Professional identity formation refers to the process by which medical trainees develop and internalize their new roles. In this work, we analyze medical student evaluations of teaching (SETs) as a window into students’ developing identities as physicians. Our data consisted of 389 open-ended comments written anonymously by first-year (pre-clerkship) students in mid- and end-of-semester evaluations of small group sessions (mandatory attendance) during one full academic year at Yale School of Medicine. Using a combination of existing frameworks on professional identity formation, the purpose of this project was to: (1) describe the characteristics of comments made by medical students about first-year courses and instructors; (2) categorize the student comments; and (3) explore the usefulness of comments as markers of students’ professional identity formation as physicians. Having established baseline information, we hope to follow the same cohort of students through their medical school career to assess if and how their evaluative comments shed light on the development of their professional identities as physicians.

INTRODUCTION

There is widespread recognition in the scholarly literature that student evaluations of teaching (SETs) are an important yet limited set of tools for assessing faculty and courses. While SETs are ubiquitous in higher education, many researchers have called into question their usefulness as a measure of teaching effectiveness [1-4]. Evidence suggests that students can dependably convey their subjective experience in a course or teaching session (e.g., how well they could hear and understand the instructor, usefulness of lecture notes or slide presentations, and availability of the teacher outside of the teaching space) but cannot credibly report on aspects of teacher or
teaching effectiveness, such as the teacher’s knowledge of the field [5]. Further, demographic and personal characteristics of teachers have been shown to influence SETs [6,7].

While most research has focused on the validity and reliability of closed-ended (multiple choice) SETs, many teaching evaluations in medical education also include opportunities for students to submit narrative comments. Reciprocally, clinical preceptors are invited to provide both closed- and open-ended feedback about trainees’ performance. Several studies have used narrative comments by medical trainees as a source of information about how trainees perceive their instructors [8]. Other research in medical education using narrative evaluations as a data source has examined whether preceptor comments are useful for clerkship students [9], and one study has investigated the ways in which faculty interpret comments about residents that are written by other faculty members [10].

In this work, we explore what SETs may reveal about medical students’ developing identities as physicians. We establish baseline information using comments from first-year medical students in hopes of subsequently following this cohort’s comments about teachers and teaching during medical school. Our hypothesis for this longer-term research is that at the beginning of medical school, students assess teaching more from the standpoint of consumers of education, but that as they progress they begin to develop professional identities and that their SETs will increasingly reflect the values, beliefs, and expectations of professional members of an academic medical community. We do not minimize the importance of medical students as consumers; students have a right to hold teachers accountable to investments (financial, time, intellectual) that students have made. Rather, our aim is to evaluate if and how SETs might inform the medical education community of ways in which students perceive and perform their progressing role as members of this professional community.

Professional Identity Formation in Medicine

Professional identity formation in this context refers to the process by which newcomers to medicine develop and internalize their professional roles. Cruess and colleagues [11,12] define a physician’s identity as “a representation of self, achieved in stages over time during which the characteristics, values, and norms of the medical profession are internalized, resulting in an individual thinking, acting, and feeling like a physician” [11]. Although no authoritative list of “characteristics, values, and norms” exists, commonly cited values of the medical profession include compassion, intellectual honesty, confidence, forthrightness, and respect [13-15].

Several frameworks are available for understanding professional identity formation in medical education. Kegan [16,17] explains that medical trainees move through distinct stages of professional identity formation, ranging from “acting” in the role of doctor without a broader understanding of the role to actually becoming physicians after having internalized the expectations, behaviors, and values of the profession and built a system of values and internal processes critical to the physician role [18]. Lewin and colleagues note that achieving an identity as a physician is a formal goal of medical education, and that it is also a “dynamic process shaped by and intertwined with the development of that person’s larger adult identity [19].”

Professional identity formation in medical education has been assessed using tools and methods that include surveys [20,21], interviews [22], narrative inquiry [23,24], structured professional identity essays [25], and expressive activities such as a mask-making exercise [26]. Each of these methods seeks to capture to what extent medical students understand and personally identify with the values, beliefs, attitudes, and expectations of the medical profession at their current stage of training. Using Kegan’s framework as a backdrop to inform how such an identity formation might occur incrementally over time has been a useful tool as we explore how students’ narrative comments may be used as indicators of these developmental stages.

