



Value and Barriers to Use of the SIMPL Tool for Resident Feedback

Megan Eaton, DO,^{*} Rebecca Scully, MD,[†] Mary Schuller, MS,[‡] Amy Yang, MS,[‡] Douglas Smink, MD,[†] Reed G. Williams, PhD,[§] Jordan D. Bohnen, MD,[†] Brian C. George, MD,^{||} Jonathan P. Fryer, MD,[‡] and Shari L. Meyerson, MD[¶]

^{*}University of Texas Health San Antonio, San Antonio, Texas; [†]Massachusetts General Hospital, Department of Surgery, Boston, Massachusetts; [‡]Northwestern University, Department of Surgery, Chicago, Illinois; [§]Indiana University, Department of Surgery, Bloomington, Indiana; ^{||}University of Michigan, Department of Surgery, Ann Arbor, Michigan; and [¶]University of Kentucky, Department of Surgery, Lexington, Kentucky

OBJECTIVE: The System for Improving and Measuring Procedural Learning (SIMPL) is a smart-phone application used to provide residents with an evaluation of operative autonomy and feedback. This study investigated the perceived benefits and barriers to app use.

DESIGN: A database of previously performed SIMPL evaluations was analyzed to identify high, low, and never users. Potential predisposing factors to use were explored. A survey investigating key areas of value and barriers to use for the SIMPL application was sent to resident and faculty users. Respondents were asked to self-identify how often they used the app. The perceived benefits and barriers were correlated with the level of usage. Qualitative analysis of free text responses was used to determine strategies to increase usage.

SETTING: General surgery training programs who are members of the Procedural Learning and Safety Collaborative.

PARTICIPANTS: Surgical residents and faculty.

RESULTS: At least 1 SIMPL evaluation was created for 411 residents and 524 faculty. Thirty percent of both faculty and residents were high-frequency users. Thirty percent of faculty were never users. One hundred eighty-eight residents and 207 faculty (response rate 46%) completed the survey. High-frequency resident users were more likely to perceive a benefit for both numerical evaluations (76% vs 30%) and dictated

feedback (92% vs 30%). Faculty and residents commonly blamed each other for not creating and completing evaluations regularly (87% of residents, 81% of faculty). Suggested strategies to increase usage included reminders and integration with existing data systems.

CONTRIBUTIONS: Frequent users perceive value from the application, particularly from dictated feedback and see a positive impact on feedback in their programs. Faculty engagement represents a major barrier to adoption. Mechanisms which automatically remind residents to initiate an evaluation will help improve utilization but programs must work to enhance faculty willingness to respond and dictate feedback. (J Surg Ed 76:620–627. © 2019 Association of Program Directors in Surgery. Published by Elsevier Inc. All rights reserved.)

KEY WORDS: graduate medical education, surgery, feedback

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

INTRODUCTION

For the past several years, the Procedural Learning and Safety Consortium has been using and studying System for Improving and Measuring Procedural Learning (SIMPL), a smartphone-based assessment tool that provides a means for facilitating ongoing assessment of resident performance and autonomy. It also provides timely, case-specific feedback intended to help residents achieve higher levels of performance and autonomy with future procedures. The American Board of Surgery (ABS) sponsored a study which demonstrated

Funding Source: The initial development of SIMPL was funded by grants from Massachusetts General Hospital, Northwestern University, and Indiana University. Later development was funded by contributions from the 26 institutional members of PLSC, grants from the American Board of Surgery and from the Association of Program Directors in Surgery.
Correspondence: Inquiries to Shari L. Meyerson, MD, Section of Thoracic Surgery, University of Kentucky, 740 S Limestone, Suite A301, Lexington, KY 40536; e-mail: shari.meyerson@uky.edu

the feasibility and validity of the SIMPL app as an evaluation tool.¹ This app uses the Zwisch scale, a 4-point scale ranging from minimum autonomy at the bottom (Show and Tell) to maximum autonomy at the top (Supervision Only). After completing an operation, either the faculty member or the resident can initiate an evaluation via the SIMPL application. The SIMPL app asks the resident and faculty member involved in an operative procedure to independently rate the amount of autonomy the resident received and the faculty member can dictate brief feedback which is accessible immediately to the resident.

