

A Novel Method for Blinding Reviewers to Gender of Proceduralists for the Purposes of Gender Bias Research

Pallaci, See, Weinstock, Riddell, et al

Research has shown gender disparity on ACGME milestone evaluations, with the largest differences in procedural competency. There are currently no established methods by which researchers can blind reviewers to gender to evaluate for bias in procedural evaluations.

Our objective was to determine if a novel recording method could effectively blind evaluators to the gender of trainees performing simulated procedures. We hypothesized that reviewers would not be able to correctly identify gender in >50% of assessments.

After removing all jewelry from their hands, proceduralists were gowned, double-gloved, and filmed by a professional videographer while performing simulated procedures. Only their double-gloved hands, gowned forearms and lower torsos were visible in the videos. Five residents (two male and three female) performed three procedures each (lumbar puncture, chest tube and central venous catheter placement), yielding 15 videos. Seven experts watched short video clips (30-45 seconds) and evaluated the perceived gender of the proceduralist on a Likert scale (1=definitely male, 3=likely male, 5=can't tell, 7=likely female, 9=definitely female). A response concordant with proceduralist gender with a confidence level of likely or higher (1-3 for males, 7-9 for females) was considered correct gender identification. Responses discordant with proceduralist gender or in the unsure range (score 4-6), were considered incorrect. Reviewer scores were described, and one-sample equality of proportions were used to assess significances against the null of >50% correct gender identification.

Of 105 total responses, 56 (53.3%) expressed confidence in the gender of the proceduralist (1-3 or 7-9), and only 64.3% of those assessments (36/56) were accurate. Across all reviewers and procedures, the proceduralist's gender was correctly identified in 30.8% (95% CI: 22.7-39.9%) of videos. This proportion was not statistically significant compared to the null of >50% correct gender identification ($p=1.00$). There were also no significant differences when the data were stratified by each procedure assessed.

The recording method used was effective in blinding reviewers to the gender of the proceduralist and represents an innovative approach to facilitate research into gender bias in procedural evaluations.