Motivations to Control Prejudice Bias Performance Feedback in Developmental Relationships

C. Malik Boykin
Brown University, c_boykin@brown.edu

Christine R. Smith
Baruch College - City University of New York, csmith@gradcenter.cuny.edu

Follow this and additional works at: https://scholarworks.bgsu.edu/pad

Part of the Human Resources Management Commons, Industrial and Organizational Psychology Commons, and the Other Psychology Commons

MOTIVATIONS TO CONTROL PREJUDICE BIAS PERFORMANCE FEEDBACK IN DEVELOPMENTAL RELATIONSHIPS

C. Malik Boykin[1] and Christine R. Smith[2]

1. Department of Cognitive, Linguistic, & Psychological Sciences, Brown University
2. Department of Psychology, Baruch College, CUNY

ABSTRACT

KEYWORDS

performance feedback, Rater bias, motivation, prejudice, mentoring

In developmental relationships, providing accurate assessments of performance is necessary to maximize the developmental benefits for those receiving the feedback. Research suggests that performance assessments for underrepresented minorities are susceptible to biases related to out-group prejudice; however, little is known about the contributions of motivations to control prejudice, particularly in face-to-face settings. Addressing this, we examined the influences of internal and external motivations to control prejudice (IMS and EMS) on the positivity of White mentor's feedback about their underrepresented minority mentee's task performance. We analyzed video-recorded interactions between 56 randomly assigned cross-racial dyads, wherein mentees performed a speech task and were given subsequent face-to-face verbal feedback from their mentor. To gain comparatively unbiased assessments of feedback positivity and of mentee performance, we used independent coders. Using structural equation modeling, our results suggested that positivity of mentors' feedback was uniquely predicted by both IMS and EMS over and above mentee performance.

Feedback on performance at work is essential for an employee's development and career advancement. *Developmental feedback* refers to valuable future-oriented information that enables individuals to learn, develop, and make improvements in their work (Zhou, 2003). In the workplace, developmental feedback can come from a variety of sources, including direct managers, coworkers, subordinates, peers outside of the organization, and mentors (e.g., Allen, Shockley, & Poteat, 2010; DeRue & Wellman, 2009; Kudisch, Fortunato, & Smith, 2006; Ng, Koh, Ang, Kennedy, & Chan, 2011; Zhou, 2003). Developmental feedback differs from administrative feedback, the latter of which is used to evaluate decisions such as salary and promotion on past performance (Aguinis, 2009; Cleveland & Murphy, 1989; Murphy & Cleveland, 1995).

As helpful as developmental feedback can be, it is also susceptible to personal motivations and biases, which can render the feedback less accurate and therefore less useful (e.g, Murphy & Cleveland, 1995; Murphy, Cleveland, Skattebo, & Kinney, 2004; Wang, Wong, & Kwong, 2010; Wong & Kwong, 2007). Although bias in performance appraisals and raters' general motivations have long been studied, less is known about how raters' motivations to appear unbiased against certain individuals can affect the feedback they give their employees—especially in face-to-face interactions. In the current study, we examine how a person's motivation to respond in a nonprejudiced manner toward a ratee from an underrepresented minority group, in fact, may produce more biased developmental feedback within a spontaneous face-to-face context.

Rater Motivations and Biased Feedback

Like any human-managed process, performance feedback is prone to contamination by human biases. Over time, the cumulative effects of subtle bias in the performance management processes can impact the professional advancement opportunities for members of underrepresent-

Corresponding author:
C. Malik Boykin
Department of Cognitive, Linguistic, & Psychological Sciences, Brown University, 190 Thayer St., Providence, RI 02912
Email: c_boykin@brown.edu

ed groups (Agars, 2004; Agars & Cazares, 2017). Given that these processes are intended to help develop, reward, and promote employees, Agars and Cazares (2017) argue that cumulative discrimination manifested from bias in performance ratings and feedback can undermine the validity of these processes and the evaluations they produce, ultimately resulting in barriers to advancement for members of stigmatized groups. Biased feedback can mean an employee will lack the appropriate information to adapt their behavior in order to be considered for future advancement opportunities (Greenhaus, 1987).

In their seminal work, Murphy and Cleveland (1995) suggested that there are a number of motivations raters can have to provide more lenient appraisals when evaluating someone's performance. That is, inaccuracies in appraisals of performance are not due to an inability to accurately evaluate but the result of an unwillingness to do so. Specifically, Murphy and Cleveland (1995) argue that raters are motivated to inaccurately appraise performance because inflated favorable appraisals can produce positive outcomes for the ratee, produce positive outcomes for the rater (e.g., positive recognition for helping the ratee improve), avoid negative reactions from the ratee towards the rater, and help maintain the organization's image. Levy and Williams (2004) extended this conceptualization by stating that a rater's appraisal of a ratee's performance also can be affected by rating motivations arising from (a) attributions made about a ratees behavior such as whether it was due to ability or effort (e.g., Struthers, Weiner, & Allred, 1998) and (b) rater accountability for making specific ratings (e.g., being more lenient when they have to answer to ratee for rating; Klimoski & Inks, 1990).