Professional Feedback in Medical Education

Giving and receiving feedback are core components of medical education. Formative feedback, defined by Ende [27] as “information that a system uses to make adjustments in reaching a goal,” is most effective when both giver and receiver are working “as allies, with common goals.” In medical education, the overarching goal is to provide trainees with the skills, knowledge, and dispositions necessary to advance to the next stage of medical training, and ultimately, to practice medicine independently. Throughout their training, medical students receive feedback from many sources: professors, attending physicians, other clinical professionals such as nurses and physician assistants, residents and interns, standardized patients, research mentors, and even peers. But it is not a one-way street; students are also invited to provide feedback to medical educators, working from the shared goal of improving the learning experience. More advanced trainees such as interns, residents, and fellows continue to receive feedback from attending physicians and mentors, but they are also responsible for delivering feedback to medical students and others. Attending physicians are expected to give feedback to trainees, and they in turn receive feedback from department chairs and
section chiefs. Even outside of academic medicine, interactions between physicians and patients often involve establishing shared goals, then giving and receiving feedback to achieve them.

Considering the importance of feedback to medical education and to the practice of clinical medicine, we believe that delivering feedback is a central aspect of a physician’s professional responsibility; therefore, learning how to deliver feedback effectively is a key component of a physician’s professional identity formation. One of the first settings in which students are invited to provide feedback—and thus begin to develop their own competencies as professional givers of feedback—is through narrative SETs. When students provide feedback in this way, they bring many perspectives reflecting their identities as young adults, consumers, learners, (future) physicians, and members of the professional community of academic medicine. As students internalize what it means to be a part of this community, one might expect that the feedback they provide would also come to resemble feedback given by a physician to a trainee or patient. Ende offers one model for effective professional feedback in medicine via eight guidelines:

1. Feedback should be undertaken with the giver and receiver working as allies, with common goals
2. Feedback should be well-timed and expected
3. Feedback should be based on first-hand data
4. Feedback should be regulated in quantity and limited to behaviors that are remediable
5. Feedback should be phrased in descriptive non-evaluative language
6. Feedback should deal with specific performances, not generalizations
7. Feedback should offer subjective data, labeled as such
8. Feedback should deal with decisions and actions, rather than assumed intentions or interpretations

At our institution, we introduce these characteristics of effective feedback to first-year students during medical school orientation. This orientation to feedback is reinforced throughout medical school by further training and opportunities in simulated and actual educational and clinical settings.

To our knowledge, our study is the first to use narrative SETs as a window into professional identity formation. SETs are an appealing source of data since they are already collected for other purposes and are therefore accessible without placing a burden on students. Additionally, the anonymous nature of SETs allows students to express themselves candidly without fear of retribution, offering faculty and administrators a unique perspective on student beliefs and values. Further, SETs are a rich source of information about professional identity formation because of the entanglement between feedback and a physician’s professional responsibilities. While other methods of assessing professional identity formation, such as structured professional identity essays, are intentionally self-reflective, the exploration of SETs is valuable because they provide, often inadvertently, examples of self-presentation or communication that students implicitly deem appropriate in an academic exchange of views.

Research Purpose

The purpose of our research is to:

1. Describe the characteristics of comments made by medical students about first-year medical school courses and instructors
2. Drawing on existing frameworks of professional identity formation, categorize descriptions of student comments
3. Explore their usefulness as markers of students’ professional identity formation as physicians

Our aim is to establish a baseline understanding that can be built upon in future research in which we follow the same cohort of medical students over time to learn if and how their SETs evolve in character and content. Ultimately, with this work, we hope to begin a conversation among medical educators regarding the usefulness of narrative SETs and student feedback generally as a gauge of professional identity formation.

MATERIALS AND METHODS

Our dataset consisted of narrative comments written anonymously by first-year medical students in mid- and end-of-semester evaluations of mandatory attendance workshops during one full academic year at Yale School of Medicine. Workshops represent class-based activities in multiple courses and subjects, and mandatory attendance workshops form a subset of didactic sessions that covers material deemed essential to the curriculum. All mandatory workshops require interactive participation of students and impart critical thinking and/or clinical reasoning skills. They engage the same faculty for each session in a block of workshops, and each section has at most 13 students.

These workshops follow educational activities that explore the meaning of the physician’s professional identity. These include a two-week course titled “Introduction to the Profession” with lectures, clinical experiences, patient simulations, and faculty-student discussions. As part of this transitional program, students participate in a “White Coat” ceremony designed to affirm their change in status and welcome them into the ranks of the profession. Students are also introduced to characteristics of effective feedback, as described above.
Mandatory workshops are distributed across ten periods of evaluation throughout the academic year; for each period in which mandatory workshops were taught, half of the students (50-55 in a total class size of approximately 105 students) were invited to complete an anonymous course evaluation survey, which included both closed- and open-ended questions about the workshops that took place during the period of evaluation. (The other half of the class was invited to complete evaluations of lectures.) Although students are required to submit evaluations to gain access to mandatory self-assessments and exams, they may leave questions blank. In total, 389 open-ended comments about mandatory workshops were received over the course of one academic year. Since the number of comments exceeds the number of students in the class, it is evident that some students submitted multiple comments, but the deidentification of the dataset makes it impossible to determine who wrote which comments.

