Confronting stereotypic biases: Does internal versus external motivational framing matter?

Mason D. Burns and Margo J. Monteith

Abstract
We investigated whether confrontations of intergroup bias that had an external (e.g., emphasizing social norms) versus internal (e.g., emphasizing values) motivational framing differentially reduced subsequent stereotyping. Internally and externally framed confrontations reduced stereotyping equally compared to a control condition, both immediately (Experiments 1 and 2) and across a 2- to 3-day delay (Experiment 1). Only weak evidence was found for a “matching hypothesis” when participants own chronic internal and external motivations to respond without prejudice were assessed. Confrontation framing did not interact with chronic motivations to affect stereotyping in Experiment 2. In Experiment 3, participants highly internally motivated to respond without bias reduced bias most with an internally framed confrontation, whereas participants who were not motivated for internal reasons reduced bias most with an externally framed confrontation. Finally, whereas both motivational framings reduced stereotyping, simply pointing bias out did not. Thus, providing some motivational framing is important for confrontation effectiveness.

Keywords
confrontation, motivation, prejudice/stereotyping, racism, sexism

Speaking up against bias in its various forms is essential for equality, justice, and tolerance. Confrontation of bias has been foundational to civil rights campaigns and organizations, and it is promoted as a way to foster inclusivity (e.g., Willoughby, 2015). However, research indicates that people often do not speak up and confront others’ biased words and actions (Swim & Hyers, 1999; Woodzicka & LaFrance, 2001).

Confronting other’s bias requires people to clear a number of hurdles, including knowing just what to say (Ashburn-Nardo, Morris, & Goodwin, 2008). Is the framing of a confrontation important, or will any confrontation do? We investigated

1University of Indianapolis, USA
2Purdue University, USA

Corresponding author:
Mason D. Burns, School of Psychological Sciences, University of Indianapolis, 1101 E 10th Street, Indianapolis, IN 46227, USA.
Email: burnsmd@uindy.edu
whether the motivational framing of a confrontation, in terms of providing either internal or external reasons for avoiding bias, differentially affected people’s postconfrontation reliance on stereotypes. Additionally, we tested whether the effects of confrontation framing on subsequent stereotyping depended on several, theoretically relevant factors. First, we examined whether confrontations with an internal or external motivation framing produced only immediate or also longer lasting reduction in stereotyping (Experiment 1). Second, we investigated whether the effect of motivational framing on stereotype reduction depended on whether race or gender bias was confronted (Experiment 2). Third, we investigated whether confrontations varying in motivational framing were differentially effective depending on recipients’ preexisting internal and external motivations to respond without prejudice (Experiments 2 and 3; see also Plant & Devine, 1998; ). Finally, we investigated the added benefit of motivational framings beyond a confrontation that did not include any motivational framing (Experiment 3).

**Confrontation of Stereotyping**

Confrontation of bias can be an effective tool for raising people’s awareness about their biases (Gulker, Mark, & Monteith, 2013) and encouraging them to monitor and reduce biased responding (Czopp, Monteith, & Mark, 2006). Because many people lack awareness of their own biases (Perry, Murphy, & Dovidio, 2015), a newfound awareness resulting from confrontation may lead people to self-regulate their biased responses (Monteith, 1993; Monteith, Ashburn-Nardo, Voils, & Czopp, 2002). Furthermore, people may be aware of their biased responses, but may not consider them problematic until they are confronted.

Research has begun to reveal factors that influence whether confrontations will be more or less impactful. For instance, confrontations appear to be perceived as more legitimate and elicit greater message acceptance when delivered by allies (e.g., White people confronting racism) than by targets of bias (Drury & Kaiser, 2014; Gulker et al., 2013). However, little research has examined how the framing of confrontations influences confrontation effectiveness. One important factor may be whether confrontations are worded in ways that emphasize internal or external reasons for perpetrators to reduce their stereotyping.

