

# A Values Affirmation Intervention to Improve Female Residents' Surgical Performance

Arghavan Salles, MD, PhD  
Claudia M. Mueller, PhD, MD  
Geoffrey L. Cohen, PhD

## ABSTRACT

**Background** Female residents in surgical training may face stereotype threat. The awareness of negative stereotypes about surgical ability based on gender may heighten stress and thus reduce performance.

**Objective** The main objective of this study was to assess the effectiveness of a brief stress-reducing writing exercise, known as a values affirmation, to mitigate the negative effects of stereotype threat on the performance of female surgical residents.

**Methods** This is a randomized, controlled trial in which 167 residents were invited to participate. A total of 45 resident volunteers, including 18 women, were randomized to the affirmation condition or the no-affirmation condition. We administered a values affirmation intervention and measured clinical evaluations data both prior to and 6 months after the intervention.

**Results** Women benefited from the affirmation. Women who had participated in the affirmation exercise earned higher clinical evaluation scores than those in the control condition ( $B = 0.34, P < .05$ ). For men, performance did not differ by affirmation condition ( $B = -0.20, P = .35$ ).

**Conclusions** Our findings suggest a benefit of values affirmation for women in surgical training, as measured by performance on clinical evaluations. This suggests that a brief psychological intervention may improve on-the-job performance for women in surgery, an underrepresented group.

## Introduction

Currently, approximately 50% of medical students in the United States are women.<sup>1</sup> Yet historically, medicine, and particularly surgery, has been dominated by men. There remains a specter of the stereotype that men make better physicians than women. Women in medicine who believe this notion may be undermined by a psychological phenomenon known as stereotype threat. This phenomenon describes the harmful effect that knowledge of a negative stereotype can have on one's level of stress and performance.<sup>2-6</sup> Recently, Burgess et al<sup>7</sup> argued that there are a number of factors, including the paucity of female role models and women's relative minority status, that support the argument that stereotype threat affects women in academic medicine. In surgical fields, only 32% of trainees nationwide are women,<sup>8</sup> and female surgeons and residents may experience stereotype threat more acutely than women in nonsurgical specialties. Indeed, data regarding stereotype perception (ie, the degree to which residents believe others to hold a negative stereotype about women's ability) corroborate this idea.<sup>9</sup> When surgical residents were asked

whether they think other residents, faculty, or the general public endorse the stereotype that men are better surgeons than women, the answer was a resounding yes. In contrast, when residents in nonsurgical specialties were asked whether they think other residents, faculty, or the general public endorse the stereotype that men are better physicians than women, the answer was a weak yes.<sup>9</sup>

Our study evaluated the effectiveness of a values affirmation intervention for reducing stereotype threat and improving performance in surgical residency. Values affirmations, in which participants reflect on their core values, have been shown to be effective at counteracting stereotype threat.<sup>10-12</sup> These affirmations decrease identity threats, such as stereotype threat, by reminding people of positive sources of meaning and support in their lives.<sup>13</sup> If timely, affirmations may stop or slow a downward spiral, in which poor performance leads to higher stress, leading to worse performance, in a recursive cycle.<sup>11</sup> Randomized field experiments have shown positive effects of values affirmation on student grades lasting up to 2 years after the intervention.<sup>10-12</sup>

The effects of stereotype threat, and the interventions to reduce them, have previously been examined mainly in schools.<sup>14</sup> Here, we performed an exploratory experiment applying the social-psychological intervention of values affirmation in a surgical residency, a work setting in which stereotype threat

DOI: <http://dx.doi.org/10.4300/JGME-D-15-00214.1>

*Editor's Note: The online version of this study contains survey questions.*

is likely to be prevalent. We predict that women will perform worse than men, due to stereotype threat, and that the values affirmation intervention will improve their performance.