To analyze our data, we employed a grounded theory approach, which Watling and Lingard [28] recommend for exploratory research in medical education that seeks to develop a theory that is “grounded” in the data. We (all three authors) began by open coding a randomly selected subset of 72 comments, representing 18.5% of the full dataset. We compared our codes to identify common themes and areas of difference, then organized our codes into a preliminary codebook. Next, two authors (EBW, DS) used the preliminary codebook to code a second randomly selected subset of 39 comments while making note of any comments that were not adequately described by the existing codes. The coding authors compared findings, then refined the codebook by adding, combining, and modifying codes as needed to better capture the range of student comments. The same two authors then used the refined codebook to code a third random subset of 20 comments, including a new category for indicating the perceived level of professional identity formation displayed by the comment using a scale from 1 (least professionally mature) to 5 (most professionally mature). During this third round of coding, the number of notes (memos) about difficulties encountered in the coding process had decreased markedly, suggesting that saturation had been reached. However, our ratings of professional maturity showed relatively poor agreement (weighted Kappa = 0.22, p = 0.0238) perhaps due to the small number of comments in the random subsample representing the high and low “extremes” of professional identity formation (e.g., only three ratings of 1 were assigned between the two coders, and one coder did not assign any ratings of 5).

Since the small number of comments made it difficult to identify salient features of comments at the extreme ends of the observed professional identity formation spectrum, we used “deviant” or “extreme” case sampling, wherein unusual examples are intentionally selected for analysis. Drawing on our expertise as medical educators and professionals, we discussed characteristics of the comments we had coded that seemed to display lower-than-typical, higher-than-typical, and typical levels of professional identity formation. Following that discussion, one author (EBW) reviewed the full dataset and purposively sampled 50 comments that exemplified different levels of professional identity formation, drawing on existing interpretive frameworks—including Kegan’s model [16,17,25], Ende’s characteristics of effective feedback [27], Mezirow’s levels of critical reflectivity [29]—to inform the selection of comments exemplifying each level.

Characteristics of the purposive sample of 50 comments, as well as the full sample of 389 comments, are displayed in Table 1. The purposive sample—composed of 17 lower-than-typical level, 22 typical level, and 11 higher-than-typical level comments—included comments that were exclusively affirmative and exclusively critical, as well as those that contained both affirmative and critical elements. One author (EBW) coded the purposive sample using the refined codebook and identified the sentiment (critical, affirmative, or both) of each comment. Finally, we organized the comments by estimated level of professional identity formation and looked for patterns in the codes applied. All three authors discussed and reached consensus on implications and interpretation of the data.

### RESULTS

Our analysis began with an iterative coding process, initially focused on describing characteristics of “typical” comments from first-year medical students—those that fall in the center of the professional identity formation

### Table 1. Characteristics of the full sample and purposive sub-sample.

|                        | Full Sample | Purposive Subsample |
|------------------------|-------------|---------------------|
| No. of comments        | 389         | 50                  |
| No. of courses represented | 5           | 5                   |
| Total no. of words     | 13,525      | 2,616               |
| Mean no. (SD) of words per comment | 34.7 (29.2) | 52.3 (37.6) |
observed other patterns (the “miscellaneous” domain) that included reporting the opinions and attitudes of other students in contrast to one’s own experiences, using superlatives (e.g., “the best,” “great,” “awesome”), and making suggestions about how the course or teaching could be improved. While comments pertaining to each

Table 2. Codebook developed from open coding process.

| Course Characteristics | • Materials (readings, etc.) and content  
|                       |   ○ Appropriate level of difficulty  
|                       |   ○ Interesting  
|                       |   ○ Relevant  
|                       | • Other  
|                       |   ○ Workshop pedagogy  
|                       |   ○ Connected to the rest of the course  
|                       |   ○ Contribute to my learning  
|                       |   ○ Good use of my time  
|                       |   ○ Scheduling conducive to learning  
|                       |   ○ Student enjoyment of workshops  
| Teaching Behaviors    | • Communication skills  
|                       |   ○ Answers questions effectively  
|                       |   ○ Clear explanations of material  
|                       |   ○ Communicates relevance  
|                       |   ○ Concise  
|                       |   ○ Gives good feedback  
|                       | • Facilitation skills  
|                       |   ○ Creates safe/positive learning environment  
|                       |   ○ Encourages student participation  
|                       |   ○ Highlights important points  
|                       |   ○ Lets students lead discussion  
|                       |   ○ Manages digression effectively  
|                       | • Organization  
|                       |   ○ Attentive to prior student knowledge  
|                       |   ○ Brings in supplemental material  
|                       |   ○ Ends class on time  
|                       |   ○ Pacing (amount of time spent on each topic) conducive to learning  
|                       |   ○ Prepared for class  
| Personal Qualities    | • Approachable  
|                       | • Cool  
|                       | • Enthusiastic, upbeat, happy  
|                       | • Helpful  
|                       | • Interesting as a person  
|                       | • Knowledgeable  
|                       | • Passionate  
|                       |   ○ About material/content  
|                       |   ○ About teaching  
|                       | • Patient  
|                       | • Puts in effort, tries hard  
|                       | • Smart  
| Other                 | • Affect  
|                       |   ○ Affirmative  
|                       |   ○ Critical  
|                       |   ○ Mixed (affirmative and critical)  
|                       | • “Good teacher/bad course”  
|                       | • Makes a suggestion  
|                       | • Reports opinions of peers  
|                       | • Uses superlatives and platitudes  