As resident training shifts toward an emphasis on competency rather than simply time in training, new methods must be developed and implemented to provide frequent and accurate assessment as well as real-time feedback to residents. Williams et al. showed that even for commonly performed procedures, such as laparoscopic cholecystectomy, at least 17 individual ratings of operative performance are required to achieve a reliable rating.² When the broad range of infrequently performed procedures in general surgery is considered, at least 40 ratings are needed for reproducibility. Our current evaluation systems, which rely on global subjective assessments and end-of-rotation evaluations, are not up to the task of providing this level of detail. The daily case-level numerical assessment of autonomy provided by the SIMPL is particularly valuable for the program to monitor progression to independence.

In addition, end-of-rotation evaluations do not provide adequate and timely formative feedback. A study by Gundle et al. surveyed residents and faculty and showed a majority of both groups felt that end-of-rotation evaluations were not adequate for surgical skills feedback.³ Residents also reported these evaluations were rarely or never filled out in a timely fashion. Overall, residents in their survey reported low frequencies of immediate feedback, but a qualitative analysis of their comments showed a great desire for case-by-case and immediate feedback. Immediate feedback is key because Williams et al. showed that evaluations completed more than 72 hours after an interaction lack detail and clarity as memories fade.⁴ Although end-of-rotation evaluations may provide a global impression of the resident's performance, they lack detailed formative feedback and are susceptible to multiple biases including the recency effect where the overall evaluation is dominated by the most recent interaction.

Despite its ease of use and the clearly described need for better strategies for evaluation and feedback, adoption of SIMPL within programs has faced barriers. This study seeks to better describe those residents and faculty within the utilizing programs who have enthusiastically

adopted SIMPL as well as those who have avoided its use and seeks to determine the perceived benefits of use and barriers to use among frequent and infrequent users at both the faculty and resident level.

METHODS

SIMPL Utilization

After Institutional Review Board approval, demographic and use data were collected from the pre-existing database of SIMPL evaluations performed during the ABS study. For resident participants, data included gender, postgraduate year of training, number of evaluations created for cases in which they participated, and mean evaluation score received. For faculty participants, data included gender, number of evaluations created for cases in which they participated, and mean evaluation score given. For both groups, the number of evaluations created in which both the resident and faculty member completed the evaluation was also calculated.

Users were divided into groups based on the number of evaluations created with the high user group including those who had greater than the mean number of evaluations created involving them. Super users were defined as having more than 1 standard deviation above the mean number of evaluations created. Nonusers were those who never had an evaluation completed by both parties named in the evaluation.

Benefits and Barriers Survey

After obtaining Institutional Review Board approval for this study, resident and faculty SIMPL users from one of the pilot programs were invited to participate in a focus group on a voluntary basis. These discussions were used to elicit opinions pertaining to perceived value of SIMPL and issues encountered with the application. This led to the development of 2 separate surveys targeting faculty and resident SIMPL users. The resident survey contained 23 questions and the faculty survey 20 questions. Respondents were asked to self-report in what percentage of cases they use the SIMPL application. To investigate the benefits of use and barriers to use, questions were designed with a statement describing a potential benefit or barrier and respondents were asked their level of agreement with the statement (strongly agree to strongly disagree). For analysis purposes, results were grouped into agree or disagree as few participants felt strongly in either direction. To capture benefits and barriers not covered in the survey, the final 2 questions were free response format asking for additional issues that prevented

TABLE 1. American Board of Surgery Study Participant Demographics and Use of SIMPL App

	Residents	Faculty
Participants with at least 1 evaluation created	411	524
Female	174 (42%)	147 (28%)
Year of Training		
1	69 (17%)	
2	87 (21%)	
3	92 (22%)	
4	85 (21%)	
5	78 (19%)	
Mean # of evals created	22 ± 26	17 ± 23
Median # of evals created	13	9
Mean # of evals completed by both participants	15 ± 18	16 ± 22
Median # of evals completed by both participants	9	8

regular use of SIMPL and any recommended changes to the app. Surveys were distributed to current resident and faculty users of SIMPL in 13 general surgery programs from the prior ABS study which are still using SIMPL (1 program withdrew from use after the study). The surveys were completed either on paper during a program's didactic time or via an email link to survey-monkey.com. Method of distribution was dependent on program preference.

