Medical Students' Views on Peer Assessment of Professionalism

Louise Arnold, PhD,[1] Carolyn K. Shue, PhD,[1] Barbara Kritt, MA,[2] Shiphra Ginsburg, MD, MEd,[3] and David T. Stern, MD, PhD[4]

[1]University of Missouri-Kansas City School of Medicine, Kansas City, Mo, USA; [2]University of Michigan, Ann Arbor, Mich, USA; [3]Faculty of Medicine, University of Toronto, Toronto, Canada; [4]University of Michigan Medical School and the VA Ann Arbor Healthcare System, Ann Arbor, Mich, USA.

BACKGROUND: Although peer assessment holds promise for assessing professionalism, reluctance and refusal to participate have been noted among learners and practicing physicians. Understanding the perspectives of potential participants may therefore be important in designing and implementing effective peer assessment.

OBJECTIVE: To identify factors that, according to students themselves, will encourage or discourage participation in peer assessment.

DESIGN: A qualitative study using grounded theory to interpret views shared during 16 focus groups that were conducted by leaders using a semi-structured guide.

PARTICIPANTS: Sixty-one students in Years 1, 3, and 4 in 2 mid-western public medical schools.

RESULTS: Three themes students say would promote or discourage peer assessment emerged: personal struggles with peer assessment, characteristics of the assessment system itself, and the environment in which the system operates. Students struggle with reporting an unprofessional peer lest they bring harm to the peer, themselves, or their clinic team or work group. Who receives the assessment and gives the peer feedback and whether it is formative or summative and anonymous, signed, or confidential are important system characteristics. Students' views of characteristics promoting peer assessment were not unanimous. Receptivity to peer reports and close positive relationships among students and between students and faculty mark an environment conducive to peer assessment, students say.

CONCLUSIONS: The study lays a foundation for creating acceptable peer assessment systems among students by soliciting their views. Merely introducing an assessment tool will not result in students' willingness to assess each other.

F aced with an urgent need, medical educators have been examining a variety of tools to assess the professionalism of students.[1] One of these tools is peer assessment[1] that, by definition, involves individuals who have attained the same general level of training or expertise, exercise no formal authority over each other, and share the same hierarchic status in an institution.[2]

Peers hold much promise in assessing professional behavior, it is thought, for they may have information that can contribute to self-improvement and enrich summative deci-

sions about promotion and retention.[3–8] Despite its theoretic promise, few successful peer assessment systems have been implemented.[4,9] Instead, reluctance and outright refusal of students, residents, and practicing physicians to participate in genuine peer appraisal have been noted.[10–14] The reluctance and refusal have not been fully explained, and the means to promote genuine participation in peer assessment have not been elucidated. Thus, this study seeks to identify factors that, according to students themselves, would encourage or discourage their participation in peer assessment.

METHOD

The study sites were 2 midwestern state-supported medical schools. At both schools, peer assessment includes student participation in an Honor Council. One school also uses peer assessment in a pharmacology course and 2 clerkships; the other used it in a research study.

Data were collected through focus group discussions as this method allows for open-ended responses, use of probes for further detail, and peer interaction to generate issues and provide an immediate check of validity.[15] Students in their first, third, and final years of medical school were invited to participate in focus groups according to year level. Because of the erratic schedules and responsibilities of students, group size ranged from 2 to 7. Groups were conducted sequentially until thematic saturation was achieved (see Table 1).

At each school, a research associate, experienced in facilitating focus groups, used a semi-structured guide to lead the discussions. Questions (available online) were asked about the actions students take in response to peer behaviors, including whether they would tell anyone about these behaviors, and the conditions that would increase or decrease their willingness to report inappropriate peer behavior. The focus group discussions were taped and transcribed verbatim with names removed by administrative assistants external to the research team.