**Motivational Framings of Confrontations**

Research on prejudice has revealed two underlying motivational factors that can encourage people to respond without bias (Dunton & Fazio, 1997; Plant & Devine, 1998). First, an internal motivation to respond without prejudice entails avoiding biased responses so as to respond consistently with an egalitarian self-concept. Second, an external motivation to respond without prejudice prompts nonbiased responding to conform to the demands of others and to avoid their disapproval. Past research has investigated whether individual differences in the extent to which people endorse internal and external motivations predict biased responding (e.g., Devine, Plant, Amodio, Harmon-Jones, & Vance, 2002; Plant & Devine, 1998). In contrast, the main focus of the current research concerns whether confrontations worded to emphasize internal or external reasons to avoid stereotyping will differentially reduce a perpetrator’s stereotyping.

We crafted confrontations with internal versus external motivational framings based on Plant and Devine’s (1998) framework for understanding internal and external motivations to respond without prejudice, and we also considered self-determination theory’s conceptualization of autonomy-supportive versus autonomy-restrictive communications (Deci & Ryan, 1985, 2000; see also Legault, Green-Demers, Grant, & Chung, 2007). Accordingly, we conceptualized a confrontation with an internal framing as pointing out the biased response(s) to the target, and then taking a constructive and supportive approach to emphasize the target’s choice to respond in more egalitarian ways to fulfill personal values and achieve
positive outcomes (e.g., a fair society). We conceptualized a confrontation with an external framing as likewise pointing out the biased response(s), but exerting pressure and demanding that the target respond with less bias or risk violating social norms and eliciting negative reactions from others.

In related research, Legault, Gutsell, and Inzlicht (2011) prepared antiprejudice brochures that urged participants to adopt nonprejudiced attitudes. In one condition, the brochure emphasized that people can choose to be nonprejudiced to fulfill their values and to foster positive relations with others (“autonomy brochure”). In another condition, the brochure emphasized that participants must be nonprejudiced to meet societal demands or otherwise experience negative interpersonal outcomes (“controlling brochure”). Legault et al. found that participants who read the autonomy brochure reported less prejudiced attitudes than participants in a control condition. In contrast, participants who read the controlling brochure showed backlash, or increased prejudiced attitudes, relative to the control condition. Our research differs from Legault et al.’s (2011) in that we studied whether confronting stereotypical responses that participants generated during an experimental task reduced participants’ subsequent stereotype application (i.e., behavior). In contrast, participants in Legault et al.’s research had not generated biased responses that were confronted; rather, they read a brochure advocating nonprejudice (or, in Study 2, were primed with autonomous vs. controlling motivation for being nonprejudiced), and the effects on attitudes were assessed. Given these differences, will our results be similar to or different from those of Legault and colleagues?

On one hand, similar to Legault et al.’s (2011), we may observe greater bias reduction with an internal than external framing. Whereas an internal framing appeals to fairness and egalitarian values that most people endorse (Monteith & Walters, 1998; Rokeach, 1973), external, other-imposed pressure can produce reactance (Brehm & Brehm, 1981; Plant & Devine, 2001). On the other hand, because our research involved confrontation of biased responses that participants had just generated, and assessed behavioral rather than attitudinal change, we may find that both an internal and external framing reduces participants’ subsequent use of stereotypes. Prior research clearly indicates that external pressures and social norms can reduce expressions of bias (Blanchard, Crandall, Brigham, & Vaughn, 1994; Crandall & Eshleman, 2003; Monteith, Deneen, & Tooman, 1996). Furthermore, our confrontation addressed subtly biased stereotypic responses that participants may not have even realized they were generating, which differs considerably from Legault et al.’s paradigm for motivating nonprejudice and measuring prejudiced attitudes. Even a confrontation with an external motivational framing may effectively stop specific biased behavior in its tracks. In sum, we tested competing predictions that confrontation with internal and external motivational framings would be equally effective at reducing participants’ use of stereotypes, relative to a control condition, versus greater effectiveness for the internal than the external framing.

