

RESEARCH ARTICLE

Pimping in Bedside Clinical Education: Questioning Style of Attending Physicians and Associations with Learners' Evaluations

Authors

Madeline Rodriguez, MD; Sean Tackett, MD, MPH; Scott M. Wright, MD; John W. McEvoy, MB, BCh, MHS, MEHP.

Affiliation

M. Rodriguez is resident, Johns Hopkins Bayview Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland

Sean Tackett, MD, MPH is assistant professor of medicine and director, International Medical Education, Division of General Internal Medicine, Johns Hopkins Bayview Medical Center, Baltimore, Maryland; ORCID: <https://orcid.org/0000-0001-5369-7225>. Stacket1@jhmi.edu

S.M. Wright is professor, Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland.

J.W. McEvoy is personal professor of preventive cardiology, School of Medicine, National University of Ireland, Galway, Ireland

Correspondence:

Madeline Rodriguez, Johns Hopkins Bayview Department of Medicine, Johns Hopkins University School of Medicine, 301 Mason Lord Drive, Baltimore, MD, 21224; telephone: (410) 550-3350; email: mrodri68@jhmi.edu

ABSTRACT

Purpose. Given the interest and momentum in medical education on psychological safety and the learning environment, we sought to evaluate whether there was an association between questioning style of attending physicians and their corresponding evaluations from learners.

Methods. Building on our prior study that developed a *pimping scale* to categorize the questioning styles of attending physicians on inpatient teaching services at two university-affiliated hospital internal medicine residency programs, this cross-sectional study examined the association between a physician's score on our scale and that physician's learner evaluations. For our primary outcome, we used the overall summative rating from teaching evaluations at both residencies, and for secondary outcomes we picked two questions from each residency's evaluation felt to be most characteristic of pimping behaviors.

Results. Including both residency programs, the range of evaluation values for the "overall attending rating" was 2-4 (potential range 1-4), and the mean for all faculty was 3.7. There was no correlation between attendings' pimping scores and their overall resident rating ($p = 0.28$). At the Osler Residency, there was no correlation between attendings' pimping scores and residents' ratings for respectfulness ($p = 0.96$) or supportiveness ($p = 0.93$). At the Bayview Residency, there was no association between pimping scores and students rating of the attendings propensity to either ask questions in non-threatening way ($p = 0.62$) or explore complications and errors without intimidation ($p = 0.78$)

Conclusion. *Pimping* questioning behaviors do not appear to be correlated with teaching evaluations of attending physicians by house staff. This may be due to the limited scope and fidelity of current teacher evaluations. It is also possible that some residents have come to accept pimping-type questioning as a cultural norm in medical education.

INTRODUCTION

In medical education, pimping refers to the process by which an attending asks a series of questions directed to a trainee, escalating beyond the scope of knowledge expected for their level, which serves to maintain the “hierarchy of academics” and evoke humility, if not humiliation.¹⁻³ While the published literature has focused on this occurrence involving attending physicians and trainees – likely because attendings provide summative evaluation – educational pimping-type questioning can transpire across any teaching dyad where a power differential exists. This practice in the socialization of medical trainees predates Brancati’s seminal article “The Art of Pimping” in 1989, and continues to be a part of the medical vernacular. The apprenticeship model remains a key educational paradigm in medicine,⁴ with the mentor as superior and in a position of authority over the mentee; this hierarchy is reinforced explicitly, and more subtly as part of the hidden curriculum. Over the years, many have advocated for the utility of pimping.^{1-2, 5-9} Proponents point to values including maintaining deference to more experienced clinicians, being pushed to expand foundational knowledge, practicing medical decision-making under stressful conditions, and having critical deficiencies in clinical reasoning identified by physician teachers. More recently, some have suggested that the terminology and practice are antiquated, even inappropriate, and that pimping as a descriptive term and humiliating practice should be eradicated entirely.³