Although many of the motivations that Murphy and Cleveland (1995) and Levy and Williams (2004) detail are interpersonal in nature, they are predominantly focused on interactions within the specific organizational environment. However, global interpersonal motivations can also impact performance appraisals. One such category of motivations are those to control one's prejudice in order to improve social interactions and interpersonal relationships by avoiding the negative consequences of being perceived as prejudiced (e.g., Croft & Schmader, 2012; Crosby & Monin, 2007). This motive maps onto Murphy and Cleveland's (1995) rater motivations to avoid delivering negative appraisals. However, the motivations to control prejudice extend beyond the ratee's perception of the rater and can include more global perceptions of the rater inside and outside of the organization. Similar to the rater motivations discussed earlier, motivations to control prejudicial behavior may impact the accuracy of evaluations and motivate the rater to produce an evaluation that is not objective.

Personal motivations may be harmful particularly to underrepresented groups, as these groups tend to receive more negative performance ratings than their majority group counterparts (e.g., Bowen, Swim, & Jacobs, 2000; Lyness & Heilman, 2006; McKay & McDaniel, 2006; Pazy & Oron, 2001; Sackett, DuBois, & Noe, 1991). Research shows raters are prone to overestimating the differences between members of underrepresented groups and majority groups, such that the former tend to have their performance rated more negatively when they were underrepresented in the domain in which they were being evaluated (Pazy & Oron, 2001; Sackett et al., 1991). Furthermore, McKay and McDaniel (2006) found in their meta-analysis that differences in subjective measures of performance (d=.51) for Black versus White employees were larger than differences in objective measures (d=.41), which suggests greater rater bias. However, in today's organizational climate in which many companies are actively promoting diversity initiatives, the focus shifts more to understanding how people's ratings change as the motivation—whether externally or internally driven—to respond (e.g., give feedback) without any prejudice increases.

Rater Motivations to Control Prejudice and Feedback

As with other rater motivations, individual differences in personal motivations to respond without prejudice can—ironically—increase bias while delivering developmental feedback. In particular, a person's internal and external motivation to respond without prejudice may impact the feedback they deliver. *Internal motivation to respond without prejudice* (IMS) results from an internalized and personally important standard to act in a nonprejudiced manner (Plant & Devine, 1998). People high in IMS are motivated to respond in an egalitarian manner and to suppress any indication of prejudices—given that responding in a prejudiced way would violate the nonprejudiced standard to which they personally hold themselves (Butz & Plant, 2009; Plant & Devine, 1998). Conversely, *external motivation to respond without prejudice* (EMS) results from perceived social pressure to comply with egalitarian norms or face social disapproval.

Essentially, IMS and EMS distinguish between a concern with *being* prejudiced (IMS) and a concern with *appearing* prejudiced (EMS; Crosby & Monin, 2007). In part, this distinction between IMS and EMS leads us to predict that rater bias will be manifested differently as it pertains to individual differences in IMS or EMS. In particular, when the motivation to respond in a nonprejudiced way is internalized, the person may feel compelled to go to extreme lengths, even overcorrecting their responses to an individual in order to demonstrate that they are not biased against members of stigmatized groups (Crosby & Monin, 2007; Dutton, 1971; Gaertner & Dovidio, 1981; Mendoza-Denton & Aronson, 2007). Evidence for these patterns of behavior were supported in Crosby and Monin's (2007) research of what they termed a "failure to warn." Within a college setting, students higher in IMS were less likely

to give a hypothetical new student honest feedback and warn them of a potentially difficult first semester course load when the hypothetical student was Black rather than White. Similarly, when White participants were evaluating whether videotaped targets were either lying or telling the truth, Lloyd, Hugenberg, McConnell, Kunstman, and Deska (2017) demonstrated that IMS lead to an overprediction of the truthfulness of Black targets. Further, Kunstman, Plant, Zielaskowski, and LaCosse (2013) found the desire to feel accepted by minority group members to be a driving mechanism in motivations to control prejudice.

Taken together, higher IMS and the associated increased desire to be liked and accepted by an outgroup could motivate raters to provide more lenient evaluations of an individual's performance in order to avoid the cognitive dissonance a ratee or third party's negative reaction to an appraisal would produce. That is, the ratee and other colleagues may perceive a more negative or critical evaluation to be associated with prejudice toward the ratee's outgroup, and as IMS increases, this reaction would be increasingly disconnected from the way that rater sees themselves. We are hypothesizing that elevation in a White rater's IMS will predict more positive feedback delivered to their minority ratee, as the raters would be more likely to want to avoid delivering negative or critical feedback to their ratee because critical feedback may make them appear more prejudiced. Specifically, we hypothesize:

Hypothesis 1: They higher the rater's internal motivation to respond without prejudice, the more likely he/she will be to rate someone's performance positively, even after controlling for that performance.