Overview of Present Research
Three experiments investigated whether participants who were initially confronted about their use of stereotypes during an experimental task were later differentially likely to reduce stereotype application, depending on the motivational framing of the confrontation. Experiment 1 examined confrontations of stereotypic responding toward Blacks, and investigated confrontation effectiveness both immediately and after a 2- to 3-day delay. Inclusion of the delayed condition allowed us to assess whether confrontation effects dissipate differently across times as a function of confrontation framing (see Deci & Ryan, 1985). Experiment 2 compared motivational framing effects of confrontation in the context of racial versus gender stereotyping. Finally, Experiment 3 (addressing gender stereotyping) included a condition where participants were confronted but without a motivational framing. Across experiments, we tested competing predictions that internal and external framings would be equally
effective versus an internal framing would be more effective.

In addition, because prior research indicates that messages are more persuasive when they match recipients’ motivations and personality than when they do not (e.g., Cesario, Higgins, & Scholer, 2008; Hirsh, Kang, & Bodenhausen, 2012), we also explored the effects of participants’ personal motivations to respond without prejudice. Specifically, participants who score higher versus lower on the Internal Motivation to Respond Without Prejudice Scale (IMS; Plant & Devine, 1998) may be more responsive to confrontations with an internal framing. Similarly, participants who score higher versus lower on the External Motivation to Respond Without Prejudice Scale (EMS; Plant & Devine, 1998) may be more responsive to an external framing.

**Experiment 1**

**Method**

**Participants.** Participants were non-Black (determined based on prescreening) undergraduate students who received course credit toward their Introduction to Psychology grade. Data from five participants who did the photo-description task incorrectly were deleted, as were data from eight participants who had considerable missing data. The final sample consisted of 166 participants (51.8% men; 97.0% White, 1.2% Asian, 0.6% American Indian, and 1.2% missing; $M_{age} = 19.55$, $SD = 1.19$). A sensitivity power analysis performed with G*Power 3.1.9.2 (Faul, Erdfelder, Buchner, & Lang, 2009) indicated that this sample size provided 80% power to detect an effect size of $f = .24$ (just below medium effect size cut-off) or greater. Given similar prior research found large confrontation effects (e.g., Chaney & Sanchez, 2018; Czopp et al., 2006), the present research appears sufficiently powered to detect effects of interest.

**Design.** A 2 (testing session: immediate vs. delayed) x 3 (condition: control vs. internal framing vs. external framing) between-participants design was used.1

**Procedure.** Up to seven participants per session were brought to the lab and seated at individual computer stations. Participants consented to the study, which purportedly investigated analytic reasoning ability. Participants first completed an inference task for eliciting stereotyping (see Burns, Monteith, & Parker, 2017). Across 20 trials, they saw a photograph of a person and a brief description, and were instructed to type a fitting one- or two-word inference. For instance, a photograph of a White man described by, “This person can be found in a theater,” might elicit “ACTOR” or “MOVIE FAN.” Participants were encouraged to respond quickly, typing the first logical response that came to mind, although a specific time restriction was not imposed.

Based on random assignment, participants received one of two sets of items, both of which included three critical trials. Critical trials included a photograph of a Black man and a description that could elicit a stereotypic response. For example, for “This person can be found behind bars,” responses such as “CRIMINAL” were stereotype-consistent, and responses such as “BARTENDER” were nonstereotypic (see supplemental materials for details).

**Confrontation manipulation.** Participants received fixed feedback that they demonstrated “slightly above average” analytic reasoning ability. This feedback was intended to be neutral given that people generally like to think of themselves as better than average (Alicke, Klotz, Breitenbecher, Yurak, & Vredenburg, 1995). Control participants proceeded to the next task. Confronted participants read that their responses revealed additional information and saw a list (e.g., deductive errors, especially slow responding) with an “X” next to “Racial Bias.” The following explanation was then provided (tailored to the critical trials actually received):

This means that you likely generated responses to photographs of Blacks that reflected racial stereotypes. You may have assumed that a Black man paired with “this person can be found behind bars” is a criminal rather than
being a bartender, that a Black person who “can be found on the street” is a bum or homeless person rather than a tourist or policeman, or that a Black person who “uses needles for recreation” is a drug addict rather than a tattoo artist.