While there are some studies that have attempted to study pimping,^{5, 15} it remains unclear how graduate medical learners perceive the pimping experience. Further, it is not known whether pimping by attending physicians influences learners’ evaluations of these teachers. This study builds on our prior work regarding pimping in clinical education, whose aim was to create an instrument - a summative pimping score – that identified a pimping phenotype based on self-reported behaviors and attitudes.¹⁰ The objective of the

current study was to analyze the associations between attendings’ pimping scores and trainees’ evaluations of their teaching. We hypothesized that attendings with higher pimping scores (reflecting higher predilection to pimping-type questioning style) would have lower cumulative ratings on resident evaluations. We thought that trainees would not appreciate the negative impact of pimping on their psychological safety or across the horizontal team dynamic.

1. METHOD

1.1 Study Design, Subjects, and Setting

For this cross-sectional study, we collected and analyzed the residents’ evaluations of department of medicine faculty teaching from two large academic medical centers, Johns Hopkins Hospital and Johns Hopkins Bayview Medical Center. Both locations provide clinical education to medical students and post-graduate medical education trainees (interns, residents, and fellows), and each has a distinct internal medicine residency training program. The faculty members were surveyed about their behaviors and attitudes related to pimping, which culminated into the Pimping Scale. This scale is an 11-item questionnaire that has been shown to have content, response process, internal structure, and relations to other variables validity evidence.¹⁰ Three items ask about frequency of pimping behaviors with responses along a 6-Point Likert scale (1=Never, 6=Always) and 8 items ask about level of agreement with statements about pimping beliefs along a 4-point Likert scale (1= Strongly disagree, 4=Strongly agree). Total scores could range from 11-50, with higher scores indicating a greater affinity to pimping. Pimping scores were generated for all faculty who had attended for ≥ 2 weeks in 2014 and 2015 on any of the following 3 teaching services at either hospital: general medicine wards, medical intensive care units, or cardiac intensive care units.

1.2 Data Collection

The initial survey that categorized attending physicians on the pinging scale was conducted in the second half of 2016.¹⁰ Learner evaluations of these same attendings, completed for their inpatient teaching activities with internal medicine residents from a similar time period, were collected using 'Evaluate' (Minneapolis, MN).¹¹ Because the dataset containing the attendings' pinging score information was de-identified, the linkage of these pinging score data with the Evaluate teaching evaluation data was performed by an independent party; all data was de-identified and at no point did the study team have access to attending identifiers. Attendings that had no evaluation data were excluded from the study.

Given that the residents belong to two distinct internal medicine residency programs within the same medical school, their faculty evaluations were similar but not identical; both teaching evaluations ended with an "overall attending rating." While the primary analysis was to focus on the association between pinging score and the overall rating of teaching performance, we also explored the relationships between two questions from each of the hospital's faculty evaluations that were believed to be most linked to the pinging behaviors. From the evaluation of faculty at Johns Hopkins Hospital, the residents rated along a 9-point Likert scale to what extent the attending was "Respectful, e.g. courteous, punctual" and "Supportive, e.g. approachable, patient, empathetic." From Johns Hopkins Bayview, residents rated the extent to which the attending "asked questions in a non-threatening way" and "explored complications and errors without intimidation" along a 5-point Likert scale. The study protocol was approved by the Johns Hopkins University School of Medicine Institutional Review Board.

1.3 Data Analysis

Demographics and other baseline variables were compared among the faculty

respondents using analysis-of-variance, Kruskal-Wallis, or Chi-square testing, as appropriate. Pinging scale totals were calculated by adding responses to all items (with one item reverse coded). For each attending, means were calculated across all evaluations, so that there was a single mean for each item for each attending. Means and standard deviations for resident evaluation items that were unique to JHH or JHBMC were calculated separately. For the outcome variable that the programs had in common, the JHH 9-point scale was normalized to a 5-point scale before means and standard deviations were calculated. Associations between pinging scale totals and resident evaluation items were determined using Pearson correlations. Pinging scale total quartiles were also created and ANOVA testing examined if there were differences in item means across the quartiles. Due to the variation in the number of evaluations completed for each attending, we performed a sensitivity analysis excluding those attendings with fewer than the median number of evaluations; this did not result in meaningful differences. All analyses were performed using STATA-13 (Stata, College Station, TX).