Conversely, when a person is high in EMS, they are more focused on not *appearing* prejudiced (Plant & Devine, 1998; 2001). As noted earlier, their motivation arises from a desire to avoid behaving in ways that might lead to social disapproval or, in the case of performance appraisals, negative reactions from the ratee or other colleagues. Specifically, in the context of feedback, Croft and Schmader (2012) demonstrate that when evaluating the students' written work, White evaluators who are higher in EMS may provide less critical feedback to minority authors than to White authors.

However, motivation from external pressures can be hard to maintain. Plant and Devine (2001) revealed that when Whites face either real or imagined social pressure to comply with norms to be more positive toward minorities, EMS drove a threatened response and negative affective backlash toward the pressure. Further, Wyer (2007) found that individuals high in EMS who attempted to suppress prejudice experienced "stereotype rebound" in which they exhibited heightened stereotyping behaviors. Each of these studies employ paradigms wherein participants engaged

with representations of hypothetical or imagined minority individuals, and allowed for controlled and corrective behaviors; however, that control likely cannot be maintained. Research suggests that self-regulating one's prejudices while interacting with minorities can be cognitively taxing on an individual (e.g., Richeson & Shelton, 2007; Richeson & Trawalter, 2005), and it may be particularly taxing for those high in EMS who may lack self-determined motivation to regulate their prejudices during the course of social interactions (Butz & Plant, 2009; Plant & Devine, 2001). As noted by Lloyd et al. (2017), espoused egalitarian values regarding prejudice may not manifest during spontaneous responses. Having to deliver the feedback directly to the evaluated individual immediately after their performance could exacerbate this issue.

Therefore, we hypothesize that after engaging with the ratee face to face within a paradigm that produces more than an hour of engagement with a minority, raters higher in EMS will be less regulated in their evaluations and therefore will exhibit less positivity when delivering developmental feedback. Specifically, we hypothesize:

Hypothesis 2: The higher the rater's external motivation to respond without prejudice, the more likely he/she will be to rate someone's performance negatively, even after controlling for that performance.

Current Study

We analyzed data from face-to-face, cross-racial mentorship dyads, a type of developmental relationship, to examine whether motivations to control prejudice would bias the feedback raters (i.e., mentors) give to the evaluated individual on a speech task. Given that we wanted to explore the positivity of rater's feedback over and above the ratee's (i.e., mentee's) performance, we first examined whether assessments of ratee's performance predicted the positivity of raters' feedback to the ratees about their speech performances. Both the assessments of the ratee's performance and of the positivity of the evaluator's feedback were derived from panels of independent judges (i.e., a panel only assessing rater behaviors and one focused only on ratee behaviors) in an effort to get a more objective evaluation of speech performance and feedback positivity. We next sought to understand whether IMS and/or EMS, our variables of interest, significantly predicted how positive or negative a rater's feedback was of their ratee's speeches over and above the ratee's actual speech performance.

METHOD

Participants

We re-analyzed data from Leitner, Ayduk, Boykin, and Mendoza-Denton (2018), wherein 112 participants were randomly assigned to 56 rater/ratee dyads. These partic-

ipants were recruited from social media advertisements, printed signs and flyers, as well as a research participation pool consisting of psychology undergraduate students. Of the 112 participants, 43 were Latino/a, 13 were Black, and 56 were White. There were 68 women and 44 men with an average age of 20.1 years old (SD_{age} = 2.22). Participants were compensated either with cash payment or with course credits. All dyads consisted of same-sex interracial pairs in which one minority (Black or Latino/a) individual was paired with one White individual.

Procedure and Measure.

A research team guided participants through the procedure.

Initial matching. Participants arrived at separate rooms, signed informed consent and media forms (i.e., forms for permissions to record the sessions), then were escorted to the common speech and evaluation task room, at which time they were seated across from their assigned mentoring partner. Because mentoring is viewed as a source of developmental feedback, using this context was a prime for participants to view their feedback as developmental. It is worth noting that all activities in the common speech and evaluation task room were video recorded for future analysis, with separate cameras dedicated to recording raters and ratees respectively. While together in the common speech and evaluation room, participants were informed of their role assignments. In each session, the White participant was assigned to the rater role, and the minority participant was assigned to the ratee role. Although the pattern of role assignments was consistent, raters and ratees were told that their assignment was random as part of the cover story.