A descriptive analysis of both resident and faculty responses was performed based on self-reported use of the application. Chi-square tests were used to assess relationships between participants' characteristics and survey responses. For self-reported use comparison, the groups were divided into those who reported never using SIMPL, those who reported using SIMPL for less than 50% of cases, and those who reported using SIMPL for at least 50% of cases.

The free response answers were compiled, and a qualitative analysis was completed separately by 2 members of the research team. The free responses were coded according to a grounded theory thematic analysis process.⁵ The ideas were grouped into categories, and themes were developed separately by each analyzer. The 2 analyses were then compared and reconciled.

Statistical Analysis

Comparisons between groups were performed using chi-square tests for categorical variables such as gender, year of training, and survey responses and *t* tests for continuous variables such as mean autonomy score. Interactions between variables were studied using univariate analysis of variance. Statistical analysis was performed using SPSS version 25 (IBM, New York). For all comparisons significance was defined as $p < 0.05$.

RESULTS

SIMPL Use

Over the period of the ABS study,¹ 411 residents and 524 faculty members at 14 programs had at least 1 operative evaluation created for a case in which they participated. Among participating residents, 174 (42%) were female and among participating faculty, 147 (28%) were female. Residents were evenly distributed between all years of clinical training (Table 1). A mean of 22 ± 26 evaluations per resident were created ranging from 1 to 184 evaluations. A portion of these was not completed by the other party in the operating room leaving a mean of 15 ± 18 completed evaluations per resident (range 1-130, response rate 68%). From those completed evaluations, the mean autonomy score was 2.6 ± 0.6 with a range of 1 (Show and Tell)—4 (Supervision Only). Only 18 (4%) residents never had an evaluation completed by both parties for them (nonusers). These nonusers did not show any demographic differences from the entire participant group with respect to gender or year of training ($p = 0.76$ for gender, $p = 0.37$ for year of training).

Further investigation was performed based on the frequency of use of the app. High users were defined as having more than the mean number of evaluations created for cases in which they participated (>22) and included 129 residents (31%). This high user group was not different with respect to the gender of the high users (43% female, $p = 0.64$), but high users were more likely to be senior residents with the largest proportion being in their fourth year of training ($p < 0.001$, Table 2). Mean autonomy score for high users was the same as for lower frequency users (2.6 ± 0.4 vs 2.6 ± 0.6, $p = 0.79$). To further investigate those who used the app most often, a superuser group was defined as having more than 1 standard deviation above the mean number of evaluations created for them (>48) and included 46

TABLE 2. Frequency of Use by American Board of Surgery Study Participating Residents

	All Users	Superusers	High Users	Nonusers
# of participants	411	46 (11%)	129 (31%)	18 (4%)
# of evaluations created	range 1-184	range 49-184	range 23-184	range 1-11
Female	174 (42%)	16 (35%, $p = 0.27$)	52 (43%, $p = 0.64$)	7 (39%, $p = 0.76$)
Year of Training		$p = 0.006$	$p < 0.001$	$p = 0.37$
1	69 (17%)	7 (15%)	19 (15%)	2 (11%)
2	87 (21%)	3 (7%)	12 (9%)	7 (39%)
3	92 (22%)	7 (15%)	31 (24%)	2 (11%)
4	85 (21%)	17 (37%)	40 (31%)	4 (22%)
5	78 (19%)	12 (26%)	26 (20%)	3 (17%)
Mean autonomy score	2.6 ± 0.6	2.6 ± 0.4 ($p = 0.71$)	2.6 ± 0.4 ($p = 0.79$)	

residents (11%). The demographics of this group was similar to the larger high user group in that there was no difference in the proportion of female residents (35%, $p = 0.27$). There remained significantly more senior residents in the superuser group with a disproportionate number of both fourth and fifth year residents ($p = 0.006$, Table 2). Mean autonomy score for superusers was not different from users of lower frequency (2.6 ± 0.4 vs 2.6 ± 0.6 , $p = 0.71$).