2. RESULTS

Of the 149 eligible faculty, 125 responded to the questionnaire related to pinging (84% response rate). Of these 125 faculty, 112 had at least one evaluation. The mean number of evaluations per faculty was 9.2 (SD 9.6). The mean age of the faculty sample was 46 years, 61% were male, and 70% were white. The median number of weeks per year that faculty reported serving as inpatient teaching attending was 6, with an interquartile range from 3-12 weeks. Approximately 41% of respondents reported they were general internal medicine physicians and 49% self-identified as clinician educators, Table 1. Demographic characteristics for resident respondents on E-Value is not collected, and thus not presented.

Table 1. Baseline demographic characteristics for the 112 faculty attendings; data shown according to self-reported behaviors and attitudes related to pimping

	All N = 112	Low Pimping Score N = 62	High Pimping Score N = 50	P-value
Age in years, mean (SD)	45.8 (9.8)	46.4 (10.0)	45.2 (9.6)	0.536
Male, n (%)	68 (60.7)	31 (50.0)	37 (74.0)	0.010
Race, n (%)				
White	78 (69.6)	43 (69.4)	35 (70.0)	0.849
Black	6 (5.4)	4 (6.5)	2 (4.0)	
Asian	23 (20.5)	13 (21.0)	10 (20.0)	
Hospital, n (%)				
JHH	52 (46.4)	23 (37.1)	29 (58.0)	0.027
JHBMC	60 (53.6)	39 (62.9)	21 (42.0)	
Appointment, n (%)				
Investigator	47 (42.0)	28 (45.2)	19 (38.0)	0.671
Educator	53 (47.3)	27 (43.6)	26 (52.0)	
Other	12 (10.7)	7 (11.3)	5 (10.0)	
Service, n (%)				
Wards	87 (77.7)	50 (80.7)	37 (74.0)	0.062
CCU	11 (9.8)	8 (12.9)	3 (6.0)	
MICU	14 (12.5)	4 (6.5)	10 (20.0)	
Weeks on service, mean (SD)	8.6 (7.9)	8.8 (8.6)	8.4 (7.2)	0.775
Number of evaluations, mean (SD)	9.2 (9.6)	9.8 (10.8)	8.4 (7.9)	0.440
Positive attitude toward pimping, mean (SD)	46 (41.1)	14.0 (22.6)	32.0 (64.0)	0.000
Pimping score total, mean (SD)	24.1 (5.4)	20.3 (3.1)	28.8 (3.7)	n/a