Speech task and feedback. After allowing for time for social interaction with their mentoring partner to build familiarity (45 minutes), both mentors (i.e., raters) and mentees (i.e., ratees) were led back to the rooms they initially arrived to and completed a series of individual difference questionnaires and filler tasks. For mentors, these questionnaires included responding to the *external and internal motivations to control prejudice* (Plant & Devine, 1998) measures. Each scale consists of five items, which were scored on a 7-point Likert scale ranging from *strongly disagree* to *strongly agree*. An example IMS item is "I am personally motivated by my beliefs to be nonprejudiced toward Latinos/as." In contrast, an EMS items is "I try to act nonprejudiced toward Latinos/as people because of pressure from others." In dyads where raters interacted with Black ratees, items instead assessed motivations toward Black people. Items assessing IMS and EMS towards Blacks and Latinos/as were pooled together for analysis, and reliability was in the acceptable range for IMS (α = .70, M_{IMS} = 5.79, SD_{IMS} = .92), as well as for EMS (α = .85, M_{EMS} = 3.68, SD_{EMS} = 1.31). Additionally, exploratory factor analysis demonstrated that IMS and EMS loaded onto separate factors.

While in their separate rooms, mentees were informed of the upcoming speech task, in which they were allowed 3 minutes to prepare a 5-minute speech pertaining to why they were qualified for their dream job. Mentees were told that their mentor would provide a written and verbal evaluation of their speech. Concurrently, mentors were informed that they would provide feedback on the speech. Dyads were then reunited in the common room for the speech task and subsequent evaluation.

Mentee speech. Mentees delivered their minimum five-minute (maximum seven-minute) speech to their mentors without notes. If mentees stopped their speech prior to using the minimum five minutes, they were provided prompts via intercom by the experimenter (ex. "What are your long-term career goals?") to help them reach the minimum time requirement. This was done to ensure that mentors had enough content to evaluate mentees' performance.

Independent assessments of mentee speech. Three independent judges from Asian and Middle Eastern ethnic backgrounds (neither Black, Latino/a, nor White to prevent in-group bias) watched videos of mentee speeches only, with each coder being blind to any information regarding the mentor's feedback, hypotheses, or paradigm. Judges were asked to assess mentee speeches along dimensions of speech quality, clarity, and likelihood of being hired based on their speech. The evaluation form was created for the purposes of this study. The quality and clarity of the speeches were each assessed on a 7-point scale ranging from *poor* to *excellent*. Likelihood of being hired was assessed on a 7-point scale ranging from *very unlikely* to *very likely*.

Independent assessments of positivity of mentor feedback. Upon mentees' completion of their speech, mentors were asked to stand and provide verbal feedback to their mentee. Three independent judges (neither Black, Latino/a, nor White) watched video recordings of mentors' feedback, and each judge was blind to information regarding the mentee, hypotheses, or paradigm. They were asked to rate the feedback mentors delivered mentees on a 7-point scale from *very negative* to *very positive*.

RESULTS

We used structural equation modeling to determine whether IMS or EMS predicted mentor's assessments of mentees' speech performance controlling for independent judgments of mentee speech quality. To establish a baseline measurement model, we predicted the positivity of mentor's feedback about mentee's speeches from a higher order factor model of independent judgments of mentee's speech performance. We began by creating first-order latent factors for each independent judge, consisting of their individual assessments of mentees' (a) likelihood of being hired based on speech, (b) speech clarity, and (c) speech quality. To account for between judge variance, we next loaded

the first-order factors from each judge's ratings onto a second-order factor. This higher order factor represented the overall panel's assessments of mentee speech performance. To assess within-rater and between-rater reliabilities, we calculated McDonald's omega estimates for each individual judge (ω_{judge1} = .93; ω_{judge2} = .93; ω_{judge3} = .95) and for the higher order judged performance factor (ω_{panel} = .88; McDonald, 1999; Zhang & Yuan, 2016).

Completing our measurement model, we next predicted the positivity of the feedback that mentors gave to mentees, measured by a latent factor created from three independent panel judgements of the positivity of mentor feedback ($\omega_{positivity}$ = .72), from the hierarchical factor representing panel judgements of mentees speech quality. We included ratee race as a control variable, given than our minority ratees were either Black or Latino. Data were analyzed using MPlus software version 8.1, and we report the results of the fully standardized solution (Muthén & Muthén, 2017). As expected, results demonstrate that independently judged speech performance is a positive predictor of feedback positivity (β = .40, p = .01). Fit indices demonstrate a moderate to good fitting baseline measurement model ($X^2(60)$ = 95.63, p = .00, CFI = .94, TLI = .93, RMSEA = .10, SRMR = .07; Bentler, 1990; Chen, Curran, Bollen, Kirby, & Paxton, 2008).

Last, to test our hypotheses of whether IMS and EMS predicted positivity of feedback we entered composite variables of the mentees mean scores on each measure into our model. As we hypothesized, IMS positively predicted the positivity of the feedback that mentors delivered (β = .51, SE = .12, p = .00; see Figure 1) over and above speech performance, and EMS negatively predicted feedback positivity (EMS: β = -.32, SE = .13 p = .01) in the same model. Results of the hypothesis testing model maintained good to adequate fit ($X^2(102)$ = 762.42, p = .00, CFI = .93, TLI = .91, RMSEA = .10, SRMR = .08; Bentler, 1990; Chen et. al., 2008). Race, entered as a covariate, was a nonsignificant predictor of our outcome variable in both the baseline and test models (p = .22, .37).