Since year of training has been shown to have a significant effect on autonomy and the superuser and high user groups contained a higher proportion of senior residents, additional analysis was performed using univariate analysis of variance to account for the effect of year of training on operative autonomy. In the high user group, the effect of year of training was significant ($p < 0.001$), but the effect of high user status was not ($p = 0.47$). The interaction between those 2 factors was not significant either ($p = 0.29$). The results were the same in the superuser group with a significant effect of year of training ($p < 0.001$) but no effect of superuser status ($p = 0.52$) and no significant interaction between the 2 factors ($p = 0.55$).

A mean of 17 ± 23 evaluations was created per faculty member Table 3. One hundred and fifty-six faculty (30%) received up to 38 requests for evaluation but never responded to a single evaluation (nonusers). This was significantly higher than the proportion of nonusers among the residents (30% vs 4%, $p < 0.001$). Of those who responded to at least 1 evaluation, response rate was high with faculty completing a mean of 16 ± 22

evaluations (94%). Forty-five superusers (9%) and 159 high users (30%) were identified among the faculty, which were similar proportions to the number in those groups from the resident cohort. High users among the faculty were significantly less likely to be female (19% vs 28%, $p = 0.002$). Mean evaluation scores given by faculty in the high user group were not different from lower frequency users (2.6 ± 0.5 vs 2.5 ± 0.6 , $p = 0.55$). Superusers also did not demonstrate a difference in scores given from lower frequency users (2.6 ± 0.5 vs 2.6 ± 0.6 , $p = 0.75$) Table 4.

Survey Results

Responses to the survey were obtained from 51% of residents (188/367) and 43% of faculty (207/484) at 13 programs. One program that was part of the ABS study (40 faculty, 44 residents) withdrew from the consortium prior to survey development. Seventy-eight residents (38%) were female and 49 faculty (26%) were female, which was not different from the demographics of the actual user group ($p = 0.26$ residents, 0.60 faculty). There was no difference in either resident response rate or self-reported frequency of use by level of training. Among residents, 74 (36%) self-identified as high users reporting that an evaluation was created for at least half of their cases. This was similar to the proportion identified as high users with greater than the average number of evaluations created in the SIMPL database (36% vs 31%, $p = 0.28$). Only 20 residents (10%) reported that they had never used the app. This was significantly

TABLE 3. Frequency of Use by American Board of Surgery Study Participating Faculty

	All Users	Superusers	High Users	Nonusers
# of participants	524	45 (9%)	159 (30%)	156 (30%)
# of evaluations created	range 1-216	range 40-216	range 17-216	range 1-38
Female	147 (28%)	8 (18%, $p = 0.10$)	30 (19%, $p = 0.002$)	46 (29%, $p = 0.63$)
Mean autonomy score	2.6 ± 0.6	2.6 ± 0.5 ($p = 0.75$)	2.6 ± 0.5 ($p = 0.55$)	

TABLE 4. Benefits and Barriers Survey Participant Demographics and Self-Reported Use of SIMPL App

	Residents	Faculty
Survey responders	207	188
Female	78 (38%)	49 (26%)
Self-identified user rate		
High user (at least 50% of cases)	74 (36%)	36 (19%)
Low user (<50% of cases)	113 (55%)	111 (59%)
Never user	20 (10%)	41 (22%)

higher than the nonusers identified from the SIMPL data (4% vs 10%, $p = 0.01$). Among faculty, 41 (22%) self-identified as high users which is significantly lower than the proportion of high users identified from the SIMPL database (22% vs 30%, $p = 0.02$). Thirty-six faculty members (19%) self-identified as never having completed a SIMPL evaluation. This is also significantly lower than the proportion seen in the SIMPL database (19% vs 30%, $p = 0.005$).

Residents and faculty were both asked about the potential benefits of using the SIMPL app on resident education. Residents were most likely to agree that the dictated feedback was valuable. The high user group saw the most value (92% agree, Table 5). The majority of the low user group (76%) agreed there was value to the dictated feedback. The never user group infrequently (30%) felt there was value in the feedback. A similar pattern was seen for the value of the numerical ratings of autonomy with greater than half of the high and low user groups agreeing but only 30% of never users. The difference in groups was larger for some of the other potential benefits. The statements that SIMPL increased the frequency and quality of feedback were positively viewed by >70% of high users but only 42% of low users and <20% of never users. Even among the high users, only a minority (38%) agreed that the use of SIMPL led to an increase in operative autonomy.