2.1 Distribution of the summative pimping score

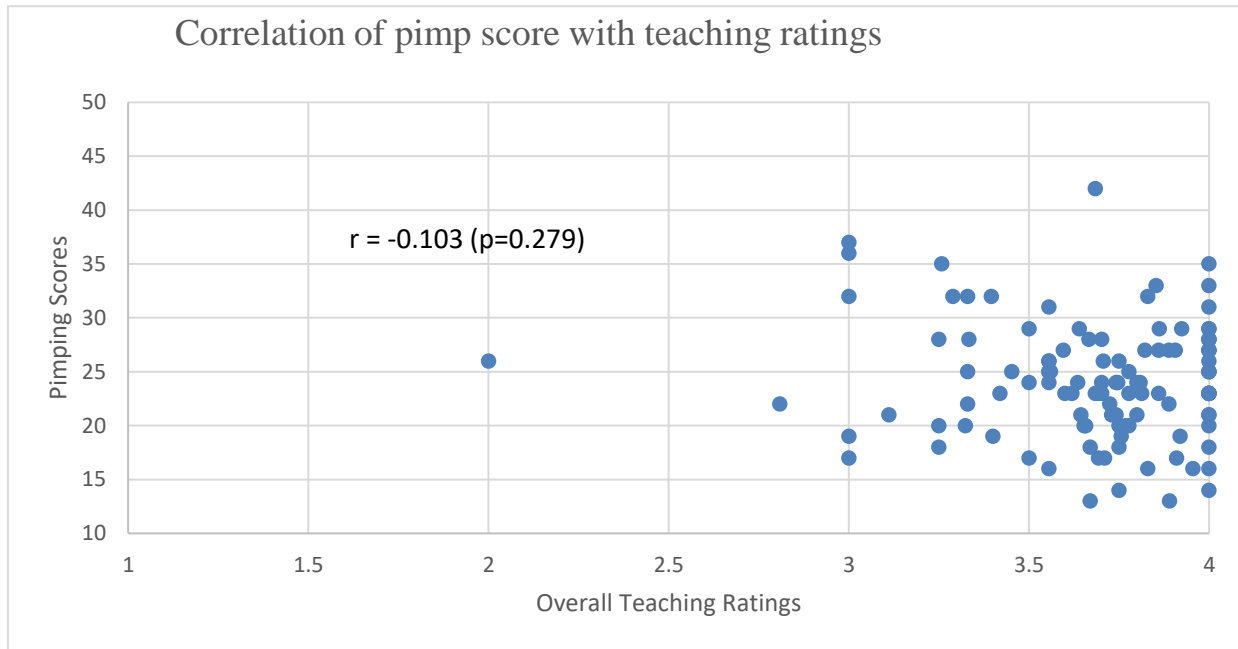
The range of values for the summative pimping score derived from faculty responses was 13 – 42. The quartile cutoffs for the scores were as follows: first quartile 13-20 (n=29), second quartile 21-24 (n=33), third quartile 25-27 (n=23), and fourth quartile 28-42 (n=27), Table 2.

2.2 Learners’ evaluations of the faculty teachers and correlations with pimping scores

The range of values on the evaluations for the “overall attending rating” was 2-4

(potential range 1-4) and the mean for all faculty was 3.7; 26 (23%) attendings had an average overall rating of 4. The mean overall rating for attendings with pimping scores corresponding to quartiles 1-4 are as follows: 3.7 (SD 0.3), 3.7 (SD 0.3), 3.7 (SD 0.4), and 3.6 (SD 0.3), Table 2. There were no differences between attendings’ overall ratings across the pimping scores quartiles (p=0.825), Figure 1. Pearson correlation coefficient was -0.103 (p=0.279).

Figure 1: Scatter plot showing both the distribution of ‘Pimping Scores’ and mean ‘Overall Evaluation Ratings’ for all 112 attending physicians



At Johns Hopkins Hospital, the range of values for the “was my attending respectful?” question was 7-9 (potential range 1-9). There was no correlation between attendings’ pimping scores and ratings for respectfulness ($r=-0.0217$, $p=0.8718$), or any difference in means across pimping score quartiles ($p=0.955$). The range of values for the “was my attending supportive?” question was 7-9 (potential range 1-9). Attendings’ supportiveness likewise had no correlation ($r=-0.0563$, $p=0.6745$), or difference across pimping score quartiles ($p=0.933$), Table.2.

At Johns Hopkins Bayview, the range of values for “asked questions in a non-threatening way?” was 2-4 (potential range 1-4). Attendings with lower pimping scores were not rated to ask questions in a less non-threatening way ($p=0.618$). The range of values for “explored complications and errors without intimidation?” was 2-4 (potential range 1-4). There was not an association between pimping scores and exploring complications and errors without intimidation ($p=0.782$), Table 2.