## Methods

### Study Design

We invited all trainees in 9 surgical specialties at 1 academic medical center to participate in a randomized, controlled experiment. Resident volunteers were recruited at resident education meetings. Recognizing that we might not have adequate power to detect treatment effects, we enrolled as many residents as we could given the limited availability of our group of interest (female surgical residents).

First, we asked all residents ( $N = 167$ ) to complete a stereotype perception self-report measure, embedded in a larger survey, to establish their perceived level of stereotype threat. Then, we asked them specifically about the extent to which they believed those around them think that men are better surgeons than women (survey questions provided as online supplemental material).<sup>9</sup> Consistent with prior studies,<sup>15</sup> we used this stereotype perception measure as an index of stereotype threat. A total of 146 respondents (53 women, 87% response rate) completed the survey.

All 167 residents were randomized to either the treatment or control condition (even if they had not completed the survey). The experimenters were blind to the condition. Following previously described procedures,<sup>10</sup> volunteers completed the intervention in their education meetings. The first page presented a menu of 12 important values, such as friends and family, sports, music, and religion. We asked those in the treatment (affirmation) condition to select the 2 or 3 values most important to them and then, on a separate page, to discuss the rationale for their choices. We asked those in the control (no affirmation) condition to select the 2 or 3 values that were least important to them, and to write about why someone else might find these values important.

The study was approved by the Institutional Review Board at Stanford University. All residents present on the day of the intervention consented to the study ( $n = 93$ ; 32 women; 56% of all 167 surgical residents). The full protocol and materials are available from the authors.

### Data Collection and Analysis

We collected faculty evaluations of residents' clinical performance 6 months before and 6 months after the intervention. Faculty were blinded to condition, and

#### What was known and gap

Female surgical residents may face stereotype threat, and the awareness of perceptions of lower ability based on gender may increase stress and result in lower performance.

#### What is new

An easily implementable values affirmation exercise may reduce the negative impact of stereotype threat.

#### Limitations

Single institution study may limit generalizability; small sample with potentially inadequate power to detect meaningful differences.

#### Bottom line

A values affirmation intervention may mitigate the effects of stereotype threat and improve female surgical residents' performance.

virtually all faculty were unaware that the study had taken place. The clinical evaluation data consisted of scores for each of the core competencies as defined by the Accreditation Council for Graduate Medical Education: patient care, medical knowledge, practice-based learning and improvement, systems-based practice, professionalism, and interpersonal communication. We standardized the data because the evaluation forms used by the residency programs were not the same. Both before and after the intervention, within each specialty, each participant received a single score for each competency. These scores were the average of all items pertaining to each competency. For example, if there were 3 items on medical knowledge, the 3 scores were averaged together to create 1 medical knowledge score. The average scores from the 6 competencies were then averaged to create an overall clinical performance score. This was then Z-scored within each specialty, such that each specialty's scores had a mean of 0 and a standard deviation of 1, to allow for cross-specialty comparisons. Comments from the evaluations were not included in the analysis.

Residents were included in final analyses if they had (1) preintervention performance data; (2) postintervention performance data; and (3) preintervention survey data. Performance data were missing either due to this being the resident's first year (such that there was no prior performance evaluation) or because of the residency program not wishing to share evaluation data (2 of 9 programs). Data were missing from residents who did not complete the survey; 45 residents (18 women, 27% of all 167 possible participants) constituted the final sample. The postgraduate year and specialties of the residents in each arm are shown in TABLE 1. FIGURE 1 shows their enrollment at each stage of the experiment.