Relying on grounded theory, the research team used open coding to develop the initial code list and subsequent coding structure.[16] Using NVivo, research associates refined the codes and codebook through frequent consultation, segment-by-segment cross-checking, and negotiation during the entire coding process. The principal investigators adjudicated any uncertainties and disagreements between the research associates' coding decisions. After all modifications of the coding scheme were finalized, the 2 research associates reviewed and updated previously coded transcripts, continued to code additional transcripts, and spot checked

Table 1. Composition and Size of Student Focus Groups Discussing Peer Assessment of Professionalism

Participants' Year Level	School A			School B		
	Number of Focus Groups	Number of Students Per Group	Total Number of Participants	Number of Focus Groups	Number of Students Per Group	Total Number of Participants
Year 1	2	Group A 5 Group B 2	7	3	Group A 7 Group B 4 Group C 2	13
Year 3	3	Group A 4 Group B 4 Group C 2	10	3	Group A 3 Group B 3 Group C 6	12
Year 4	2	Group A 5 Group B 2	7	3	Group A 5 Group B 3 Group C 4	12

each other's coding. A project consultant provided an external check of the coding process. This combination of continuous internal and external auditing strengthened the codes' reliability.

Code reports were used to identify emerging themes and sub-themes. Two members of the research team (L. A., C. K. S.) independently checked the interpretation of themes and sub-themes by reviewing every fifth interpretation. They agreed with 92% of the initial interpretations and resolved discrepancies through discussion.

Whether the themes were discussed by the focus groups at each school and year level was noted. However, systematic analysis of potential differences owing to students' school or year level was not conducted because of small student numbers from each year at each school. The themes, presented in the results section, are illustrated by quotations from the transcripts.

This project received approval from the schools' institutional review boards. All students in the focus groups indicated their willingness to participate by signing an approved consent form.

RESULTS

A total of 61 students in Years 1, 3, and 4 of medical school participated in 16 focus groups at the 2 schools. No group completely embraced or rejected peer assessment. Rather, the groups recognized both the benefits and costs to participating in peer assessment or spontaneously stated that participation in peer assessment would depend on a number of factors surrounding the assessment. The groups then discussed the factors that would promote or discourage peer assessment. These factors can be categorized around 3 main themes: students' struggles with peer assessment, the characteristics of a peer assessment system that would be used, and the environment in which the system would operate.

Students' Struggles with Peer Assessment

In every focus group, regardless of school or year level, students discussed how their struggles with peer assessment would impact their willingness to assess and report the professionalism of their peers.

> It would be nice if a system was set up that wouldn't negatively impact anybody, at least the person making the report . . . and that

there would be an outcome that would be beneficial for making the report in the first place . . . .But to set up a system where everybody was required to sort of participate in kind of a half-assed way, and they're just going to be sort of unwilling to act, in general it's not going to be beneficial. Year 4, School A

> I think there is no question that it absolutely should be brought to someone's attention, but you don't want to be the person that is tattling on someone else either. Because I know I wouldn't want to be responsible for putting something negative in someone's file . . . Year 4, School B

As these quotes suggest, students grapple with the prospect of looking bad themselves if they bring a peer's unprofessional behavior to the attention of someone. They also are concerned for the well-being of peers. They wish to give peers the benefit of the doubt, especially in ambiguous situations where they cannot be sure they are interpreting the behavior of a peer correctly. They worry that something bad will happen to a peer because of a negative report. Additionally, students are reluctant to damage personal relationships by hurting their peer's feelings or incurring a peer's anger. Similarly, they do not wish to disrupt relationships among teammates or members of work groups:

> If it's a personal matter where it's not carrying your share of the workload . . . you're not as likely to say something basically because you don't want to rock the boat. Year 4, School A

Finally, students are uneasy about reporting unprofessional behavior because they feel powerless in influencing peers.

> What's to say they [peers] are going to listen to you. They could just totally blow you off and then it's like what was the point in telling them? Year 1, School B

These personal struggles, then, discourage students from participating in peer assessment.

Characteristics of a Peer Assessment System

Students' preferences for characteristics of a peer assessment system also impact their willingness to participate in peer assessment. System characteristics that were discussed by all focus groups from both schools and at all year levels are: who is involved in the assessment, the use of the assessment, and its mechanics. All the groups but a Year 3 group at School A addressed the content of peer assessment, and all groups but a Year 3 and a Year 4 group at School B discussed anonymity.

