

There was a majority of positive comments (Figure 5). Figure 5 shows a histogram of numbers of positive comment cards per student. The range is 1 to 15. There are modes at 7, 9, and 10. Nine students had 10 positive comments on their cards. One student turned in 15 cards all with positive comments. Figure 6 shows a histogram of negative comments. Note that 20 students received no negative comments.

Acceptability survey results

Only 2 students returned surveys regarding using the CECs. These were favorable, however, too few to report conclusions.

Activity	Rotation	Sum	Mean	SD	p
Procedure	SICU	40	2	1.6	<.001*
	Sub-I	87	6.2	2.7	
History/Physical	SICU	35	1.7	1.9	.013+
	Sub-I	11	.8	1.0	
Daily Rounds	SICU	167	8.3	3.8	<.001+
	Sub-I	70	5.0	2.4	
Consult	SICU	0	0	0	.16*
	Sub-I	2	.14	.4	
Clinic	SICU	0	0	0	.02*
	Sub-I	6	.4	.9	
Other	SICU	21	1.1	1.6	.02*
	Sub-I	2	.14	.5	
+ Poison regression *Fisher's exact test					

Figure 4. Activity differences between rotations.
SICU, Surgical Intensive Care Unit; Sub-I, Subinternship.

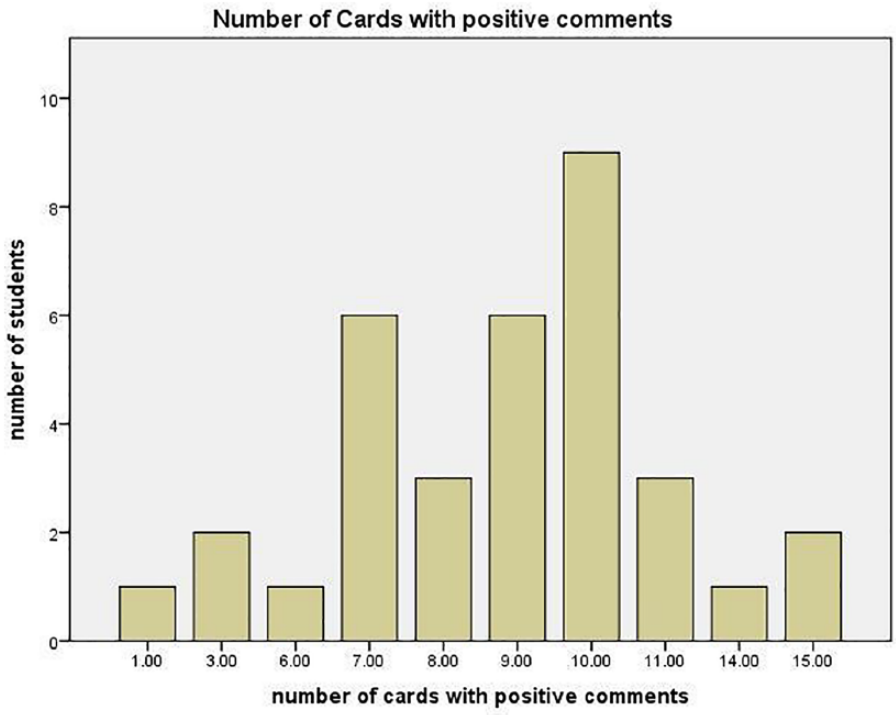


Figure 5. Distribution of the number of cards with positive comments.

Of the 21 (40%) evaluators, 8 turned in surveys. For the demographic questions, the most common location from the evaluator perspective was rounds, followed by operating room (OR), then verbal presentations, then clinic. For efficiency questions, mean Likert-type scores were as follows: "easy to read"—4.8, "easy to fill out"—4.7, "created own comments"—4.8, and "time used was reasonable"—4.8.

For the content interaction questions on the evaluator survey, "I was comfortable giving feedback," the mean response was 4.7. For "I found it easier to point out areas for improvement," the mean response was 4.7. For "I found no difference with feedback from CECs vs online evaluations," the mean score was 3.2. For "I found it easier to give positive feedback," the mean response was 4.0. For "CECs are fair and unbiased method of grading," the mean response was 4.0. The survey item "When a student gave me a card to complete, I felt comfortable giving feedback to the student at the same time" had an average Likert-type scale score of 4.7.