Linear regression analyses were performed to assess gender differences in the clinical evaluations and

**TABLE 1**  
Number of Residents Comprising Final Sample

| Specialty and PGY                    | Male <sup>a</sup> | Female <sup>a</sup> | Male <sup>b</sup> | Female <sup>b</sup> |
|--------------------------------------|-------------------|---------------------|-------------------|---------------------|
| Cardiothoracic                       |                   |                     |                   |                     |
| PGY-1                                | 0                 | 0                   | 1                 | 0                   |
| General surgery                      |                   |                     |                   |                     |
| PGY-1                                | 2                 | 2                   | 3                 | 0                   |
| PGY-2                                | 2                 | 2                   | 0                 | 0                   |
| PGY-3                                | 0                 | 0                   | 1                 | 1                   |
| PGY-4                                | 1                 | 1                   | 0                 | 0                   |
| PGY-5                                | 1                 | 1                   | 0                 | 1                   |
| Ophthalmology                        |                   |                     |                   |                     |
| PGY-2                                | 0                 | 1                   | 1                 | 0                   |
| PGY-3                                | 1                 | 1                   | 0                 | 0                   |
| Orthopedic surgery                   |                   |                     |                   |                     |
| PGY-1                                | 0                 | 1                   | 0                 | 0                   |
| Otolaryngology–head and neck surgery |                   |                     |                   |                     |
| PGY-2                                | 1                 | 0                   | 1                 | 0                   |
| PGY-4                                | 0                 | 1                   | 0                 | 0                   |
| PGY-5                                | 1                 | 0                   | 1                 | 1                   |
| Research                             | 0                 | 1                   | 0                 | 1                   |
| Plastic surgery                      |                   |                     |                   |                     |
| PGY-4                                | 0                 | 0                   | 1                 | 0                   |
| PGY-5                                | 1                 | 0                   | 0                 | 0                   |
| PGY-6                                | 1                 | 0                   | 0                 | 0                   |
| Urology                              |                   |                     |                   |                     |
| PGY-1                                | 1                 | 0                   | 1                 | 1                   |
| PGY-2                                | 1                 | 0                   | 0                 | 0                   |
| PGY-3                                | 0                 | 0                   | 1                 | 0                   |
| PGY-4                                | 0                 | 0                   | 1                 | 1                   |
| PGY-5                                | 1                 | 0                   | 0                 | 0                   |
| Vascular surgery                     |                   |                     |                   |                     |
| PGY-1                                | 0                 | 1                   | 0                 | 0                   |

Abbreviation: PGY, postgraduate year.

<sup>a</sup> Control condition.<sup>b</sup> Treatment condition.

whether the affirmation intervention improved the performance of the negatively stereotyped group, women. Analyses were conducted in Stata SE version 10.1 (StataCorp LP, College Station, TX).

## Results

The baseline characteristics of the women who met criteria for inclusion in the analysis are presented in TABLE 2. There were no statistically significant differences in United States Medical Licensing Examination scores or stereotype perception between the treatment and control conditions.

**TABLE 2**  
Women's Average Baseline Data by Affirmation Condition<sup>a</sup>

|                       | Affirmation Condition, Mean (SD) | Control Condition, Mean (SD) |
|-----------------------|----------------------------------|------------------------------|
| Age                   | 30.63 (2.67)                     | 29.54 (2.63)                 |
| Stereotype perception | 5.52 (0.73)                      | 5.12 (0.89) <sup>b</sup>     |
| USMLE                 | 240.11 (16.50)                   | 236.46 (14.30)               |
| Clinical evaluations  | −0.45 (1.11)                     | 0.25 (1.05) <sup>c</sup>     |

Abbreviation: USMLE, United States Medical Licensing Examination.

<sup>a</sup> A total of 18 to 23 women included per row. Number per row varies slightly based on availability of data.<sup>b</sup> Rated on a 7-point scale, with higher numbers indicating more stereotype perception.<sup>c</sup> Scores were standardized within specialty.

Our initial prediction was that women's clinical performance under stereotype threat, as measured by evaluations, would be worse than that of their male counterparts. In contrast, the standardized clinical evaluation scores for men and women before and after the intervention showed no difference in clinical evaluation performance by gender at either time point (both  $P > .18$ ; TABLE 3).

When we examined only women, the group we believed to be most affected by negative stereotyping, we also found no relationship between evaluation scores and stereotype perception ( $P = .07$ ).

