Major Compulsory Revisions

1. I’m not sure that your title aligns with the real purpose of the study. In the introduction you state that you want to better understanding entering residents’ perspectives on professionalism and you add the bit about gender and institutional differences almost as an afterthought, it certainly doesn’t seem to be central to your study. I also note below that there are problems with your treatment of the data in regards to testing for differences in gender and institution. I’ll also argue that your most compelling data is the relationship between residents’ professionalism ratings and their participation in those behaviors. I’d consider rewriting to either remove the focus on gender and institution or substantially reducing its importance in the paper.

2. I think you can do a better job at providing reliability and validity evidence for the survey you used. You make a couple of vague statements in the methods section about your survey being based on another survey previously published and that the vignettes align with the authors’ experience. Just because a survey has been previously published, doesn’t mean it is a high quality survey or a survey that believably answers your research questions. I’d like to see some discussion of the reliability coefficients you observed during this study. At minimum, I would have liked to see you look to the literature and/or a panel of experts to determine if your survey has representatively sampled the domain of unprofessional behaviors.

3. I don’t generally think it is appropriate to use hypothesis tests on each individual item on an instrument. I think it weakens the validity of the instruments and of the conclusions drawn from the statistical analysis, particularly when a correction for type I error inflation is not used. Please see the paper cited below for a more complete discussion on the problems in analyzing item-level data.


4. I’m not well versed in the Cochran-Mantel-Haenszel test statistic, but based on my brief reading, I have doubts as to its appropriateness for this data. I would advise having an expert statistician review.

5. I’m not sure what it means that entering residents from the two institutions had significant differences. I’d be more interested in seeing if there were differences between graduating learners of different institutions. Are you trying to make the
case that the two different institutions have different standards in terms of professionalism in their selection process or that the professionalism of learners who favor attending one institution over another is different? Either way, there are better ways to go about addressing those questions. I'm not sure that this set of findings adds a lot of value to the literature.

6. Your qualitative results are rather cursory. I think you could have provided much more detail or at minimum, some examples of comments that illustrate your categories. You did a nice job of explaining your qualitative data analysis, but then provide extremely little in your results section.

7. Page 9, end of first complete paragraph, I think these are all likely contributing factors. I also think that sometimes people just make mistakes, recognize them and then correct future behavior. It would have been interesting to ask how frequently the participants reported engaging in the behaviors.

8. I was surprised to see the participants list "balancing home and work life" as a professionalism issue. I would have really liked to see some of the comments that participants made. I haven't come across this idea in my own research in professionalism, nor have I read about in others' research. I think it is intriguing and would have liked the authors to explore it in more depth.

Minor Essential Revisions

1. First sentence, second paragraph of the introduction needs a period.

2. First full sentence on page 5 ("The survey included..."), I found difficult to read, I had to read it a couple of times to understand what you were trying to say.

3. Page 5, first full paragraph, you mentioned that residents were informed about that the study was voluntary twice. I think you could reword to remove the redundancy.

4. Page 5, last full paragraph, you are missing a decimal point, “p-value of 05” should read, “p-value of .05.”

5. I would like to see percentages, in addition to, or instead of frequencies in table 2.

6. I’d like a bit of explanation as to why you chose to focus on the top 15 behaviors to compare the professionalism ratings with whether or not a respondent participated in the behavior.

7. Bottom of page 8, “50% of responders rating” should read “50% of responders rated.”

8. Role modeling is very likely a contributing factor as to why the respondents participated in the behaviors in the first place, so improving role modeling could very well serve to reduce the behavior's incidence; however, it is much easier to say that role models need to act more professionally than it is to actually change behavior. Your statement on page 9, "may be all that is required" minimizes the difficult task of getting role models to change their behaviors. I think a revision that doesn’t minimize the significant task of changing physician behavior is warranted.
9. You had a couple of references that were incomplete; 16, 40.
10. Figure 1 needs some editing, the percentages are jumbled together or overlap with the bars in a few places.

Discretionary Revisions

1. Page 4, second complete paragraph, first sentence hyphenate “day to day.”
2. I can see why you briefly describe the two institutions at the end of the introduction, but I would rather see the description in the methods section.
3. I don’t think Table 1 adds much to the paper, you already do a nice job of explaining the response rates in the text, I suggest you consider deleting Table 1.
4. Page 6, after you discuss Table 2 you reiterate that you collapsed response categories, you already discussed that above, you don’t need to reiterate it here.
5. The title of Figure 1 is awkward; consider “Behaviors in which 10% of respondents participated”
6. You seem to reference the millennial generation quite frequently in the paper. In my opinion, these generational differences are often overblown and we tend to see much more variability within a generation than between generations. I suggest that you cut back on your references to the millennials.

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Needs some language corrections before being published

**Statistical review:** Yes, but I do not feel adequately qualified to assess the statistics.

**Declaration of competing interests:**

I declare that I have no competing interests.