At the start of the project, the research team identified 2 barriers to evaluation: lack of time and grading self-efficacy, the belief that evaluators could give a fair grade with limited exposure. The mean Time Scale score was 4.78 (±0.31). Reliability

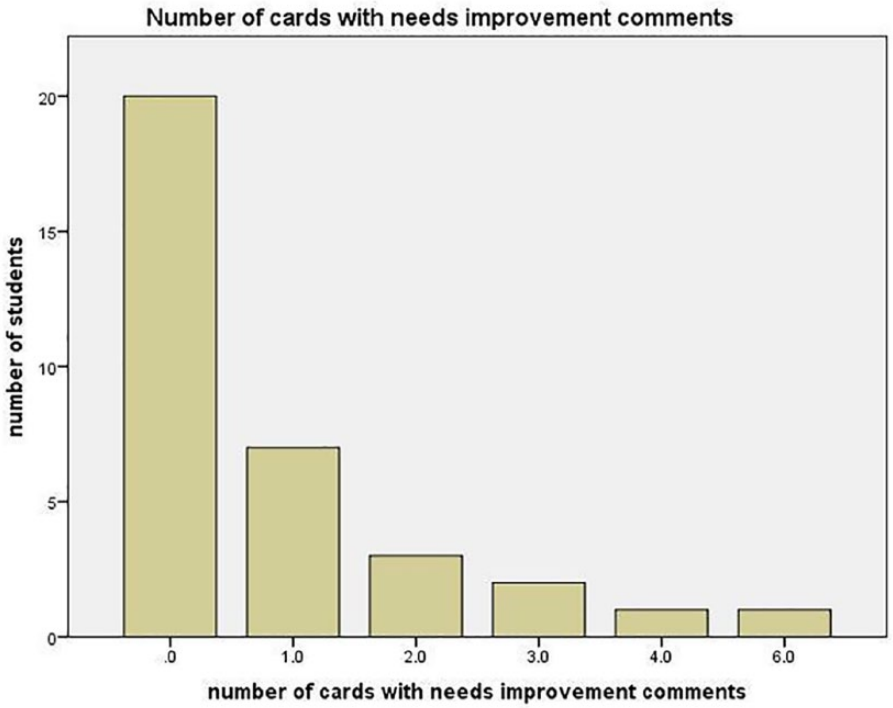


Figure 6. Distribution of the number of cards with needs improvement comments.

analysis for the Time Scale results produced a Cronbach alpha for the Time Scale of 0.959. The Comfort Grade Scale was created by adding the variables "When a student gave me a card to complete, I felt comfortable giving feedback to the student at the same time," "I found it easier to point out areas for improvement to the student using the CECs," "I found it easier to give positive feedback using the clinical evaluation cards," and "I believe that the clinical evaluation cards are a fair and unbiased method of grading" and dividing this sum by 4. Mean Comfort Grade Scale was 4.24 (±0.64). Cronbach alpha for Comfort Grade Scale was 0.81. We are limited in this analysis by the small sample size of only 8 persons filling out this survey; however, most respondents answered the same way. The standard deviation for the mean Time Scale score was only 0.31 and for the comfort scale score it was only 0.81. Thus, the survey respondents indicated that CECs overcame 2 important barriers to evaluation: lack of time and grading self-efficacy.

Discussion

The POE CEC system dramatically increased the number of student evaluations in the 2 fourth-year student rotations: SICU and Sub-I. To our knowledge, this is the first study specific to using CECs as a POE system during fourth-year surgical clerkships. Using the constructs of the HBM, especially eliminating barriers to evaluations and increasing cues to action, helped overcome the lack of evaluations. Constructs from the HBM were useful in constructing and explaining how the CEC system worked. Evaluators found the CECs efficient and easy to use and reported comfort giving both positive and negative feedback on the survey. These responses on the survey and the strong response to the program show that "barriers" were overcome, especially the time barrier. The Time Scale showed a high mean score for time-related items, with a high Cronbach alpha. This showed that the time barrier was overcome.

Another barrier was evaluator comfort with the grading system, having the belief that they would give a fair grade. The return of more than 300 cards per year across 2 rotations and multiple settings shows an increase in self-efficacy by the graders.

