

The Effect of Gender on 360-Degree Evaluations at One Academic Center

Franklyn Cladis, MD – The Childrens Hospital of Pittsburgh of UPMC

Erica Sivak, MD – The Childrens Hospital of Pittsburgh of UPMC

Introduction: Gender bias is a potential issue in the education environment. It may affect how medical students, residents, and fellows are educated, evaluated, and provided feedback. Some studies have demonstrated that female trainees receive less independence, are graded differently from male trainees and receive more critical feedback. There is currently limited data investigating the influence of the gender of the *evaluator* on male and female trainees. Do male evaluators grade female trainees more harshly than male trainees or do female evaluators grade female trainees more harshly? We hypothesize that the evaluations of male and female trainees would be different from male and female evaluators. Specifically we hypothesized female evaluators would be more critical than male evaluators when grading female trainees.

Methods: At the Children's Hospital of Pittsburgh of UPMC, gender of the evaluator has been tracked as part of the 360-degree evaluation tool for the pediatric anesthesiology fellowship. Fellows receive evaluations from PACU nurses, preoperative nurse practitioners, anesthesia technicians, and administrative staff as part of their 360-degree evaluation process. They are evaluated on a Likert scale (1=unsatisfactory-9, =superior) on qualities related professionalism and communication, not patient care or medical knowledge. Tracking gender of the evaluator allows comparison of evaluator gender to fellow gender. An overall grade was collected for each fellow from each evaluator.

Results: Deidentified data from 2014 to 2017 was retrospectively reviewed for 360-degree evaluations. From 2014 to 2017 a total of 577 evaluations were performed on 36 fellows (20 male, 16 female). 354 evaluations were performed for male fellows and 223 were performed for female fellows. A global score was collected for each evaluation. The average global score for females was higher than for male (8.52 vs. 8.41, respectively) but this was not statistically significant ($p=0.129$). The global score for male and female fellows were separately investigated from male and female evaluators. The results are outlined in Table 1.

Discussion: The global score from a 360-degree evaluation tool demonstrated higher scores for female fellows than for male fellows in one academic center and this difference is even higher from *male* evaluators. This is different from the hypothesis that gender bias may result in lower scores for female fellows and suggests that gender bias may not be as prevalent in this specific population with this specific tool. It is possible that the evaluation tool being used affected the results and was biased in favor of female fellows.

Table 1: Effect of Evaluator Gender on Fellow 360-degree Global Score

Fellow Gender (total number)	Female Evaluator Grade	P-value	Male Evaluator Grade	P-value
Female (223)	8.47	0.29	8.61	0.03
Male (354)	8.42		8.42	