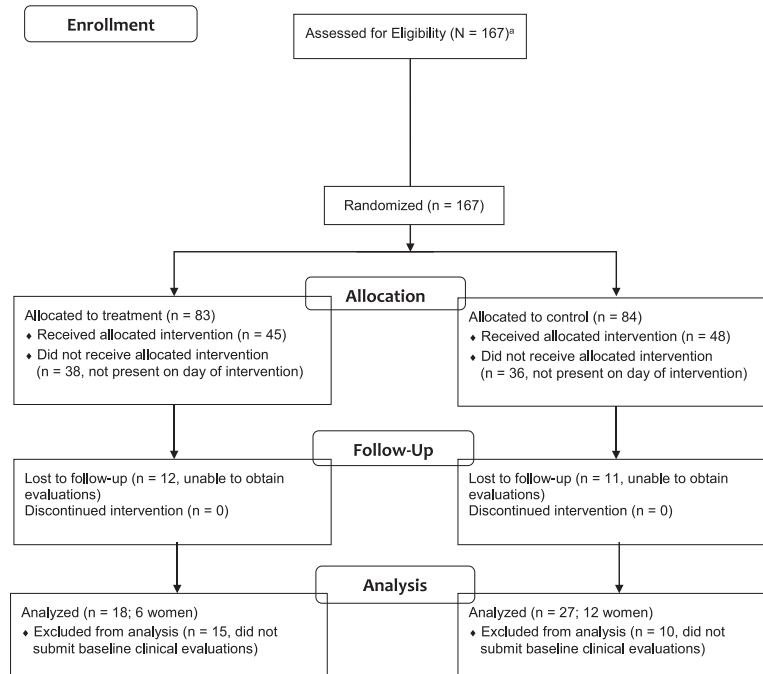
Our second hypothesis was that affirmations would improve women's performance as measured by evaluation scores. FIGURE 2 shows that, controlling for baseline evaluation scores, stereotype threat, and specialty, affirmation had a significant positive effect on women's clinical performance (main effect  $B = 0.34$ ;  $P < .05$ ; effect size = 0.17). The interaction of stereotype perception and affirmation condition did not reach significance ( $B = 0.22$ ,  $P = .42$ ). Thus, regardless of the degree of stereotype perception, women performed better, on average, if they had been affirmed.

Consistent with prior research, men, who do not experience consistent stereotype threat in surgery, were unaffected by affirmation. All main and interactive effects were insignificant (all  $P > .35$ ),

**TABLE 3**  
Preintervention and Postintervention Standardized Clinical Evaluation Scores for Men and Women<sup>a</sup>

|       | Preintervention, Mean (SD) | Postintervention, Mean (SD) |
|-------|----------------------------|-----------------------------|
| Men   | 0.13 (0.85)                | −0.02 (0.86)                |
| Women | −0.02 (1.00)               | −0.08 (1.03)                |

<sup>a</sup> Data available for 65 participants (24 women) preintervention and 68 participants (25 women) postintervention.

**FIGURE 1****CONSORT 2010 Flow Diagram**

<sup>a</sup> The total number of residents eligible here includes all surgical residents at this institution. However, only approximately 55% of those residents were present at the educational meetings where the intervention took place. All those who were present participated in the intervention. Of this group, the only residents excluded from the final sample were those who did not have evaluations. In these cases, either the data did not exist (the resident had not been at the institution at the time of interest) or the residency program chose not to share residents' evaluation data despite residents' consent (2 of 9 programs).

and specifically the main effect of affirmation condition was not significant ( $B = -0.20$ ,  $P = .35$ ).