The faculty survey showed a similar pattern of responses (Table 5). High users were most likely to agree that SIMPL led to more frequent feedback (76% vs 25% of never users, $p < 0.001$) and that SIMPL evaluations improve resident learning (90% vs 28% of never users, $p < 0.001$).

The most common individual barriers to completing SIMPL evaluations for both the faculty and residents were forgetting and being too busy (Table 6). The residents and faculty also identified each other as significant barriers. Eighty-seven percent of residents across all groups reported faculty did not respond to evaluation requests consistently and 86% reported faculty infrequently dictated feedback. Faculty, on the other hand, reported that residents did not regularly send them evaluations to complete (81% of all faculty respondents).

Qualitative analysis of the free response questions generated 3 main themes. The first theme was those who reiterated the barriers already discussed including blaming the other group for their lack of use, being too busy, or forgetting to use. The second theme was suggested changes that could increase usage including more reminders to use the application, and several respondents even suggested daily cell phone notifications to use the app. Others suggested decreasing the administrative burden by linking SIMPL to a pre-existing resident chore such as case logging to limit the number of things residents need to do on a daily basis. The third theme was suggested changes to the app which would increase value. Technical suggestions included making a favorites list to simplify case entry and allowing written feedback instead of dictated. Both residents and faculty suggested additional questions which could be asked. The residents sought to increase the quality of feedback by asking more specific questions such as “what technical skills should they work to improve?” Faculty also wanted more specific questions about different components of the case (knowledge, technical ability, etc.) to allow them to better identify areas for improvement.

TABLE 5. Identification of Benefits of SIMPL App Showing Number (percentage) of Respondents Who Agree With Each Statement

Resident Survey Question	High Users	Low Users	Never Users	p Value
I value the numerical feedback of SIMPL	56 (76%)	65 (58%)	6 (30%)	<0.001
I value the dictated feedback of SIMPL	68 (92%)	86 (76%)	6 (30%)	<0.001
The frequency with which I receive feedback has increased since SIMPL	50 (70%)	47 (42%)	3 (15%)	<0.001
The quality of feedback I receive has increased since SIMPL	54 (73%)	47 (42%)	4 (20%)	<0.001
My operative autonomy has increased due to SIMPL	28 (38%)	29 (26%)	3 (15%)	0.07
The frequency of verbal feedback has declined since SIMPL	26 (35%)	31 (27%)	3 (15%)	0.18
Faculty Survey Question				
Residents receive more frequent feedback since SIMPL	31 (76%)	35 (42%)	9 (25%)	<0.001
Residents more often ask for feedback since SIMPL	19 (46%)	26 (31%)	7 (19%)	0.04
Resident operative autonomy has increased since SIMPL	20 (49%)	32 (27%)	6 (17%)	0.01
SIMPL evaluations improve resident learning	37 (90%)	52 (62%)	10 (28%)	<0.001

TABLE 6. Identification of Barriers to Use of SIMPL App Showing Number (Percentage) of Respondents Who Agree With Each Statement

Resident Survey Question	High Users	Low Users	Never Users	p Value
Technical issues	28 (38%)	26 (23%)	5 (25%)	0.08
I forget to enter the evaluations	65 (88%)	105 (93%)	9 (45%)	<0.001
I'm too busy to enter the evaluations	54 (73%)	86 (76%)	6 (30%)	<0.001
I feel like it bothers my faculty to have to fill out the evaluations	36 (49%)	59 (52%)	5 (25%)	0.08
Not worth the effort because faculty don't respond consistently	70 (95%)	98 (87%)	12 (60%)	<0.001
Not worth the effort because faculty often don't dictate feedback	68 (92%)	98 (87%)	11 (55%)	<0.001
Faculty Survey Question				
Technical issues	11 (27%)	22 (27%)	9 (25%)	0.90
I forget to enter or respond to the evaluations	29 (71%)	52 (62%)	14 (39%)	0.01
I'm too busy to enter or respond to the evaluations	32 (78%)	51 (61%)	24 (67%)	0.22
Residents are not receptive to feedback	8 (20%)	17 (20%)	11 (31%)	0.52
Residents do not send me evaluations to fill out regularly	30 (73%)	74 (88%)	21 (58%)	<0.001
I have concerns about how this data could be used in legal cases	2 (5%)	8 (10%)	4 (11%)	0.56