Who. Who conducts the peer assessment, who receives it, and who delivers the feedback are issues that could potentially impact students' inclination to participate. Students comment that they would feel more comfortable assessing a junior student than a student in their own class and considerably less comfortable assessing a more advanced student.

Who receives the peer assessment once it is completed is a common concern for students. Although they recognize that it is difficult and potentially damaging to relationships to talk to peers about their behavior, they also believe that direct communication with the peer is a necessary fair first step in reporting behavior. A fourth-year student at School B provided this analogy:

> If somebody's sitting next to you on the bus, and they have something in their teeth, "you say 'you have something in your teeth," and they fix it. You don't tell the guy at the front of the bus that Bob Jones has a piece of celery in his teeth, then he makes a call over the microphone . . . If you'd just tell him, he'll say "thanks, I appreciate that."

However, if the peer is not receptive to the feedback, the behavior does not change, or the behavior is seriously affecting patient care, then faculty or a designated authority figure should be told, as a Year 3 student at School B stated:

> Ideally, first would be to approach your peer, if it's your peer, and say, "Hey, I don't appreciate this or this is not appropriate behavior . . ." Then if the student doesn't change, a resident or an attending should speak to them.

In contrast, some students believe that lodging peer reports with faculty instructors, physicians, and attendings would be counterproductive. They fear that faculty will not act, will sanction them for tattling, and/or will break the anonymity or confidentiality of the report. Other students believe that giving the peer assessment directly to the peer is undesirable. They do not wish to face the peer, especially if they have to work with him/her in the future. Indeed, some students said that faculty need to evaluate students' behavior because it is the faculty's responsibility, and not the students'.

If peers themselves do not provide direct peer feedback, students suggest that residents, faculty, or a council should provide it. Older students offering feedback or a work group sharing its concerns with a peer are other ways that might encourage students to participate in peer assessment. The presence of a mediator might also enable students to feel more comfortable with giving feedback to peers.

Use. Students' preferences for the use of peer evaluations range from a strictly formative purpose with no consequences for a peer to a summative purpose even if peer evaluations might prevent a student from entering a residency program.

Evaluations that provide strictly formative feedback should be directed, most often, to individual peers so they can learn about professionalism, understand expectations for professional behavior, correct inappropriate behaviors, receive reinforcement of their positive behaviors, and grow professionally—as long as no negative consequences ensue for the peer. Indeed, for some students, negative outcomes for the peer constitute a strong deterrent to participation in peer assessment.

Conversely, there are students who feel that peer evaluations should carry consequences. According to these students, if the assessment is negative, it should affect the peer's grades in courses and even graduation.

> . . . Maybe a student would realize that someone's actually paying attention to their behavior . . . that affects their grade and that would be the motivation for them. Year 3, School A

Some students add that negative peer assessments should have consequences only under certain conditions, for example, when a peer does not change his/her unprofessional behavior despite previous warnings or when a peer's lapse is serious.

> We worry about not wanting to affect someone's grade or evaluation or status in school, but if they're doing this [behavior] consistently and multiple [people] are seeing the exact same thing, then something needs to be done. If it comes to the point of them not being a doctor, I'm sorry, but . . . if the student is not able to correct it, then . . . they have no business dealing with patients. Year 3, School B

Mechanics. One aspect of the mechanics of peer assessment that impacts students' participation concerns whether it is informal and optional or formal and required, as this quote illustrates:

> I think if it's really informal, . . . it's a lot easier. Like last week, the first time I was working with a [younger] student, I was telling him every time he did something right or it was very easy if he did something wrong to point out that it is wrong [and] this is how you do it . . . Year 3, School A

Some students believe that informal peer assessment, such as day-to-day comments shared with peers as part of peer teaching and group discussions of peer behavior to improve group functioning, would be facilitative. They comment that informality is most appropriate for formative peer assessment.

Other students strongly contend that peer assessment should be formal and is required to ensure that all students will participate.