bell curve based on the full spectrum of first-year medical student narrative SETs. The codebook we developed is displayed in Table 2. Our coding process revealed that student comments typically addressed one or more of three domains: course characteristics, teaching behaviors, and personal qualities of the instructor. We also
of the domains appeared across the spectrum of levels of professional identity formation, we began to notice differences in the frequencies with which we had applied certain codes to comments that we felt displayed a relatively sophisticated level of professional identity formation as opposed to those which reflected earlier stages of professional development. These differences, which we explored more thoroughly after the selection of the purposive subsample, are described in subsequent sections.

Although our refined codebook captured the content and sentiment of the SETs in our dataset, as indicated by the marked decrease in the coder memos written during the analysis of the third random subsample, we felt that drawing on other frameworks for understanding and interpreting the comments might help us further articulate the other characteristics of student comments reflecting typical, lower-than-typical, and higher-than-typical levels of professional identity formation. Examining the comments in relation to Ende’s [27] characteristics of effective feedback, in particular, resonated with our subjective assessments of which comments seemed to be the least and most professionally mature. While the Kegan stages of professional identity formation [16,17,25] and levels of critical reflectivity [29] initially seemed promising as ways of interpreting the student comments, we found that most comments were simply too short or limited in scope to interpret and code within these frameworks. Furthermore, it was difficult to identify specific values or virtues of the medical profession [13-15] since most comments more directly addressed the workshop leaders’ strengths and weaknesses as teachers rather than as medical professionals (cf. “Physician as Teacher” vs “Physician as Physician” [8]).

Below, we use an amalgamation of interpretive frameworks to describe the characteristics of student comments from narrative evaluations of mandatory workshops, drawn from the purposive subsample of 50 comments. We first present “typical” cases, followed by “extreme” cases representing both comments that we coded as showing impressive progress towards identity formation as physicians (“higher-than-typical”) as well as those showing less evidence of professional identity formation (“lower-than-typical”). Comments below are presented with minor edits for length and clarity and have been anonymized to ensure the privacy of both students and faculty involved. Table 3 presents a summary of representative comments across the observed spectrums of professional identity formation and sentiment (exclusively critical, both critical and affirmative elements, or exclusively affirmative).

**Typical Comments**

A majority of the comments represented what we came to identify as a “typical” level of professional identity formation among first-year medical students. In general, we found that the comments reflected respondents’ various (and sometimes competing) needs, values, and identities as young adults, learners, and (future) physicians, while displaying several characteristics of effective feedback [27]. Although some comments offered generalizations rather than identifying particular behaviors, most students mentioned one or more specific ways in which the workshop leaders were effective (or ineffective) as teachers for them at this stage in their education as medical professionals. Students often began their comments with superlatives (“the best,” “amazing”) before offering supporting details:

“[The instructor] was absolutely incredible. His enthusiasm for the science being discussed in the papers was very contagious, and he often nuanced our understanding of the papers with a more thorough and disciplined take on the science based on his own knowledge.”

“[The instructors] were fabulous workshop leaders who engaged with students while also presenting the information in a clear and accessible manner.”

“[The instructor] has been an outstanding workshop leader. I really enjoy how he leads them. We talk through each sign/symptom and what it rules in or rules out or moves on our differential. It is incredibly helpful to think that way, and it’s how I’d like all workshops to be.”

Students frequently commented on the perceived relevance, value, or usefulness of material covered in workshops, and particularly appreciated when workshop leaders focused on high-yield material and provided “clinical pearls.”

“[The instructor] was great at explaining the diseases and was also great at giving clinical correlations and clinical pearls. He also kept well to the workshop time and pace and diverted unnecessary/irrelevant questions from students very well.”

“Wednesday afternoons are so much more bearable because of [the instructor]. She makes sure to cover everything we need, and she actually teaches us useful things. Also, very appreciative that she is respectful of our time and manages to finish on time every session.”