Table 2. Data from the learner evaluations about the teaching attendings categorized by pimping quartiles

	Pimping Quartile #1 N = 29	Pimping Quartile #2 N = 33	Pimping Quartile #3 N = 23	Pimping Quartile #4 N = 27	P-value
Pimping Score, mean (SD)	17.6 (2.2)	22.7 (1.1)	26.0 (0.9)	31.2 (3.5)	n/a
Overall Teaching Rating, mean (SD)	3.7 (0.3)	3.7 (0.3)	3.7 (0.4)	3.6 (0.3)	0.83
Respectfulness, mean (SD)	8.4 (0.6)	8.4 (0.4)	8.5 (0.6)	8.4 (0.5)	0.96
Supportiveness, mean (SD)	8.4 (0.6)	8.4 (0.3)	8.4 (0.4)	8.4 (0.5)	0.93
Asked questions in non-threatening way, mean (SD)	3.8 (0.2)	3.6 (0.3)	3.6 (0.6)	3.6 (0.5)	0.62
Explore complications/errors without intimidation, mean (SD)	3.7 (0.3)	3.6 (0.3)	3.5 (0.6)	3.6 (0.6)	0.78

DISCUSSION

This study examines the association between attending physicians’ self-reported pimping behaviors and their summative teaching evaluations by internal medicine residents across inpatient settings at two academic teaching hospitals. Notably, our data did not show a correlation between these attendings’ characterization of their own pimping attitudes and behaviors (pimping score) and learners’ teaching evaluation of attendings.

Much of the published literature about pimping in medicine are commentaries and perspective pieces that delve into its history, appropriateness, and acceptance by learners.^{1-2, 5-7, 12-13} Previous studies have assessed learners’ characterization of pimping behaviors as generally good or bad, but did not explore pimping’s effects on other educational outcomes - such as attendings’ teaching evaluations. Of note, most of these prior discernments have been solicited from medical students.^{5, 14-15} This study represents the first published work to describe the characterization of attending pimping behaviors by graduate medical learners, and also offers insight into

whether attending physicians with ‘pimping-questioning styles’ are assessed differently by trainees on their summative teaching evaluations.

The evaluations of teaching attendings completed by learners in this study - from two respected residency training programs - are likely similar to those used at most hospitals. The assessments address multiple domains such as role modeling, provision of feedback, compassion towards patients, and respectfulness of all, in addition to an overall summative rating. In looking to correlate pimping scores with the overall teacher assessments, as well as other questions, we saw just how upwardly skewed the responses are. With the mean overall score being 3.7 out of 4, and so many attendings being rated the top box (4 out of 4), it becomes apparent that these evaluations are of limited use and certainly impractical for research purposes and statistical correlations. The skewed responses may indicate that all attendings included in this study are equally appreciated, or that pimping-type questioning is not weighted heavily in the summative evaluations of teaching attendings. Additionally, residents are often filling out

evaluations from a prior rotation some weeks later while they are busy on a new service. Whether it is due to recall bias, apathy, fear of retribution, or poorly constructed instruments, the data that is being collected appears to have limited fidelity. Additionally, it is understood that learners' evaluations of clinician-educators are largely formative. For such assessments to be impactful, the feedback must be timely and linked to specific details with examples of how the attending achieved or faltered in a particular competency.

Pimping exposes the brazen hierarchy within medical training. Yet, some affirm that pinging-type questioning may be an effective means of encouraging self-directed learning.^{5-6, 17} As suggested by studies in undergraduate medical education,^{5,14} learners may also augment their emotional reaction to pinging based on the perceived intent of the questioning; if there is clear scholastic intent, it may be more tolerable – especially among learners who are particularly gritty.¹⁷⁻²⁰ More recent literature calls into question the label and practice of pinging, with concern for microaggressions and even bullying.^{3,7,21} Clearly, the expectation should be for attendings to ask questions in a “non-threatening” manner and to explore errors “without intimidation”; these are essential for professional working relationships. Attendings should only inspire excellence in trainees along a horizontal team dynamic, one that encourages all team members to consider their own deficiencies or “learning edges” while emphasizing a shared growth mindset that values learning together.²²