> It's important to make it required because it forces everybody to be [involved]. Part of the practice of medicine is continually stopping to assess . . . and reassess yourself, and figure out how much progress you're making. That's what the point of peer evaluations is. Year 3, School B

Still others recommend a combination of informal mid-experience peer evaluation and formal end-of-experience evaluation.

Other aspects of the mechanics of peer assessment that students believe would encourage their participation in the process concern length (the shorter the better), specificity (the more concrete the questions and peer comments the better), and computerization.

Content. To participate in peer assessment, students believe that they should evaluate specific behaviors in domains in which they have expertise. These behaviors include whether peers arrive promptly for class, do their fair share of work, and display teamwork—all especially germane to professionalism and related to access that students have to observe a variety of peer behaviors—and exclude many of the technical aspects of medicine.

With respect to the types of behavior contained in peer reports, students prefer that peer assessment focus on both positive and negative behaviors of peers. As a Year 1 student at School A shared:

> . . . it's great to hear positive feedback on what you do well, but it is also really helpful to hear . . . honest feedback about "you really could improve this," and . . . "here are some specific things you can do to make it better."

Other students favor reporting only negative peer behaviors, particularly offensive, serious, or consistent problems, to convey to peers when their actions are unprofessional.

> But I think if it's going to be a peer sort of negative stuff, it's got to be the big stuff and it's got to go to town. You can't go back and forth about whether somebody is late. Year 1, School B

Some students believe that a focus on negative behaviors of peers would do little to promote peer assessment. They find formal reporting of peers' professional lapses difficult because of potential untoward consequences that the report might have for peers and for themselves. Moreover, the subjective nature of the assessment of professionalism contributes to students' discomfort in reporting negative peer behavior. Unclear questions on an assessment form can compound their uneasiness.

Other students suggest that peer assessments should focus only on positive professional peer behavior. Reports of positive behavior, these students say, inflict no harm, give students a boost, are easier to write, and reinforce appropriate behavior.

> . . . it's important to try to give positive feedback to reinforce what we should do, not just what we shouldn't do . . . it's just encouraging, I think for people to realize that it is affecting other people, when they are taking that extra effort to do things . . . , that it's having an impact and being noticed. Year 1, School B

A few students believe that a focus on only positive professional behavior would discourage participation in peer assessment. They think reports of positive behavior are unnecessary, for positive behavior already influences students' grades.

Anonymity. Anonymity is an important, frequently mentioned way to encourage genuine participation in peer evaluation. Students believe that anonymity can protect both the student evaluator and the peer who is being evaluated. In particular, the evaluator who lodges a negative anonymous report does not have to face the peer directly, an uncomfortable prospect for many. Moreover, the student evaluator cannot be accused of tattling, an action not readily condoned. Students also contend that anonymity may reduce the prospect of disrupting teamwork and interpersonal relationships while promoting more candid honest assessments.

Students who favor anonymity express several concerns about totally anonymous peer assessment. The prospect of retaliation raises students' anxieties about anonymity. They think that the assessment system should incorporate oversight to verify information contained in the report about a peer, hold student evaluators accountable for their reports, and avoid student venting and vendettas.

> I feel like anything that goes on your record that's a statement of your work or your . . . professional behavior, [and] should be accounted for . . . .If people are going to evaluate you, then it should not be a system in which students can say anything about anyone and not be responsible for what they say. Year 3, School A

Some students predict that peers will not take anonymous evaluations as seriously as those that are signed. They also say that anonymous evaluations do not prepare them well for a future task of confronting peers face to face when they are in practice. Finally, these students believe that they could thank peers for writing a positive evaluation if it were signed.

A confidential system may be an acceptable compromise, students suggest. In their eyes, confidentiality offers the prospect of verification and thus accountability but retains the anonymity of the student evaluator.

Environmental Factors

Every focus group in each school and at each year level considered environmental factors that would encourage students' participation in peer assessment, including the school's stance toward peer reports, relationships among students, faculty, and administrators, the school's evaluation philosophy and standards, and its educational program for peer assessment and professionalism.