“[The instructor] does a good job of bringing in only the most relevant information; he really does let us lead most of the session ourselves.”

Personal and affective characteristics that contributed to a positive learning experience also appeared frequently:

“[The instructor] is simply the best! She is whip smart and really nice and cool and just a great representation of how one balances obscure [basic science] knowledge with real-life clinical skills.”

“[The instructor] is an absolute gem to have as a leader. She is so incredibly kind, thoughtful, and always makes everyone feel warm and welcome. There is never
"We don’t really have much of a discussion, just answer the questions on the sheet. It would be nice to have more of a journal club style discussion. [The instructor] is very nice and cares about the students."

"[The instructor] is a great workshop leader. I like when he repeats certain points over numerous weeks. I sometimes get the idea he doesn’t want to teach the workshops though and would rather leave and get back to his other job."

**Extreme Comments**

We identified several comments that stood out as reflecting elements of highly developed professional identities—at a “higher-than-typical” level for first-year medical students. In these comments, students revealed a cognizance of something larger than their own individual learning and satisfaction with the course and teaching—namely, a sense of responsibility to contribute to the teaching and learning community by critically assessing course content, pedagogy, and teaching effectiveness as a way of perfecting these elements for future classroom activities. Such comments mentioned specific teaching

| Critical only | Both affirmative and critical elements | Affirmative only |
|---------------|----------------------------------------|------------------|
| **Lower-than-typical** | "Pretty useless, why were these mandatory?" | "[The instructor] was the absolute best. The information that we gained, however, could have easily been provided via PDF. Meeting at 8am for 4 days was not necessary at all." | "[The instructor] is super interesting as a person and good at explaining things." |
| | "I honestly learn [sic] nothing." | |
| **Typical** | "[The instructor]...was unable to clearly communicate concepts and move through material in an effective manner..." | "[The instructor] is extremely knowledgeable about almost every possible topic... I think he could improve by remaining on topic so that we can at least cover all the cases." | "[The instructors] were amazing at making the concepts easy for a first year medical student to understand. I enjoyed [the class] so much more because of them." |
| **Higher-than-typical** | "I’m not getting much out of the workshops with [the instructor] so far unfortunately. He clearly knows a lot and I’d be happy for him to teach us more, but that hasn’t been happening a lot. I wish our section was more helpful in putting all the ideas together to help synthesize them, because that seems like it would be the most beneficial part of the workshops." | "I really liked the way [the instructor] did a broad overview and review of the topic for the day before jumping into the workshop cases... The second session of some of the back-to-back workshops felt slightly rushed, but he made sure to go over the highlights we missed at the beginning of the next session." | "I really appreciated that [the instructor] asked for feedback in the middle of the workshops so she could adjust real time to our feedback and that she actually incorporated our suggestions in her next session." |

a dumb question, and she makes sure that we are challenged to come up with ideas on our own, while supporting us along the way.”

In some cases, students were able to acknowledge their own perspective as learners (e.g., as first-year medical students) or as individuals (e.g., as people with particular interests or backgrounds that might differ from their peers):

"[The instructors] were amazing at making the concepts easy for a first-year medical student to understand. I enjoyed [the class] so much more because of them."

"[The instructor] is very passionate about scientific inquiry and it shows, as he always made sure to provide additional context to the papers we were reading. I enjoyed the content of the workshops, but I’m biased because I was a [subject] major in college."

Comments balanced praise and accolades with critical feedback or suggestions for improvement:

"[The instructor] is extremely knowledgeable about almost every possible topic in [the field]. I think he could improve by remaining on topic so that we can at least cover all the cases."
behaviors and qualities of the learning environment, such as allowing students to lead discussion, creating a classroom environment that is welcoming of questions, communicating the clinical relevance of material, and emphasizing the most important information:

“The quality of these workshops... is excellent. [The instructor] gives the students in our workshop the freedom to lead the discussions and answer each other’s questions without much interruption, jumping in only when clinical or deeper basic science knowledge is necessary to understand the answers. I think this creates a more collaborative environment in these workshops and definitely helps the students learn more efficiently.”

“[The instructor] is an amazing facilitator. She challenges our thought and ensures that we are highlighting important topics in the paper. She also gives really good and helpful feedback.” [This student also pointed out the benefit of having an instructor from a demographic background which had previously been underrepresented in workshops.]

“I really liked the way [the instructor] did a broad overview and review of the topic for the day before jumping into the workshop cases. These were nice refreshers and he did a very good job of focusing on the most important aspects of the topic at hand. The second session of some of the back-to-back workshops felt slightly rushed, but he made sure to go over the highlights we missed at the beginning of the next session.”

Some comments recognized the instructor and course at a point in time, capable of improvement, or acknowledged that students also play a role in defining and shaping the class dynamic:

“I really appreciated that [the instructor] asked for feedback in the middle of the workshops so she could adjust real time to our feedback and that she actually incorporated our suggestions in her next session.”