Although the specific stereotypic responses noted in the feedback were not tailored to exactly what participants’ had typed, most people generated at least one and often more stereotypic responses in this task, and found the feedback to be believable (Burns et al., 2017).

Internal framing:

Such racial biases are unfortunate among people who strive for social justice. You can choose to think about Blacks in more open-minded ways. If you choose to avoid racial stereotyping, you will be able to benefit more from positive interactions with diverse people and contribute to an egalitarian society. You will also be contributing to a free, fair, and peaceful society.

External framing:

Such racial biases are unacceptable among people who strive for political and social correctness. They violate societal norms to not be racist. People are not going to like you, they may not hire you, and they may prevent you from joining their organizations if you continue to think about Blacks in stereotypical ways. You really should not rely on stereotypes of Blacks if you want to fit in with today’s anti-racism norms.

Next, all participants completed 10 filler items to boost the cover story.\(^2\)

**Testing session manipulation.** Participants in the immediate condition completed the dependent measures next, whereas participants in the delayed condition were informed they would receive an email in 2 days to complete other tasks online. On average, these participants completed the experiment 2.54 (\(SD = 1.08\)) days after the initial session and attrition was minimal (\(n = 4\)).

**Dependent measures.** Participants in both testing session conditions learned that data were being collected on a new set of photo-description pairs for norming purposes, and responses would go into an anonymous database. Participants were informed that they would not receive feedback based on their responses. Participants then completed a second set of 20 photo-description pairs, including three trials that served as the measure of postconfrontation stereotyping. Participants who had completed Set 1 during the preconfrontation phase, completed Set 2 postconfrontation, and vice versa.

After completing the second set of photo-description pairs, recall for the confrontation framing was assessed. Across two items, confronted participants reported (1 = *strongly disagree*, 7 = *strongly agree*) their recollection of their confrontation framing’s focus (“My feedback emphasized that if I choose to respond in ways that are not biased, I will have more positive interactions with diverse people and contribute to an egalitarian society”; “My feedback emphasized that racial biases violate norms of what is socially and politically correct, so I might not be hired or accepted into organizations if I respond in racially biased ways,” respectively).

Finally, as in all experiments herein, participants were thoroughly debriefed and were given the option to disallow use of their data.

**Results**

**Confrontation Framing Recall**

A 2 (testing session: immediate vs. delayed) x 2 (confrontation: internal vs. external framing) ANOVA predicting the extent to which participants recalled their feedback as internally or externally framed was conducted with only participants in the confrontation conditions. The interaction between confrontation condition and testing session was significant, \(F(1, 105) = 6.12,\)
In the immediate condition, the internally framed confrontation ($M = 5.34, SD = 1.57$) was recalled as focusing on internal reasons more than the external framing confrontation ($M = 3.93, SD = 1.88$), $t(105) = 3.27, p = .002, d = 0.81, 95\% \text{CI} [0.27, 1.36]$. However, in the delayed condition, the internal ($M = 5.36, SD = 1.42$) and external ($M = 5.48, SD = 1.61$) framing confrontations were recalled as focusing on internal reasons similarly ($p = .784$). This finding suggests that, over time, participants in the external framing condition recalled their confrontation as appealing to internal reasons more than it actually did.

When predicting participants’ recollection of their feedback’s external focus, the effect of confrontation was significant, $F(1, 105) = 10.69, p = .001, \eta^2_p = .09$. The internal framing ($M = 4.18, SD = 1.90$) was recalled as less externally focused than the external framing ($M = 5.31, SD = 1.67$).