Students emphasize that a school's responsiveness to peer reports is key to promoting peer assessment. The presence of trustworthy as well as receptive people designated to receive a peer report is another critical environmental characteristic:

> But it depends partly on, like if you know the people who you're reporting it to, if you can trust them to actually do something, or to be able to make a decision . . . So I guess that's one advantage of having students in the honor council. Year 1, School A

Close relationships among students and faculty and among students themselves are also part of an environment conducive to peer assessment. Students say that positive relationships with faculty can promote productive discussions about peer behavior and more generally about professional and unprofessional behavior. The presence of student groups whose members are amenable to teaching each other and to exploring professionalism issues can contribute to a collaborative and productive critique of peer performance, in contrast to groups where competition, protectionism, and trench mentality discourage honest assessment.

The philosophy that undergirds the school's approach to evaluation also impacts students' acceptance of peer assessment. Evaluation aimed at improvement would be especially helpful, students comment. They also believe that the presence of clear standards of professionalism and consistent enforcement are fundamental to the success of peer assessment even though they recognize the difficulty of delineating behavioral expectations for learners.

Education for peer assessment, a particular aspect of the institutional environment, would promote willingness to assess peers, students remark. They also say that education must enable students to learn the meaning of professionalism and specify the expectations that faculty hold for professional behavior. Topics for peer assessment education should also include conflict resolution and training in giving and receiving feedback. Students ask that the process of peer assessment be clearly described to them.

> . . . it should be made clear to the students there's an actual process you go by, and it is unacceptable to all of a sudden . . . write a bad peer evaluation and expect that to go straight into someone's file. Year 3, School B

Students also believe that peer assessment education should highlight its importance and usefulness, especially via faculty role modeling. Some students mention that the act of participating in peer assessment in medical school, in itself, would prepare them for the peer assessment that they will be expected to conduct when they are in practice.

> . . . If we don't take action now on those kinds of things [inappropriate behavior] there's not a whole lot of chance we ever

will . . . .it's hard to confront people. And if you get in a habit of not doing it, and just overlooking things that bother you about other people's behavior and not even mentioning it to them, then I think you get to a point where you just don't really care about how your colleagues act. Year 1, School A

DISCUSSION

Focus group discussions identified and illuminated a number of factors that could promote as well as deter peer assessment. These factors go beyond students' personal struggles with reporting peer behaviors and include characteristics of a peer assessment system itself and the environment in which the system would operate.

Examination of these factors suggests that successful peer assessment entails more than introducing a peer assessment instrument. Who receives the information, its use, and the issue of anonymity emerge as system characteristics particularly influential in determining students' willingness to participate in peer assessment. Generally, students prefer informal peer-to-peer feedback for formative purposes, except when frequent or egregious unprofessional behavior is the issue. The prospect of confidential, rather than totally anonymous, reports that can be verified could also promote participation in peer assessment.

The educational environment plays a significant role in students' willingness to assess peers, just as it does in the development of students' professionalism generally.[17] Students contend that administrators, faculty, and students themselves should be appropriately responsive to peer reports of unprofessional behavior. Student teams or groups open to honest critique would provide a particularly welcoming setting for peer assessment.

Choosing system characteristics that address student concerns and take into account key environmental factors is a logical first step in designing a system to assess professional behavior of peers. For example, a system guaranteeing confidentiality or one used only for formative purposes might ease students' personal struggles with peer assessment and promote their participation.

Building students' preferences into a peer assessment system may not be easy or straightforward. For each potential system characteristic, there seem to be contrary views within individual students and among students who articulate arguments for their particular positions. The tensions in students' perspectives on who should receive their peer assessments, what uses peer assessment will have, and whether peer assessments should be anonymous are especially complex and deeply felt.

How to resolve students' opposing preferences remains a critical issue for our continuing work. The stance that an institution takes toward professionalism, its philosophy of evaluation, and the education it offers students about professionalism and peer assessment may help to resolve some of the tensions among students' ideas about peer assessment.