DISCUSSION

Analysis of data from the ABS study of SIMPL use shows a group of resident superusers (11%) with more than 1 standard deviation above the mean number of evaluations created. Characteristics of these superusers did not show any gender differences from the users as a whole. The superuser and larger high user groups were more likely to contain senior residents. This may be due to senior residents seeing more value in the evaluations and feedback. However, the more likely reason is that senior residents simply do more operative cases and therefore have a larger denominator of cases performed. Unfortunately, data on total cases performed were not available so it is unclear if the senior residents are evaluating a higher percentage of their cases than the junior residents or evaluating the same or lower percentage. In trying to tease out why superusers are superusers, one could hypothesize that they create evaluations more often because they are stronger residents and therefore more likely to receive positive evaluations. This is not borne out by the data as superusers were evaluated as receiving the same amount of autonomy as lower level users. Use among the faculty followed a different pattern. The percentage of superusers was smaller (9%), and high users were less likely to be female compared to the whole user group. This gender difference could have multiple possible causes including decreased response rate from female faculty (not true), decreased denominator of cases performed by female faculty (no data available), or selection bias where the residents are selectively not creating evaluations for cases performed with female faculty. The last option could be due to unintentional bias or perhaps female faculty more often give face-to-face feedback and the residents do not feel as much need to seek additional feedback. Much like the residents, it could be hypothesized that faculty who give

more autonomy either due to teaching style or case mix would be more likely to fill out evaluations. However, again there was no difference in the average autonomy given between superusers and lower frequency users.

The second interesting group of users is the never users. These are users who despite signing up for the app and receiving training in its use never responded to an evaluation. Never users were rare among residents (4%). There was no difference from the whole user group with respect to gender or level of training. Among the faculty, the never user group was quite large including 30% of all users. Despite female faculty being less likely to be high users, they were equally as likely as males to be never users. On the survey portion of this study, resident never users were well represented (10% of survey participants). Faculty never users were slightly underrepresented on the survey at only 22% of respondents.

The survey showed significant differences between high users, low users, and never users with respect to their perceptions of the benefits of use and barriers to use. Resident users overall saw more value in the dictated feedback than the numerical evaluation of autonomy. This makes sense since they can also evaluate their own autonomy, and previous work shows approximately 60% concordance between faculty and resident evaluations of autonomy.⁶ Autonomy data are more useful to programs for evaluation of resident progress and faculty teaching performance. Dictated feedback, on the other hand, is something residents often struggle to receive on a regular basis. High users were much more likely to see value in either component of the evaluation with 90% of high users valuing dictated feedback. This is in contrast to the never users who saw little value in either component of the evaluation. High users were also much more likely to perceive benefits to the training program as a whole from SIMPL. Nearly three-quarters of high using residents

perceived an improvement in the frequency and quality of evaluations received since using SIMPL in contrast to less than 20% of never users. The same pattern was seen on the faculty survey. High users were much more likely to perceive the benefits of use. One must be careful to note that this is an association and does not imply causation. We would like to think those who use the app become believers and see the benefits of it, therefore, increasing their use. Unfortunately, it is just as likely that users who believed in the premise of the evaluation system prior to its implementation were more likely to use it. The second hypothesis is likely to explain the never users. These may be faculty and residents who made up their minds early that there was inadequate value to the app to encourage them to even try it.

The second focus of the survey was identifying barriers to use. Among the high and lower user groups, there were clearly identified strategies which could be adopted. Two commonly cited barriers were forgetting to create or respond to evaluations and being too busy to add another administrative task to their daily schedule. Respondents made several useful suggestions to address forgetting to use the app including daily reminders which could be incorporated into the app at a predetermined time. Ideally, the app could be tied to the electronic medical record and generate an evaluation automatically once the circulating nurse in the operating room enters that the case is completed. This would eliminate the step of someone needing to intentionally create the evaluation.