“[The instructor] is a kind and friendly instructor that encourages discussion and makes everybody feel comfortable during the sessions. The questions in the workshops since last evaluation did feel conducive of more scientific debate and discussion.”

“I wish more of my classmates were willing to participate and contribute to the discussion. I often felt sorry for the workshop leader because the group was so silent. Thank you for your patience!”

When critical or making a suggestion, students were able to define the impact on themselves or other students, and working from shared goals of maximizing student learning:

“I’m not getting much out of the workshops with [instructor] so far unfortunately. He clearly knows a lot and I’d be happy for him to teach us more, but that hasn’t been happening a lot. I wish our section was more helpful in putting all the ideas together to help synthesize them, because that seems like it would be the most beneficial part of the workshops.”

“The energy of my particular section is subdued. We tend to have answers displayed on the board - I wonder if the other groups that worked out problems on the board were more energetic?”

“[The instructors] were well-prepared for the workshop and had made a PowerPoint for the workshop session. I felt that it was a good use of time to learn about [the topic], but I know many of my peers in the workshop group felt that it got repetitive and that reading the assigned papers was enough.”

At the other end of the continuum, comments that we identified as reflecting a “lower-than-typical” level of professional identity formation tended to read more like anonymous online posts than feedback coming from medical professionals. Frequently, these comments included general platitudes without supporting detail, and comments about personal attributes that were not directly related to the faculty member’s effectiveness as a teacher, role model, or medical/scientific professional. Also common were disparaging remarks about the course and course materials without recommendations for how these aspects of the course could be improved. Examples of comments that were brief or vague, lacking specific details to support their authors’ claims, include:

“I had [instructor name], I think that she did a good job.”

“I honestly learn [sic] nothing.”

“No specific critiques, just never found [this subject] to be very helpful.”

“[The instructor] is super interesting as a person and good at explaining things.”

“I like [the instructor]. He is very sincere.”

Of note is that these comments could be perceived as affirmative, critical, or neutral, but their common trait is that they contain no information that would lead recommendations or changes that would enhance the values and mission of a practice community—in this case, a community of medical educators and learners. Some lower-than-typical level comments included statements about the mandatory nature of the workshops with no input for the instructor or course director to enhance or improve aspects of the workshops or course that are within their purview to adjust. In brief, they gave evidence of few or no qualities of effective feedback. For example, students wrote:

“Pretty useless, why were these mandatory?”

“I have [instructor]. While the course is interesting, I personally don’t think it should be made mandatory.”

Interestingly, some comments at both extremes of the professional identity formation spectrum shared a particular formula in which a positive statement about the teacher was followed by a complaint or critique about
the workshop or overall course. We categorized feedback following this formula as “good teacher/bad class.” A “lower-than-typical” example in which comments were very general and less useful for understanding needed improvements or possible enhancements, includes:

“The instructor is very kind and does her best with the material she is given. Unfortunately, I think this course is extremely poorly designed for our learning and therefore she isn’t able to be very successful. It is not her fault by any means, however.”

At the other end of the continuum, while “good teacher/bad class” feedback was still present, the “good teacher” part included details about specific teaching behaviors or relevant teacher characteristics, rather than just a superlative or platitude:

“The instructor is a great workshop leader who always brings in charts and worksheets to help us with understanding the concepts that we are covering. He is also able to step back and allow the students to lead the conversation, only intervening when there are questions that the student leaders cannot answer. However, I think the workshops themselves aren’t super useful. We don’t have enough time to cover the paper that we have to read and go over questions for discussion, so we don’t get to explore the paper in depth.”

DISCUSSION

Our review of SETs reveals perspectives that are consistent with our expectations of first-year students and with the referenced models of professional development. In alignment with Kalet [25] our findings suggest that most first-year medical student comments reflect awareness of the behaviors, values, and expectations characteristic of the medical profession, although these behaviors, values, and expectations are not yet fully internalized. This may be an indication of a staged development of professional identity discussed by Kegan [16-18], but because most of the comments were relatively brief (52 words long on average) and of limited scope, it was difficult to interpret and code comments solely within the full five-stage framework that Kegan provides. Nonetheless, we likely will find Kegan’s model useful for our longitudinal study of SETs over students’ medical school career, and therefore, we want to keep it accessible for future analysis. We ultimately determined that Ende’s characteristics of feedback [27] offered a more helpful set of standards by which to assess the student reviews. As argued by Ende, comments likely giving evidence of professional development assume a set of shared goals between teacher and trainee and are based on “first-hand data.” Moreover, these comments make few specific points and are limited to remediable behaviors; they are generally descriptive and most subjective data is labeled as such; and they cite actions or decisions rather than assume intentions. In this analysis, we distinguished sentiment (critical vs affirmative) from level of professional identity development. Put differently, well-reasoned and insightful comments may have been highly critical of the course or the instructor, and poorly developed feedback was sometimes quite complimentary.