Evaluators averaged 4.24 out of a possible 5 on this scale, indicating a high level of comfort with evaluating students using the CECs. Using the HBM constructs to explain these results, one can conclude that the negative outcome of having inaccurate evaluations was overcome by evaluating each clinical encounter at the POE and not having evaluators worry about global student performance.

In results, we noted that comparing the 2 rotations, SICU and Sub-I, there was a difference in personnel who completed cards, with more faculty in the SICU and more PGY3-5 in the Sub-I. There are 2 reasons for this: First, faculty are present in SICU for more hours at one time, because rounds typically take at least 3 hours. Hence, there are more opportunities for the students to hand the faculty a card. There are no PGY5s usually in the SICU, except for one 2-month time period where a resident was doing extra SICU rotations in the PGY5 year. Hence, in SICU, there were less PGY5 evaluations. In contrast, in Sub-I, faculty are present during surgery, but, when scrubbed, are using sterile technique, and hence do not fill out cards until the end of surgery. At that time, if a faculty is called out before a student can hand a card, then that student may hand the card to a senior resident. PGY1-5 residents are rotating with the

students on the Sub-I rotations, so likely the residents were more available. In assessing the evaluations, we felt that more senior residents would be good evaluators, because after PGY3, in other specialties, these persons would be old enough to be an attending. In summary, the likely increased number of faculty evaluators during SICU was due to ease and availability.

This feedback is especially important during fourth-year student surgical rotations, as many of these students are interested in becoming surgeons. Students handed more cards to residents compared with faculty, probably because there were far more residents on rotations, but possibly because the students perceived the resident to grade more leniently.

We evaluated the comments and noted that there were far more positive comments than needs improvement comments. This issue has been seen in other surgery evaluation systems. McQueen et al[10] noted, in a surgery residency program, that a focus group of faculty evaluators tended to give more positive comments than negative.

Our study results are similar to those previously published,[2,4-7,9] despite the differences in our card design and population. Our cards allowed for more narrative feedback and had fewer predesignated areas. Our system was focused on rapid and strategic feedback rather than the comprehensive format used by Dunnington et al.[7] Another difference was that there was slightly more "needs improvement" feedback in Schum et al's[9] study. In general, this study's evaluation system was designed to be rapid and to pick up brief encounters, to add to formative evaluation.

Two recent studies also used a feedback design meant to be brief and featured student-initiated feedback, as in our student. A study of third-year medical students on a surgery clerkship was also evaluated by a rapid evaluation system designed for brief encounters, not comprehensive. Hughes et al[15] designed a "Minute Feedback System." This system was designed to be 1 minute or less, designed to be initiated by the student, similar to our system. The Minute Feedback System had very similar grading items as our CECs: location (floor, emergency department, clinic, OR) and which clinical skills the student was requesting feedback on (physical examination, patient history, oral presentation, technical skills, general performance). Their system worked by having a student fill out name, location, which skills, and sending it to faculty by email. They also noted an increase in evaluations in the intervention group. This system was tested in the first 6 students, then in 31 students.[10] Another study of student-initiated feedback in a third-year surgery clerkship featured verbal feedback meant to be quick and for shorter episodes of clinical work.[11] Students were instructed to ask for feedback with any faculty with whom the student spent more than 3 hours. Students were given a video on how to ask for feedback. Faculty were also instructed. Feedback was meant to be 5 to 10 minutes. There was a crossover design, with students going to the planned daily feedback vs usual custom feedback every 2 weeks, with surveys done after every 2 weeks. This study had a small sample size of 33 students and had some selection bias in that students had to elect to be in the study. Only 62% of eligible students enrolled (33 of 53 total eligible) and had a higher amount of students who were more interested in surgery as a career, compared with non-participants. There was a significant increase in weekly feedback sessions in the experimental group ($p = .002$).[11] These findings in 2 third-year clerkships with student-initiated brief feedback systems are similar to our study, yet ours is in fourth-year students, who have shorter rotations. Our system did not have selection bias, in that all students participated.

Formative feedback is essential to medical student learning and tends to be qualitative. This is especially true of shorter rotations. More educators are realizing the importance of formative feedback, especially before the mid-clerkship time, so that students may improve during the clerkship.[12] There is noted "grade inflation" on many surgical clerkships, especially during the fourth year.[13] In a recent study of orthopedic surgical fourth-year clerkships, it was noted that 80% to 100% of students received honors. When program residency coordinators responded to the survey, the items most consistently ranked as highly important for grading were "enthusiasm and interest" (100% highly important), "professionalism" (96%), and student relationship with residents (93%). These are all qualitative measures, which would be captured by a CEC. "Knowledge base" was the fourth most likely item to be ranked as "highly important," by 89% of respondents.[13] In this study, orthopedic students, while numerically graded higher, were apparently judged more qualitatively. The CECs for our study captured the aforementioned type of qualitative information, especially enthusiasm, interest, and professionalism.