DISCUSSION

The expectation is that better speech performance would lead to more positive feedback, and the strength of this association helps preserve a fundamental goal of feedback. Additionally, within the context of cross-racial mentorships with minority mentees, it can be important for mentees development that both positive and negative feedback be delivered with warmth (Boykin, Mendoza-Denton, & Patt, 2015; Leitner et. al. 2018). However, we found that, unfortunately, mentors' motivations to control their prejudice systematically biased the feedback delivered to their mentees. That is, we found that rater motivations were a common source of bias in the evaluation of one's perfor-

mance, and these findings align with previous research (e.g., Murphy & Cleveland, 1995).

To some degree, these results signal that for people who are internally motivated to control prejudice, management of a nonprejudiced self-concept contributes to the inflation of the positivity of feedback provided to minorities. This upward bias in performance assessments, which is not actually related to mentees' performance, can complicate mentees understanding of what is expected of them and diminish their ability to seize opportunities to use feedback to strengthen their skills. Conversely, our results suggest that external motivations to control prejudice can downwardly bias feedback, which can have similar deleterious effects as the upward bias, as it obscures the accuracy of the assessment. Ultimately, neither motivation should play a factor in systematically contributing to the positivity or negativity of mentors' feedback to their mentees, because both could diminish the value of receiving feedback in the first place and, in turn, erode trust in the process.

Implications

In terms of theoretical implications, the present work adds dynamic face-to-face interactions to the observed domains in which motivations to control prejudice can bias feedback delivered to minorities. Existing research in this area has relied largely on photographs, essays, and other proxies for minority contact to understand whether or how IMS and EMS influence behavior. Next, the present work converges with previous findings that IMS can positively bias feedback delivered to minorities (Croft & Schmader, 2012; Crosby & Monin, 2007). Conversely, the present work provides support for previous suggestions that EMS, under the stress of face-to-face interaction, can lead to negativity and in turn negatively biased feedback (Butz & Plant, 2009; Plant & Devine, 2001; Wyer, 2007). Additionally, the findings for both IMS and EMS also expand our understanding of rater motivations in the workplace (Levy & Williams, 2004; Murphy & Cleveland, 1995). Although the main focus of the rater motivation literature to date has focused on rater–ratee dynamics in the workplace, the exploration of IMS and IMS as a rater motivation illustrates the impact more global motivations have on situation-specific evaluations. Although mechanisms underlying these findings need further exploration, their convergence with the motivation to control prejudice literature shows promise for future dyadic work in this area.

In terms of more practical implications, the current study expands on our understanding of how feedback can be affected by factors beyond one's performance, specifically, the influence of a rater's personal motivations to control their prejudices. As more organizations move towards performance management "without ratings" (i.e., processes with more continuous feedback not associated with an annual rating; Adler, Campion, Colquitt, Grubb, Murphy,

FIGURE 1.