This study is limited by the same issues of range, depth, and candor of discussion characteristic of any focus group design. Moreover, this paper suggests, but leaves for further analysis, the difficulties likely encountered with any generalized prescriptive application of a peer assessment system. It may be, for example, that because of their reliance on aspects of the educational environment, different optimal systems of peer assessment would necessarily exist in different settings. Identifying the variability in school environments and student perceptions of peer assessment across multiple schools and year levels is the focus of the authors' further research. Likewise, the variability in faculty perceptions and endorsement of professional behaviors is not known. Finally, the question of whether the concept of professionalism that students use in assessing each other relates only to their studenthood or maps conceptually to professional behavior of physicianhood deserves exploration.

In the meantime, this study has laid a foundation for creating acceptable peer assessment systems for medical students. Drawing upon the perceptions of students in different curricular years at 2 schools with different uses of peer assessment, this research has amassed a rich set of conditions that may encourage students to assess each others' professionalism. More specifically, it has identified a set of choices that must be faced with respect to the environment surrounding the assessment and the characteristics of the assessment system itself to address students' concerns with peer assessment. If effective peer assessment can be established, it may offer further insight into students' professional lapses that, when identified by faculty, have been linked to later risk of censure by state licensing boards.[18]

REFERENCES

1. **Arnold L.** Assessing professional behavior: yesterday, today, and tomorrow. Acad Med. 2002;77:502–15.
2. **Arnold L, Stern DT.** Content and context of peer assessment. In: Stern DT, (ed.) Measuring Medical Professionalism. London: Oxford University Press; 2005:(forthcoming).
3. **Montgomery BM.** An interactionist analysis of small group peer assessment. Small Group Behav. 1986;17:19–37.
4. **Small PA, Stevens CB, Duerson MC.** Issues in medical education: basic problems and potential solutions. Acad Med. 1993;68(suppl): 89S–98S.
5. **Hundert EM, Douglas-Steele D, Bickel J.** Context in medical education: the informal ethics curriculum. Med Educ. 1996;30:353–64.
6. **Ginsburg S, Regehr G, Hatala R, et al.** Context, conflict, and resolution: a new conceptual framework for evaluating professionalism. Acad Med. 2000;75(suppl):6S–11S.
7. **Ramsey PG, Wenrich MD, Carline JD, et al.** Use of peer ratings to evaluate physician performance. JAMA. 1993;269:1655–60.
8. **Norcini JJ.** Peer assessment of competence. Med Educ. 2003;37:539–43.
9. **Arnold L, Willoughby L, Calkins V, et al.** Use of peer evaluation in the assessment of medical students. J Med Educ. 1981;56:35–42.
10. **Linn BS, Arostegui M, Zeppa R.** Performance rating scale for peer and self-assessment. Br J Med Educ. 1975;9:98–101.
11. **Thomas PA, Gebo KA, Hellmann DB.** A pilot study of peer review in residency training. J Gen Intern Med. 1999;14:551–4.
12. **Van Rosendaal GMA, Jennett PA.** Resistance to peer evaluation in an internal medicine residency. Acad Med. 1992;67:63.
13. **Rennie SC, Crosby JR.** Students' perceptions of whistle blowing: implications for self-regulation. A questionnaire and focus group survey. Med Educ. 2002;36:173–9.

14. **Friedson E.** Profession of Medicine; a Study of the Sociology of Applied Knowledge. New York: Dodd, Mead, and Co.; 1970.

15. **Morgan DL.** The Focus Group Guidebook: Focus Group Kit 1. Thousand Oaks: Sage Publications; 1998.

16. **Glaser BG, Strauss AL.** The Discovery of Grounded Theory: Strategies for Qualitative Research. New York: Aldine De Gruyter; 1999.

17. **Inui T.** The Flag in the Wind: Educating for Professionalism in Medicine. Washington, DC: Association of American Medical Colleges; 2003.

18. **Papadakis MA, Hodgson CS, Teherani A, Kohatsu ND.** Unprofessional behavior in medical school is associated with subsequent disciplinary action by a state medical board. Acad Med. 2004;79:244–9.

Supplementary Material

The following supplementary material is available for this article online:

Appendix. Examples of Questions posed for Focus Group Discussions of Peer Assessment of Medical Students' Professionalism.