Limitations of this study are that we have a smaller group of students than the studies on third-year clerkships. This is typical of fourth-year medical student rotations. We attempted to assess this evaluation system by both student survey and evaluator survey, we did not get enough student surveys back to evaluate, but we did get 40% resident and faculty surveys back so we could evaluate the usefulness of the survey. We had a large volume of cards: 339 cards were returned, this number is like similarly sized studies; at an average of 10 cards per student, this study is comparable with other POE systems. In addition, the 10 card per student average shows we met our original objective. Finally, we did not put "competencies" on the card, instead we listed clinical skills. The clerkship director translated the clinical skills into the competencies for final evaluations. For example, a great rating on a PowerPoint presentation showed both a good score on professional communication and also practice-based learning. A comment on "great team player" showed that the student had excellent interpersonal skills and patient care. "Great fund of knowledge" and a score of "3" or "above MS4" on "knowledge" translated to medical knowledge competencies. One potential limitation could be that there is less information on each card than on a comprehensive examination. However, clinical encounter cards

have been evaluated and, if evaluators are educated, can provide as much information in aggregate as a comprehensive evaluation.[8]

We used constructs from the HBM, which is usually used to explain or understand motivations to comply with health behaviors. The constructs in this model applied well to our problem of evaluation. Other models of motivation might be used. One model could be Heckhausen's expectancy-value model: in this model, the expected consequences of one's actions determine the likelihood of complying with a behavior.[14] For example, if there are not any negative expected consequences and the behavior is cumbersome, it is less likely to be complied with. In this model, situations factor importantly, as behaviors which might be expected to have an outcome which is positive in one situation would be negative in another. This model would not apply well to our problem, because the main issue was barriers to action, not expected outcomes. Once the barriers were removed, the compliance with formative evaluations being finished went from 0 to more than 300 evaluations in 1 year. The constructs from the HBM explained and predicted the success of using CECs.

Applying the constructs of the HBM, the barriers mentioned when talking to faculty, of lack of time and the feeling that each faculty could not perform a comprehensive evaluation, were removed by making shorter evaluations which evaluated a shorter episode of clerkship activity. Self-efficacy was enhanced by explaining to the evaluators how to use the cards and the purpose of each part. Perceived severity and perceived susceptibility to consequences, which were mostly important because it explained why people were not motivated to complete evaluations, were even less important as now people were filling out evaluations. Another way to approach the problem would be to create some severe consequences for not filling out the longer evaluations, for example, threats of bad evaluations to faculty, but this would likely breed some animosity toward having yet another task to do. By making evaluations shorter, it created a win-win situation, in that people were compliant to fill out cards, and wrote comments which could be used on the student's applications to residency letters.

Future directions

This intervention was deemed successful because it increased formative feedback and quantity of evaluations and overcame the barriers of time and self-efficacy in faculty evaluators by providing a way to quickly and frequently obtain feedback for students. Other innovations could be to make the same information electronic as in the Minute Feedback System,[15] or as in a no-cost smartphone-based resident evaluation system.[16]

In summary, the needs for formative evaluation for learners on surgery rotations often pit student's needs against the time demanded of surgeon-evaluators for patient care and related activities. Using the framework of the HBM, the clerkship team was able to overcome barriers to filling out evaluations, increase evaluators' self-efficacy about evaluating students, and allay concerns that the evaluator did not have comprehensive knowledge of a student performance. By evaluating observable parts of the student performance and only having to evaluate these parts, faculty and residents felt comfortable doing evaluations. This comfort increased overall evaluations, and by adding together multiple encounters, a more complete picture of student performance over the clerkship by more participants in evaluation was accomplished.

Conclusions

Institution of a POE CEC system in 2 surgical rotations in the fourth year of medical school enhanced the evaluation process by providing timely quantitative and qualitative evaluations across a variety of clinical activities and settings. Constructs of the HBM of health behavior change were used to provide a framework to change the evaluative behaviors. Fourth-year student grading was enhanced by these cards by providing timely and personal formative evaluation.