Test Model of IMS, EMS, and Speech Performance Predicting Feedback Positivity. $x^2(102) = 762.42$; $p = .00$; CFI = .93; TLI = .91; RMSEA = .10; SRMR = .08. Note. $^*p < .05$, $^{**}p < .01$.
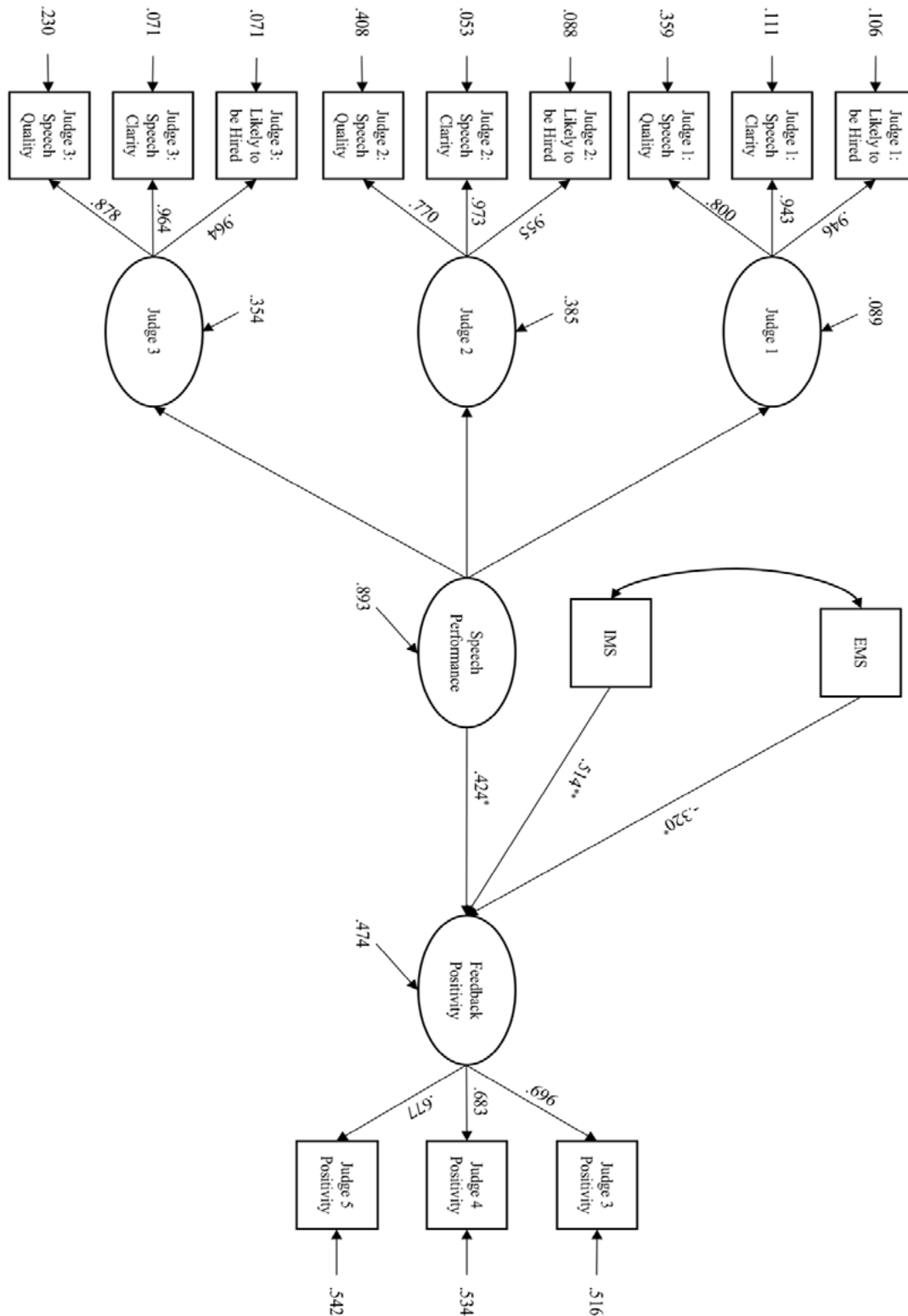
Ollander-Krane, & Pulakos, 2016), these feedback systems need to be monitored closely to understand how motivations may bias the feedback. Furthermore, given the growing focus on strengthening the diversity of organizational leadership pipelines and the impact feedback can have on career advancement (Agars, 2004; Agars & Cazares, 2017; Greenhaus, 1987), it is of particular importance to understand how bias and prejudice may influence feedback given to members of underrepresented groups.

Limitations and Future Research

Conducted in a lab setting, the present study lacks the realism to meet the standard of ecological validity; future research might be conducted in the context of existing organizational mentorship relationships, thus establishing greater robustness and generalizability of these findings. As organizations continue to increase their focus on diversity, employees are likely to feel an increased pressure to respond to different work situations in unprejudiced ways. That is, they are likely to experience increases to their internal and/or external motivations to control prejudices in order to behave in a workplace-sanctioned way.

A second limitation is that although the current study explored the motivations to control prejudice, it did not explore the individual's actual prejudices. As noted earlier, motivation to control prejudice does not equate to the motivated individual actually being prejudiced (Dunton & Fazio, 1997; Plant & Devine, 1998); however, given that the current study does not explore the individual's prejudice, the impact of the intersection of rater prejudice and the rater motivations to control that prejudice on the feedback given cannot be explored and should be explored in future research. Similarly, prior research has shown the quality of feedback is also susceptible to intra- and interpersonal influences such as perceived similarity (Eby et al., 2013), ratee attachment style (Allen et al., 2010), and rapport between rater and ratee (Leitner et al., 2018); however, although such variables were not explored in this study, future research should consider how these variables interact with motivations to control prejudice to impact feedback.

Although the current sample was too small to test such relationships, future research should examine how different profiles of IMS and EMS could create different feedback responses. Given that IMS and EMS are theoretically independent motivations (Crosby & Monin, 2007), a person can be high in both or low in both, and these different profiles may lead individuals to deliver more or less accurate feedback because of it. Additionally, Devine, Forscher, Austin, and Cox (2012) have demonstrated that interventions could help mitigate the impact of motivations to control prejudice on behavior. It is possible that such interventions could help individuals better manage these motivations to deliver more accurate feedback to their ratees.

Conclusions

Accurate assessments of task performance in developmental relationships are vital to optimizing the benefits of feedback for future performance. Performance assessments for underrepresented minorities can be skewed by bias related to outgroup prejudice; however, the potential contributions of motivations to control prejudice have been underexplored, particularly in face-to-face settings. The current work provides support for our assertion that IMS can positively bias feedback while EMS can negatively bias feedback, both of which can compromise feedback accuracy. For individuals, increased societal and organizational focuses on implicit and explicit biases and prejudices can lead to increases in both internal as well as external motivations to control prejudice. Further, increased knowledge of biases may result in inaction, suppression of bias, avoidance of intergroup interaction, or other behaviors that are detrimental to cross-racial engagement broadly (Daumeyer, Onyeador, Brown, & Richeson, 2019; Onyeador, 2017). It may important to focus on destigmatizing bias in favor of recognizing that all individuals can grow into better allies and support systems for underrepresented minorities and in turn provide more accurate developmental feedback.