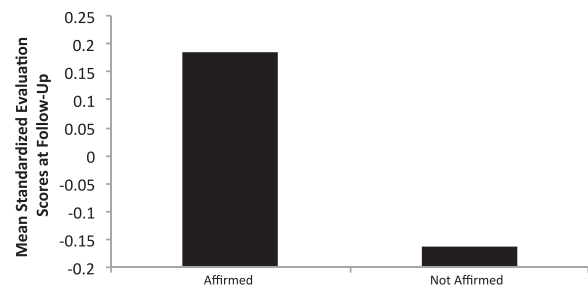
## Discussion

As predicted, women, the group at risk of experiencing stereotype threat in surgical training, performed better after a values affirmation intervention. The timing of the evaluations (6 months after the intervention) suggests that the affirmations have a long-lasting effect. This is consistent with prior studies,<sup>10–12</sup> which have shown improved outcomes up to 2 years later. Although men also experience stress during training, they do not face an identity threat, and we expected affirmations to have no effect on their performance. Minority residents, who likely also face stereotype threat, might benefit from values affirmations.

Given the extensive literature establishing the relationship between stereotype threat and decreased performance, we were surprised to find no gender difference in clinical performance. There are several explanations for this. First, perhaps women in this study experienced reactance, causing them to perform better, rather than worse, in the face of explicit stereotype threat.<sup>16</sup> Second, it is possible that our small sample size made it impossible to detect a

difference in performance. Finally, men's and women's equal performance may be consistent with stereotype threat. Prior research has shown that when members of a stereotyped group perform under less stereotype threat, they can actually outperform the nonstereotyped group.<sup>17</sup>

Although performing a values affirmation is simple, in an experimental setting collecting performance measures, survey measures, and the responses to the affirmations is a significant undertaking. Outside of the context of a research program, however, values affirmations are not onerous. They require very little

**FIGURE 2****Women's Evaluations by Affirmation Condition**

Note: These values are derived from a linear regression predicting scores for women based on affirmation condition (treatment or control) and stereotype threat controlling for baseline evaluation scores and specialty.

time (15 minutes) and are well tolerated. The main challenge to this type of work is that if participants know why they are doing the affirmations, the desired impact is not found.<sup>18</sup> We framed the affirmations as an attempt on the part of the administration to understand what residents valued. The intervention should also be done early enough in the academic year to prevent a downward spiral. We performed the intervention in early fall. By lifting the stress associated with stereotype threat, the affirmations may unlock the aptitude and positive resources for growth latent in the person and in his or her environment.<sup>19</sup>

There are 2 limitations in this research. First, our study had a small sample and was conducted at 1 academic center, so there may be meaningful effects we did not capture. For example, although not statistically significant, the mean performance of women high in stereotype perception tended to be lower than that of women low in stereotype perception. Similarly, although there was no statistically significant interaction of stereotype perception with treatment condition, the direction of the data was consistent with a stronger treatment effect for those facing higher levels of stereotype perception. With a larger sample size, we might have found data supporting these hypotheses. It also is not clear whether improved evaluation scores would translate to improved education or patient care outcomes. These questions about the potential impact of this type of intervention on measures such as resident learning, well-being, and patient care are only answerable with replication and a larger number of participants.

We have since expanded our research to include another social-psychological intervention, centered on belonging, and its effects on residents' performance and well-being.

## Conclusion

A values affirmation intervention can mitigate the effects of stereotype threat among female surgical residents. Our data suggest that this low-cost, low-input intervention may improve female residents' performance.