Author Contributions

SFM (MD), originated the idea for clinical evaluation cards to be used in a Point-of-Encounter System. And wrote the first draft and was main writer and editor for all subsequent drafts submitted to the journal.

NL assisted with creating the card system and also assisted with major revisions to the second draft and read and approved all submissions.

MF contributed to the creation of the CEC card system and reviewed the original draft and all drafts.

AA performed statistical analysis and wrote the section on statistical analysis.

All authors were sent and approved the final draft.

ORCID iDs

Susan F McLean https://orcid.org/0000-0002-9659-1727
Naomi L Lacy https://orcid.org/0000-0001-7885-7539

REFERENCES

1. LCME ed. *Standard 9: Teaching, Supervision, Assessment, and Student and Patient Safety*. Washington, DC: Association of American Medical Colleges and the American Medical Association; 2017.
2. Paukert JL, Richards ML, Olney C. An encounter card system for increasing feedback to students. *Am J Surg*. 2002;183:300–304. doi:10.1016/S0002-9610(02)00786-9.
3. Skinner CS, Tiro J, Champion Victoria L. Health belief model chapter 5. In: Glanz K, Rimer BK, Viswanath K eds. *Health Behavior: Theory, Research and Practice*. 5th ed. San Francisco, CA: Jossey-Bass Public Health; John Wiley & Sons; 2015:75–94.
4. Gandy RE, Richards WO, Rodning CB. Action of student–resident interaction during a surgical clerkship. *J Surg Educ*. 2010;67:275–277. doi:10.1016/j.jsurg.2010.07.003.
5. Richards ML, Paukert JL, Downing SM, Bordage G. Reliability and usefulness of clinical encounter cards for a third-year surgical clerkship. *J Surg Res*. 2007;140:139–148. doi:10.1016/j.jss.2006.11.002.
6. Ozuah PO, Reznik M, Greenberg L. Improving medical student feedback with a clinical encounter card. *Ambul Pediatr*. 2007;7:449–452. doi:10.1016/j.ambp.2007.07.008.

7. Dunnington G, Reisner L, Witzke D, Fulginiti J. Structured single-observer methods of evaluation for the assessment of ward performance on the surgical clerkship. *Am J Surg*. 1990;159:423–426. doi:10.1016/S0002-9610(05)81289-9.
8. Cheung WJ, Dudek N, Wood TJ, Frank JR. Daily encounter cards—evaluating the quality of documented assessments. *J Grad Med Educ*. 2016;8:601–604. doi:10.4300/JGME-D-15-00505.1.
9. Schum TR, Krippendorf RL, Biernat KA. Simple feedback notes enhance specificity of feedback to learners. *Ambul Pediatr*. 2003;3:9–11. doi:SFNESO>2.0.CO;2.
10. 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. doi:10.1016/j.amjsurg.2015.10.002.
11. Garner MS, Gusberg RJ, Kim AW. The positive effect of immediate feedback on medical student education during the surgical clerkship. *J Surg Educ*. 2014;71:391–397. doi:10.1016/j.jsurg.2013.10.009.
12. Hochberg M, Berman R, Ogilvie J, et al. Midclerkship feedback in the surgical clerkship: the "professionalism, reporting, interpreting, managing, educating, and procedural skills" application utilizing learner self-assessment. *Am J Surg*. 2017;213:212–216. doi:10.1016/j.amjsurg.2016.08.001.
13. Johnson JP, Patel N, Wang P, Mulcahey M. Grading medical students during their fourth year orthopaedic surgery rotations [published online ahead of print January 2017]. *MedEdPublish*. doi:10.15694/mep.2017.000133.
14. Eccles JS, Wigfield A. Motivational beliefs, values, and goals. *Annu Rev Psychol*. 2002;53:109–132. doi:10.1146/annurev.psych.53.100901.135153.
15. Hughes DT, Leininger L, Reddy RM, Sandhu G, Englesbe M. A novel minute feedback system for medical students. *Am J Surg*. 2016;213:330–335.
16. Green JM. An innovative, no-cost, evidence-based smartphone platform for resident evaluation. *J Surg Educ*. 2016;73:e14–e18. doi:10.1016/j.jsurg.2016.07.016.