REFERENCES

Adler, S., Campion, M., Colquitt, A., Grubb, A., Murphy, K., Ollander-Krane, R., & Pulakos, E.D. (2016). Getting rid of performance ratings: Genius or folly? A debate. Industrial and Organizational Psychology: Perspectives on Science and Practice, 9(2), 219-252.

Agars, M. D. (2004). Reconsidering the impact of gender stereotypes on the advancement of women in organizations. Psychology of Women Quarterly, 28, 103–111.

Agars, M., & Cazares, E. (2017). A cycle or a ceiling? The cumulative effects of subtle discrimination through the lens of performance management. Industrial and Organizational Psychology : Perspectives on Science and Practice, 10(1), 97-100.

Aguinis, H. (2009). An expanded view of performance management. In J. Smithers & M. London (Eds.), Performance management: Putting research into action (pp. 1-43). San Francisco, CA: Jossey-Bass.

Allen, T. D., Shockley, K. M., & Poteat, L. (2010). Protégé anxiety attachment and feedback in mentoring relationships. Journal of Vocational Behavior, 77, 73-80.

Bentler, P. M. (1990). Comparative fit indexes in structural models. Psychological Bulletin, 107, 238–246.

Bowen, C., Swim, J. K., & Jacobs, R. R. (2000). Evaluating gender biases on actual job performance of real people: A meta-analysis. Journal of Applied Social Psychology, 30, 2194–2215.

Boykin, C. M., Mendoza-Denton, R., & Patt, C. (2015). Prejudice

and stigma: Implications for student achievement and mentoring. In C. Rubie-Davies (Ed.), The social psychology of the classroom international handbook (pp. 93-101). New York, NY: Routledge.

Butz, D.A., & Plant, E.A. (2009). Prejudice control and interracial relations: The role of motivation to respond without prejudice. Journal of Personality, 77, 1311-1341.

Chen, F., Curran, P. J., Bollen, K. A., Kirby, J., & Paxton, P. (2008). An empirical evaluation of the use of fixed cutoff points in RMSEA test statistic in structural equation models. Sociological Methods & Research, 36, 462-494.

Cleveland, J. N., & Murphy, R. E. (1989). Multiple uses of performance appraisal: Prevalence and correlates. Journal of Applied Psychology, 74, 130-135.

Croft, A., & Schmader, T. (2012). The feedback withholding bias: Minority students do not receive critical feedback from evaluators concerned about appearing racist. Journal of Experimental Social Psychology, 48, 1139-1144.

Crosby, J. R., & Monin, B. (2007). Failure to warn: How student race affects warnings of potential academic difficulty. Journal of Experimental Social Psychology, 43, 663-670.

Daumeyer, N. M., Onyeador, I. N., Brown, X., & Richeson, J. A. (2019). Consequences of attributing discrimination to implicit vs. explicit bias. Journal of Experimental Social Psychology, 84. Advance online publication. https://doi.org/10.1016/j.jesp.2019.04.010

DeRue, D. S. & Wellman, N. (2009). Developing leaders via experience: The role of developmental challenge, learning orientation, and feedback availability. Journal of Applied Psychology, 94, 859-875.

Devine, P. G., Forscher, P. S., Austin, A. J., & Cox, W. T. L. (2012). Long-term reduction in implicit race bias: A prejudice habit-breaking intervention. Journal of Experimental Social Psychology, 48, 1267-1278.

Dutton, D. G. (1971). Reaction of restaurateurs to Blacks and Whites violating restaurant dress requirements. Canadian Journal of Behavioral Science, 3, 298–302.

Dunton, B. C., & Fazio, R. H. (1997). An individual difference measure of motivation to control prejudiced reactions. Personality and Social Psychology Bulletin, 23, 316-326.

Eby, L. T., Allen, T. D., Hoffman, B. J., Baranik, L. E., Sauer, J. B., Baldwin, S., … Evans, S. (2013). An interdisciplinary meta-analysis of the potential antecedents, correlates, and consequences of protégé perceptions of mentoring. Psychological Bulletin, 139, 441-476.

Gaertner, S. L., & Dovidio, J. F. (1981). Racism among the well-intentioned. In E. G. Clausen & J. Bermingham (Eds.), Pluralism, racism, and public policy: The search for equality (pp. 208–222). Boston, MA: G.K. Hall and Company.

Greenhaus, J. H. (1987). Career management. Hinsdale, IL: Dryden Press.

Klimoski, R., & Inks, L. (1990). Accountability forces in performance appraisal. Organizational Behavior and Human Decision Processes, 45, 194-208.