Most of the SET data suggest that student comments are constructive, with notable results at the extremes. Some comments at the “higher-than-typical” level of professional identity formation indicate an unexpected maturity and represent an aspirational set of attitudes for first-year medical students. The feedback at this level takes a holistic view of the course and of students’ own level of knowledge and places these elements in the context of the competing social demands of small group discussions: the ability of teachers to put students at ease, to encourage intellectual exploration, and to even assess the participation of classmates.

Nonetheless, some SET data at the “lower-than-typical” level highlight the continued need for attention to professional development. These comments bring into question the level of awareness that respondents have regarding the purpose and usefulness of feedback in a professional context. Again, as per Ende, feedback in the medical school context should always be in reference to the shared goal of performance improvement or enhancement. On the other hand, and in fairness to such respondents, medical schools do not always make clear the purpose of course and teacher evaluations. They should explain how such feedback will be used and give evidence that it is actually used to improve teaching and the curriculum. Without such guidance, it is easy for students to treat SETs—typically administered electronically and anonymously—in the same way as comments on social media. If SETs are used in the future as a tool to help assess professional identity formation of a cohort of students, they should be done so with clear direction from educators regarding the purpose of SETs.

We believe SETs are an important tool in the medical curriculum and represent one way to assess the professional growth of students. Course evaluations also provide an opportunity for students to treat them—typically administered electronically and anonymously—in the same way as comments on social media. If SETs are used in the future as a tool to help assess professional identity formation of a cohort of students, they should be done so with clear direction from educators regarding the purpose of SETs.
ularly for new teachers. Development of medical student comments away from such unhelpful content and towards that which is substantively useful for performance improvement may well be one marker of professional identity formation.

Limitations

As with any research, there are factors that may limit the generalizability of our findings. First, the data we analyzed (i.e., narrative SETs) were not originally collected for the purpose of measuring professional identity formation. This is both a strength and weakness. Unlike the professional identity essay used by Kalet and colleagues [25] to assess medical students' identity formation, comments about workshops drawn from mid- and end-of-semester course evaluations are neither a systematic nor complete representation of a student's professional identity. Throughout the coding process, we noted that it was difficult to disentangle evidence of students' developing professional identities from their ability or willingness to provide written feedback about their instructors in this setting.

Second, our study considered data from one medical school cohort during one academic year at a single institution; therefore, findings are not generalizable. In qualitative research, however, the primary aim is to obtain a deeper understanding of a given phenomenon, rather than producing generalizable results [30]. Further, the exploratory nature of our study means that the credibility and dependability of our method of analyzing student comments as a measure of professional identity formation are unknown. While the procedures and frameworks laid out in this paper may be helpful for medical school administrators and faculty who would like to explore how SETs might shed light on typical levels of professional identity formation for trainees at different stages in their medical education, we do not currently recommend using this method to judge or grade the professionalism or professional identity formation of individual students. Future work that is more systematic in nature and considers additional student contexts would be necessary to determine the usefulness of SETs as a proxy for more formal ways of assessing professional identity formation.

CONCLUSIONS AND FUTURE WORK

This exploratory work has shown that SETs are a fertile yet complicated source of information about medical students’ professional identities as physicians. In addition, there is a degree of “candidness” in anonymous SETs that might be useful in shedding light on competing interests, needs, and motivations stemming from medical students’ multiple identities (e.g., consumer, learner, professional) and their negotiation of professional identity formation during the first year of medical school. Indeed, in some sense, the very usefulness of our data is predicated on the fact that students do not know that they are being observed in this way.

To develop this approach and our understanding of its usefulness, we propose the following. First, make certain that students are given guidance about the schools’ expectations for them as contributing members of a medical community as they comment on teaching and curriculum. Next, students should be given training in ways of providing constructive feedback in SETs prior to being asked to evaluate courses. Already at our institution, early in the first year of medical school (and in addition to the training offered in orientation mentioned earlier), students are offered training on Ende’s model as a framework for delivering feedback about courses and teaching as well as for gauging the quality of feedback they are receiving from educators. This training is voluntary, and over the last few years, 30-35% of each first-year class has chosen to participate. In the future, the training may be a required component of an existing professional development course and may also provide more examples of written evaluation comments that are most actionable within a teaching-learning context.