**Stereotypic Responses**

Responses to the photo-description tasks were independently coded by both authors in this and subsequent experiments (nonstereotypic = 0, stereotypic = 1) with consistently high agreement ($\geq 98\%$). Discrepancies were resolved through discussion.

**Preconfrontation stereotyping.** On average, participants provided 1.75 ($SD = 0.87$) stereotypic responses across the three preconfrontation prompts, and few ($n = 12$) provided none. Across all experiments, excluding participants who provided no preconfrontation stereotypic responses did not change the results, and were retained.

A 2 (testing session: immediate vs. delayed) x 3 (confrontation: control vs. internal framing vs. external framing) ANOVA performed on preconfrontation stereotyping revealed that stereotyping did not differ across confrontation condition ($p = .439$), nor were any other effects significant.

**Postconfrontation stereotyping.** A 2 (testing session: immediate vs. delayed) x 3 (confrontation: control vs. internal framing vs. external framing) ANCOVA predicting postconfrontation stereotyping was performed, with preconfrontation stereotyping entered as a covariate. Preconfrontation stereotyping was a significant covariate, $F(1, 159) = 10.67, p < .001, \eta^2_p = .06$. As expected, the effect of confrontation was significant, $F(2, 159) = 8.50, p < .001, \eta^2_p = .10$ (see Table 1). The two motivational framing conditions did not differ from one another ($p = .456$). Neither the effect of testing session ($p = .185$), nor the interaction between confrontation and testing session ($p = .425$) was significant.

In sum, the external framing reduced stereotyping just as much as the internal framing. Moreover, both motivational framings were influential whether stereotyping was assessed immediately and in the same laboratory session, or 2–3 days later and in another setting. Interestingly, participants in the delayed testing session recalled the internal and external framings as similarly focusing on the internal reasons to avoid stereotyping. Perhaps the external framing confrontation was reconstrued across time, or perhaps participants simply forgot its content. Regardless, the results support the efficacy of motivational confrontations drawing on external and internal
reasons for curbing stereotyping both immediately and across time.

**Experiment 2**

Experiment 2 investigated confrontations of both sexism and racism. Past research suggests that people feel less guilty after imagining a confrontation involving sexism than racism (Czopp & Monteith, 2003), and are less persuaded by confrontations concerning gender than racial bias (Gulker et al., 2013). These differences have been explained by noting that racism is normatively viewed as more serious than sexism (Fiske & Stevens, 1993; Woodzicka, Mallett, Hendricks, & Pruitt, 2015), and that the prescriptively “positive” nature of stereotypes toward women leads people to dismiss contentions of sexism (Glick & Fiske, 1996). Given weaker norms against sexism, we considered whether a confrontation emphasizing external reasons to avoid gender bias would be relatively ineffective when compared to racial bias. On the other hand, recent research has found that gender-bias confrontations that present people with concrete evidence of their bias, thereby making it difficult to dismiss, can prompt concern over responding in gender-biased ways (Parker, Monteith, Moss-Racusin, & van Camp, 2018). Because our confrontation paradigm involved pointing out stereotypic responses to the photo-description task and noting nonstereotypic alternatives, internally and externally focused confrontations might be similarly effective at reducing gender- and race-based stereotyping.

Additionally, Experiment 2 investigated whether participants’ chronic internal and external motivations to respond without prejudice moderated the effect of confrontations varying in motivational focus. People may be more responsive to a confrontation with a motivational framing that matches their personally held motivations.