Kudisch, J. D., Fortunato, V. J., & Smith, A. F. R. (2006). Contextual and individual difference factors predicting individuals' desire to provide upward feedback. Group & Organization Management, 31, 503-529.

Kunstman, J. W., Plant, E. A., Zielaskowski, K., & LaCosse, J. (2013).

Feeling in with the outgroup: Outgroup acceptance and the internalization of the motivation to respond without prejudice. Journal of Personality and Social Psychology, 105, 443-457.

Leitner, J. B., Ayduk Ö, Boykin C. M., & Mendoza-Denton, R. (2018) Reducing negative affect and increasing rapport improve interracial mentorship outcomes. PLoS ONE, 13(4).

Levy, P. E., & Williams, J. R. (2004). The social context of performance appraisal: A review and framework for the future. Journal of Management, 30, 881-905.

Lloyd, E. P., Hugenberg, K., McConnell, A. R., Kunstman, J. W., & Deska, J. C. (2017). Black and white lies: Race-based biases in deception judgments. Psychological Science, 28, 1125-1136.

Lyness, K. S., & Heilman, M. E. (2006). When fit is fundamental: Performance evaluations and promotions of upper-level female and male managers. Journal of Applied Psychology, 91, 777–785.

McDonald, R. P. (1999). Test theory: A unified treatment. Mahwah, NJ: Lawrence Erlbaum.

McKay, P. F., & McDaniel, M. A. (2006). A reexamination of Black–White mean differences in work performance: More data, more moderators. Journal of Applied Psychology, 91, 538-554.

Mendoza-Denton, R., & Aronson, J. (2007). Psychosocial processes in the cultivation of intellective competence: The interpenetration of affective, cognitive, and situative processes in intellective behavior. In E.W. Gordon & B.L. Bridglall (Eds.), Affirmative development: Cultivating academic ability (pp. 189-205). Lanham, MD: Rowman and Littlefield.

Murphy, K., & Cleveland, J. (1995). Understanding performance appraisal: Social, organizational and goal-oriented perspectives. Newbury Park, CA: Sage.

Murphy, K. R., Cleveland, J. N., Skattebo, A. L., & Kinney, T. B. (2004). Raters who pursue different goals give different ratings. Journal of Applied Psychology, 89, 158-164.

Muthén, B., & Muthén, L. (2017). Mplus version 8.1 [software]. Los Angeles, CA: Author.

Ng, K., Koh, C., Ang, S., Kennedy, J. C., & Chan, K. (2011). Rating leniency and halo in multisource feedback ratings: Testing cultural assumptions of power distance and individualism-collectivism. Journal of Applied Psychology, 96, 1033-1044.

Onyeador, I. N. (2017). Presumed unintentional: The ironic effects of implicit bias framing on Whites' perceptions of discrimination (Unpublished doctoral dissertation). University of California, Los Angeles, CA.

Pazy, A., & Oron, I. (2001). Sex proportion and performance evaluation among high-ranking military officers. Journal of Organizational Behavior, 22, 689–702.

Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. Journal of Personality and Social Psychology, 75, 811–832.

Plant, E. A., & Devine, P.G. (2001). Responses to other-imposed pro-black pressure: Acceptance or backlash? Journal of Experimental Social Psychology, 37, 486-501.

Richeson, J. A., & Shelton, J. N. (2007). Negotiating interracial interactions: Costs, consequences, and possibilities. Current Directions in Psychological Science, 16, 316–320.

Richeson, J. A., & Trawalter, S. (2005). Why do interracial interactions impair executive function? A resource depletion account. Journal of Personality and Social Psychology, 88, 934–947.

Sackett, P. R., DuBois, C. L., & Noe, A. W. (1991). Tokenism in performance evaluation: The effects of work group representation on male-female and White-Black differences in performance ratings. Journal of Applied Psychology, 76, 263–267.

Struthers, C. W., Weiner, B., & Allred, K., (1998). Effects of causal attributions on personnel decisions: A social motivation perspective. Basic & Applied Social Psychology, 20, 155-166.

Wang, X. M., Wong, K. F. E., & Kwong, J. Y. Y. (2010). The roles of rater goals and ratee performance levels in the distortion of performance ratings. Journal of Applied Psychology, 95, 546-561.

Wong, K. F. E., & Kwong, J. Y. (2007). Effects of rater goals on rating patterns: Evidence from an experimental field study. Journal of Applied Psychology, 92, 577.

Wyer, N. A. (2007). Motivational influences on compliance with and consequences of stereotype suppression. Journal of Experimental Social Psychology, 43, 417–424.

Zhang, Z., & Yuan, K. H. (2016). Robust coefficients alpha and omega and confidence intervals with outlying observations and missing data: Methods and software. Educational and Psychological Measurement, 76, 387-411.

Zhou, J. (2003). When the presence of creative coworkers is related to creativity: Role of supervisor close monitoring, developmental feedback, and creative personality. Journal of Applied Psychology, 88, 413-422.