As subsequent research steps, it would be useful to “follow” the first-year cohort involved in this study, continuing to assess their comments in learning activities during their medical school curriculum. Additionally, our research would benefit from extending our assessment to multiple cohorts over several academic years to assure a greater credibility and dependability of our methodological approach. Another future approach is to develop a more systematic method of assessing narrative SETs, e.g., by adapting the Narrative Evaluation Quality Instrument (NEQI) [31] for this purpose.

REFERENCES

1. Boring A, Ottoboni K, Stark P. Student Evaluations of Teaching (Mostly) Do Not Measure Teaching Effectiveness. ScienceOpen Research [Internet]. 2016 Jan 7 [cited 2020 Jan 6]. Available from: https://www.scienceopen.com/hosted-document?doi=10.14293/S2199-1006.1.SOR-EDU.AETBZC.v1
2. Uttl B, White CA, Gonzalez DW. Meta-analysis of faculty’s teaching effectiveness: student evaluation of teaching ratings and student learning are not related. Stud Educ Eval. 2017 Sep;54:22–42.
3. Hessler M, Pöpping DM, Hollstein H, Arnemann PH, Massoth C, et al. Availability of cookies during an academic course session affects evaluation of teaching. Med Educ. 2018 Oct;52(10):1064–72.
4. Wallace SL, Lewis AK, Allen MD. The State of the Literature on Student Evaluations of Teaching and an Exploratory Analysis of Written Comments: Who Benefits Most? Coll Teach. 2019 Jan;67(1):1–14.
5. Hornstein HA. Student evaluations of teaching are an inadequate assessment tool for evaluating faculty performance. Cogent Education. 2017 Jan;4(1):1304016. https://doi.org/10.1080/2331186X.2017.1304016.

6. Smith BP, Hawkins B. Examining Student Evaluations of Black College Faculty: Does Race Matter? J Negro Educ. 2011;80(2):149–62.

7. Freng S, Webber D. Turning Up the Heat on Online Teaching Evaluations: Does “Hotness” Matter? Teach Psychol. 2009 Jun;36(3):189–93.

8. Lim DW, White JS. How Do Surgery Students Use Written Language to Say What They See? A Framework to Understand Medical Students’ Written Evaluations of Their Teachers. Acad Med. 2015 Nov;90(11 Suppl):S98–106.

9. Lye PS, Biernat KA, Bragg DS, Simpson DE. A pleasure to work with—an analysis of written comments on student evaluations. Ambul Pediatr. 2001 May-Jun;1(3):128–31.

10. Ginsburg S, Regehr G, Lingard L, Eva KW. Reading between the lines: faculty interpretations of narrative evaluation comments. Med Educ. 2015 Mar;49(3):296–306.

11. Cruess RL, Cruess SR, Boudreau JD, Snell L, Steinert Y. Reframing medical education to support professional identity formation. Acad Med. 2014 Nov;89(11):1446–51.

12. Cruess RL, Cruess SR, Boudreau JD, Snell L, Steinert Y. A schematic representation of the professional identity formation and socialization of medical students and residents: a guide for medical educators. Acad Med. 2015 Jun;90(6):718–25.

13. Levoy B. Traits of the ideal doctor. Veterinary Economics. 2007 Jul;48(7):17.

14. Notzer N, Soffer S, Aronson M. Traits of the ‘ideal physician’ as perceived by medical students and faculty. Med Teach. 1988;10(2):181–9.

15. Pellegrino ED. Professionalism, profession and the virtues of the good physician. Mt Sinai J Med. 2002 Nov;69(6):378–84.

16. Kegan R. The Evolving Self. Harvard University Press; 1982. 340 pp.

17. Kegan R. In Over Our Heads: The Mental Demands of Modern Life. Harvard University Press; 1994. 420 pp.

18. Jarvis-Selinger S, Pratt DD, Regehr G. Competency is not enough: integrating identity formation into the medical education discouse. Acad Med. 2012 Sep;87(9):1185–90.

19. Lewin LO, McManamon A, Stein MT, Chen DT. Minding the Form That Transforms: Using Kegan’s Model of Adult Development to Understand Personal and Professional Identity Formation in Medicine. Acad Med. 2019 Sep;94(9):1299–304.

20. Houseknecht VE, Roman B, Stolfi A, Borges NJ. A Longitudinal Assessment of Professional Identity, Wellness, Imposter Phenomenon, and Calling to Medicine Among Medical Students. Med Sci Educ. 2019 Jun;29(2):493–7.

21. White MT, Borges NJ, Geiger S. Perceptions of Factors Contributing to Professional Identity Development and Specialty Choice: A Survey of Third- and Fourth-Year Medical Students. Ann Behav Sci Med Educ. 2011 Mar;17(1):18–23.

22. Vivekananda-Schmidt P, Crossley J, Murdock-Eaton D. A model of professional self-identity formation in student doctors and dentists: a mixed method study. BMC Med Educ. 2015 Apr;15(1):83.