Several limitations of this study should be considered. First, the limitations applied to the original data set describing attending attitudes towards pinging and self-assessment of pinging behavior apply to the current study; these include the effect of social desirability on accurate self-reporting and the use of a pinging score as a surrogate for teaching behavior rather than direct observation of pinging behavior. Second, resident evaluations of attending physicians may be influenced by the recall bias

introduced by delays in completing evaluations. Being further removed from a rotation with an attending with a stronger pinging phenotype may lead to flawed remembering of important interactions. Further, very few residents complete the evaluations thoroughly, with detailed comments, which might actually provide perspective into how attendings' pinging behaviors were interpreted. Third, by analyzing the mean evaluations for each attending physician, our data may be exploring the association between ‘team consensus’ summative evaluations and pinging. While this may be a strength, it can also be a limitation if one learner was ‘pimped’ most harshly and frequently. Finally, with regards to external generalizability, one cannot assume that the perspectives in this cohort would match those of faculty affiliated with other departments or schools. We understand that program-level attitudes, cultures, and norms may have a greater impact on the learning environment than do individual, faculty-level, behaviors.

In conclusion, pinging behaviors were not correlated with summative evaluations of attendings from residents at two internal medicine residency programs. This finding offers opportunities to examine our current evaluations of teaching physicians, and how these attendings can support the development of graduate learners in ways that promote psychological safety and well-being. Teaching evaluations should provide meaningful feedback so that an attending physician can understand whether learners thought that he or she successfully created a safe and compassionate learning environment that facilitated the trainees' professional development. Attendings may better support trainees' development by setting the expectation for a shared growth mindset, and acknowledging that learning requires vulnerability, but not humiliation. Most questions should be formative, rather than evaluative, with the intent to promote improvement, expansion, and reflection on knowledge and skills.

ACKNOWLEDGMENTS AND DISCLOSURES

Dr. Wright is the Anne Gaines and G. Thomas Miller Professor of Medicine and is supported through the Johns Hopkins Center for Innovative Medicine.

FUNDING/OTHER SUPPORT

Dr. McEvoy is supported by the following grants, NIH/NIDDK- R01DK089174, NIH/NIDDK- R01DK108784, and AHA-17MCPRP33400031.

DISCLOSURES

None reported

ETHICAL APPROVAL

This study was approved by the Johns Hopkins School of Medicine Institutional Review Board on April 20th 2016 (reference number 00098488).

PREVIOUS PRESENTATIONS

Oral abstract presentation in Baltimore, Maryland on July 2017 at the JHU Masters of Education in the Health Professions (MEHP) Summer Conference: Advancing Careers through Interprofessional Health Professions Education- Teach, Research, Lead.