Recognition of the additional administrative workload imposed by the app is important. Almost three-quarters of residents, even among high users, reported being too busy to complete evaluations. If the SIMPL evaluation could be combined with another administrative task such as case logging which residents are already mandated to perform, this may increase the percentage of cases for which evaluations are created. The biggest barrier identified by both groups was each other. Faculty reported that residents do not send them evaluations to complete and residents reported that faculty members do not respond to evaluations when they are created. The high percentage (30%) of never users among the faculty gives credence to the residents' perception. The structural changes noted above may increase the percentage of cases in which residents create an evaluation.

This structural approach will not, however, address the never users among the faculty who are not responding to created evaluations. This group of users does not currently perceive value in the app and seems unwilling to try it. Reaching this group will require a more focused effort. Possible strategies could include asking high users to talk with these faculty members, helping residents demonstrate to faculty how much they value the feedback or creating external benefits such as decreasing the length of

end-of-rotation evaluations for those who use the app regularly. Some faculty could be overestimating the amount of feedback they give and see SIMPL as being redundant. But, a study by Jensen et al. showed resident satisfaction with timing, frequency, amount, and specificity of feedback was significantly lower than faculty satisfaction.⁷

CONCLUSIONS

This study of benefits and barriers to use of the SIMPL smartphone app showed that there is a subset of users who are motivated to use the app more frequently. They are characterized by their belief that there is value in the evaluations and use of the app benefits the program as a whole. Even within this group, they noted structural changes which could enhance their frequency of usage. There is also a subset that does not use the app despite receiving requests for evaluation. This phenotype is more likely to exist among faculty members than residents. Different strategies focusing on their preexisting lack of perceived value of the evaluations will need to be addressed to engage them.

REFERENCES

1. Bohnen JD, George BC, Williams RG, et al. The feasibility of real-time intraoperative performance assessment with SIMPL (system for improving and measuring procedural learning): early experience from a multi-institutional trial. *J Surg Educ.* 2016;73:e118-e130. <https://doi.org/10.1016/j.jsurg.2016.08.010>.
2. Williams RG, Swanson DB, Fryer JP, et al. How many observations are needed to assess a surgical trainee's state of operative competency? *Ann Surg.* 2017. <https://doi.org/10.1097/SLA.0000000000002554>.
3. Gundle K, Mickelson D, Hanel D. Reflections in a time of transition: orthopaedic faculty and resident understanding of accreditation schemes and opinions on surgical skills feedback. *Med Educ Online.* 2016;21:30584. <https://doi.org/10.3402/meo.v21.30584>.
4. Williams RG, Chen X, Sanfey H, Markwell S, Mellinger J, Dunnington G. The measured effect of delay in completing operative performance ratings on clarity and detail of ratings assigned. *J Surg Educ.* 2014;71:e132-e138. <https://doi.org/10.1016/j.jsurg.2014.06.015>.
5. Chapman A, Hadfield M, Chapman CJ. Qualitative research in healthcare: an introduction to grounded theory using thematic analysis. *J R Coll Physicians Edinb.* 2015;45:201-205. <https://doi.org/10.4997/JRCPE.2015.305>.

6. Meyerson SL, Sternbach JM, Zwischenberger JB, Bender EM. Resident autonomy in the operating room: expectations versus reality. *Ann Thorac Surg*. 2017;104:1062-1068. <https://doi.org/10.1016/j.athoracsur.2017.05.034>.
7. Jensen A, Wright A, Kim S, Horvath K, Calhoun K. Educational feedback in the operating room: a gap between resident and faculty perceptions. *Am J Surg*. 2012;204:248-255. <https://doi.org/10.1016/j.amjsurg.2011.08.019>.

SUPPLEMENTARY INFORMATION

Supplementary material associated with this article can be found in the online version at doi: [10.1016/j.jsurg.2019.01.012](https://doi.org/10.1016/j.jsurg.2019.01.012).