## References

1. Association of American Medical Colleges. The state of women in academic medicine, the pipeline and pathways to leadership. 2015. <https://members.aamc.org/eweb/upload/The%20State%20of%20Women%20in%20Academic%20Medicine%202013-2014%20FINAL.pdf>. Accessed April 1, 2016.
2. Blascovich, J, Spencer SJ, Quinn D, Steele C. African Americans and high blood pressure: the role of stereotype threat. *Psychol Sci*. 2001;12(30):225–229.
3. Steele CM. A threat in the air: how stereotypes shape intellectual identity and performance. *Am Psychol*. 1997;52(6):613–629.
4. Nguyen HD, Ryan AM. Does stereotype threat affect test performance of minorities and women? A meta-analysis of experimental evidence. *J Appl Psychol*. 2008;93(6):1314–1334.
5. Spencer SJ, Steele CM, Quinn DM. Stereotype threat and women's math performance. *J Exp Soc Psychol*. 1999;35:4–28.
6. Shih M, Pittinsky TL, Ambady N. Stereotype susceptibility: identity salience and shifts in quantitative performance. *Psychol Sci*. 1999;10(1):80–83.
7. Burgess DJ, Joseph A, van Ryn M, Carnes M. Does stereotype threat affect women in academic medicine? *Acad Med*. 2012;87(4):506–512.
8. Viola KV, Bucholz E, Yeo H, Piper C, Bell RH Jr, Sosa JA. Impact of family and gender on career goals: results of a national survey of 4586 surgery residents. *Arch Surg*. 2010;145(5):418–424.
9. Salles A, Mueller CM, Cohen GL. Exploring the relationship between stereotype perception and residents' well-being. *J Am Coll Surg*. 2015;222(1):52–58.
10. Cohen GL, Garcia J, Apfel N, Master A. Reducing the racial achievement gap: a social-psychological intervention. *Science*. 2006;313(5791):1307–1310.
11. Cohen GL, Garcia J, Purdie-Vaughns V, Apfel N, Brzustoski P. Recursive processes in self-affirmation: intervening to close the minority achievement gap. *Science*. 2009;324(5925):400–403.
12. Sherman DK, Cohen GL. Accepting threatening information: self-affirmation and the reduction of defensive biases. *Curr Dir Psychol Sci*. 2002;11(4):119–123.
13. Creswell JD, Welch WT, Taylor WE, Sherman DK, Gruenewald TL, Mann T. Affirmation of personal values buffers neuroendocrine and psychological stress responses. *Psychol Sci*. 2005;16(11):846–851.
14. Woolf K, McManus IC, Gill D, Dacre J. The effect of a brief social intervention on the examination results of UK medical students: a cluster randomized controlled trial. *BMC Med Educ*. 2009;9:35.
15. Marx DM, Stapel DA. Distinguishing stereotype threat from priming effects: on the role of the social self and threat-based concerns. *J Pers Soc Psychol*. 2006;91(2):243–254.
16. Kray LJ, Thompson L, Galinsky A. Battle of the sexes: gender stereotype confirmation and reactance in negotiations. *J Pers Soc Psychol*. 2001;80(6):942–958.
17. Walton GM, Spencer SJ. Latent ability: grades and test scores systematically underestimate the intellectual

ability of negatively stereotyped students. *Psychol Sci*. 2009;20(9):1132–1139.

18. Sherman DK, Cohen GL, Nelson LD, Nussbaum AD, Bunyan DP, Garcia J. Affirmed yet unaware: exploring the role of awareness in the process of self-affirmation. *J Pers Soc Psych*. 2009;97(5):745–764.
19. Yeager DS, Walton GM. Social-psychological interventions in education: they're not magic. *Rev Educ Res*. 2011;8(2):267–301.



**Arghavan Salles, PhD, MD**, is a Fellow, Department of Surgery, Washington University, St Louis; **Claudia M. Mueller, PhD, MD**, is Assistant Professor, Stanford Health Care; and **Geoffrey L. Cohen, PhD**, is Professor, Stanford Graduate School of Education.

**Funding:** This research was funded by the Stanford University School of Education Dissertation Support Grant, the Vice Provost for Graduate Education Diversity Dissertation Research Opportunity, and the National Center for Research Resources and the National Center for Advancing Translational Sciences, National Institutes of Health, through grant 5 KL2 RR025743.

**Conflict of interest:** The authors declare they have no competing interests.

The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

**Corresponding author:** Arghavan Salles, MD, PhD, Washington University School of Medicine in St Louis, Campus Box 8109, 660 South Euclid Avenue, St Louis, MO 63110, 650.387.5275, arghavan@alumni.stanford.edu

Received May 25, 2015; revisions received December 18, 2015, and February 4, 2016; accepted February 22, 2016.