REFERENCES

1. Brancati FL. The art of pimping. *JAMA*. 1989;262(1):8990. <http://dx.doi.org/10.1001/jama.1989.03430010101039>.
2. Detsky AS. The art of pimping. *JAMA*. 2009;301(13):1379-1381. Accessed Sep 12, 2019. doi: 10.1001/jama.2009.247.
3. Kost A, Chen FM. Socrates was not a pimp: Changing the paradigm of questioning in medical education. *Acad Med*. 2015;90(1):20-24. doi: 10.1097/ACM.0000000000000446.
4. Pelling M. Managing uncertainty and privatising apprenticeship: Status and relationships in english medicine, 1500–1900. *Soc Hist Med*. 2019;32(1):34-56. <https://academic-oup-com.proxy1.library.jhu.edu/shm/article/32/1/34/4094884>. doi: 10.1093/shm/hkx059.
5. Wear D, Kokinova M, Keck-McNulty C, Aultman J. Pimping: Perspectives of 4th year medical students. *Teach Learn Med*. 2005;17(2):184-191. doi: 10.1207/s15328015tlm1702_14.
6. Healy JM, Yoo PS. In defense of "pimping". *J Surg Educ*. 2015;72(1):176-177. Accessed Sep 12, 2019. doi: 10.1016/j.jsurg.2014.06.012.
7. Reifler DR. The pedagogy of pimping: Educational rigor or mistreatment? *JAMA*. 2015; 314(22):2355-2356. <http://dx.doi.org/10.1001/jama.2015.14670>. doi: 10.1001/jama.2015.14670.
8. Scott KM, Caldwell PH, Barnes EH, Barrett J. "Teaching by humiliation" and mistreatment of medical students in clinical rotations: a pilot study. *The Medical journal of Australia*. Aug 17 2015;203(4):185e 181-186. doi: 10.5694/mja15.00189.
9. Hautz WE, Schröder T, Dannenberg KA, et al. Shame in medical education: A randomized study of the acquisition of intimate examination skills and its effect on subsequent performance. *Teach Learn Med*. 2017;29(2):196-206. doi: 10.1080/10401334.2016.1254636.
10. McEvoy JW, Shatzer JH, Desai SV, Wright SM. 2019. Questioning Style and Pimping in Clinical Education: A Quantitative Score Derived from a Survey of Internal Medicine Teaching Faculty. *Teaching and Learning in Medicine*. 31(1):53-64. doi:10.1080/10401334.2018.1481752.
11. MedHub Web site. <http://www.medhub.com/>.
12. van Schaik KD. A piece of my mind. Pimping socrates. *JAMA*. 2014;311(14):1401-1402. <https://www.ncbi.nlm.nih.gov/pubmed/24715073>. doi: 10.1001/jama.2014.2155.
13. McCarthy CP, McEvoy JW. Pimping in medical education: Lacking evidence and under threat. *JAMA*. 2015;314(22):2347-2348. <http://dx.doi.org/10.1001/jama.2015.13570>.
14. Lo L, Regehr G. Medical students' understanding of directed questioning by their clinical preceptors. *Teach Learn Med*. 2017;29(1):5-12. doi: 10.1080/10401334.2016.1213169.
15. Williams, E. A., Miesner, A. R., Beckett, E. A., & Grady, S. E. (2018). "Pimping" in Pharmacy Education: A Survey and Comparison of Student and Faculty Views. *Journal of Pharmacy Practice*, 31(3), 353–360. <https://doi.org/10.1177/0897190017715393>
16. Fluit, Cornelia R. M. G., Bolhuis S, Grol R, Laan R, Wensing M. Assessing the quality of clinical teachers: A systematic review of content and quality of questionnaires for assessing clinical teachers. *J Gen Intern Med*. 2010;25(12):1337-1345. doi: 10.1007/s11606-010-1458-y.
17. Haizlip J, May N, Schorling J, Williams A, Plews-Ogan M. The negativity bias, medical education, and the culture of academic medicine: why culture change is hard. *Acad*

Med. 2012;87(9):1205–1209. doi:
10.1097/ACM.0b013e3182628f03.
18. Bynum, WE and Haque, TM. 2016. Risky
Business: Psychological Safety and the Risks of
Learning Medicine. *Journal of Graduate
Medical Education*. 8(5): 780-782. doi:
10.4300/JGME-D-16-00549.1.
19. Stoddard HA, O’Dell DV. Would Socrates
Have Actually Used the “Socratic Method” for
Clinical Teaching? *J Gen Intern Med*. 2016; 31
(9):1092-6. doi: 10.1007/s11606-016-3722-2.
20. Sargeant J, Mann K, Sinclair D, Van der
Vleuten C, Metsemakers J. 2008a.

Understanding the influence of emotions and
reflection upon multi-source feedback
acceptance and use. *Adv Health Sci Educ
Theory Pract*. 13:275–288. doi:
10.1007/s10459-006-9039-x.
21. Ayyala MS, Rios R, Wright SM. Perceived
bullying among internal medicine
residents. *JAMA*. 2019;322(6):576-
578. [https://search.proquest.com/docview/2279
797299](https://search.proquest.com/docview/2279797299). doi: 10.1001/jama.2019.8616.
22. Dweck C. *Mindset: The New Psychology of
Success*. New York: Random House; 2016.