**Method**

**Participants**

Participants were recruited from MTurk in return for $0.30. Participation was restricted to people who had not completed other research from our laboratory that had similar materials or involved confrontation. Thirteen participants completed the photo-description task incorrectly, and seven participants marked that their data could not be used on a postsession consent form that explained the use of deception. These data were evenly spread across conditions, and they were deleted. This left 245 participants (71% women; 81.9% White, 3.4% Black, 8.3% Asian, 4.5% Hispanic, and 1.9% “other”; $M_{age} = 38.68$, $SD = 14.65$). A sensitivity power analysis was performed with G*Power 3.1.9.2 (Faul et al., 2009), specifying a fixed model linear regression for $R^2$ increase, with $n = 245$, $\alpha = .05$, and 80% power. The analysis was performed for the highest order interaction (i.e., Confrontation x Bias Type x IMS x EMS), and it indicated that the minimum effect size that could be detected was $f^2 = .04$ (i.e., a small effect). Similar prior research (Parker et al., 2018) has
found small- to medium-sized interaction effects between IMS, EMS, and confrontation conditions. Thus, Experiment 2 appeared adequately powered to detect interaction effects if present.

**Design**

A 2 (type of bias: gender vs. race) x 3 (confrontation: none vs. internal framing vs. external framing) between-participants design was used, and IMS and EMS varied continuously.

**Procedure**

The procedure was identical to Experiment 1 except where noted.

**Bias manipulation.** Participants randomly completed a race or gender version of the procedure. The same critical photo-description pairs from Experiment 1 were used in the race-bias condition. Critical trials were replaced by gender photo-description pairs in the gender-bias condition, which played on role restrictive, lower status, stereotypes about women. For instance, “This person can often be found in a hospital” may yield a stereotypic response like “NURSE” rather than a nonstereotypic alternative response like “DOCTOR.”

**Confrontation manipulation.** Regardless of bias type, all participants saw that they performed “slightly above average” and nothing else. Confronted participants received a list of potential errors, with a red “X” next to “Racial Bias” or “Gender Bias,” depending on condition. As in Experiment 1, an explanation of what constitutes stereotypic responding on task, with examples, followed. Confronted participants then read the internal or external framing from Experiment 1 tailored depending on bias type condition. Participants then completed 10 filler items consistent with the cover story.

**Dependent measures.** Next, participants completed 20 postconfrontation photo-description pairs (Set 1 or 2, depending on preconfrontation set).

Participants in the race-bias condition then completed the race-relevant IMS and EMS items (Plant & Devine, 1998). Participants in the gender-bias condition completed gender-relevant versions (e.g., Klonis, Plant, & Devine, 2005). IMS ($M = 5.58, SD = 1.42, \alpha = .94$) and EMS ($M = 3.38, SD = 1.59, \alpha = .89$) were completed on a 7-point scale (1 = strongly disagree, 7 = strongly agree). Finally, confronted participants completed the same items from Experiment 1 concerning recall of the confrontation.5

**Results**

**Preliminary Analyses**

We performed 2 (bias type: race vs. gender) x 3 (confrontation: control vs. internal framing vs. external framing) ANOVAs to test whether condition predicted IMS and EMS scores, which would indicate they could not be used as predictors. No significant results were obtained ($F$s < 1, $p$s > .386), and IMS and EMS were mean centered for subsequent analyses. Note that hierarchical linear regression analyses were used in subsequent analyses involving IMS and EMS because these variables were continuous predictors.

**Confrontation Framing Recall**

Recall data among participants in the confrontation conditions were analyzed in a hierarchical regression with bias type (race bias = 0, gender bias = 1), confrontation type (internal framing = 0, external framing = 1), IMS, and EMS entered on Step 1, two-way interactions entered on Step 2, and three-way interactions as well as the four-way interaction on Step 4. Only the main effects for confrontation were significant. The internal framing was recalled as more internally focused ($M = 5.15, SD = 1.49$) than the external framing ($M = 4.15, SD = 2.36$), $t(161) = 3.32, B = -0.99, 95\% CI [-1.59, -0.40], \beta = -0.25, p = .001$. The external framing was recalled as more externally focused ($M = 5.68, SD = 1.87$) than the internal framing ($M = 3.58, SD = 2.16$), $t(161) = 7.22, B = 2.15, 95\% CI [1.56, 2.74], \beta = 0.47, p < .001$. 


Plan of Analyses

Confrontation condition was represented by two dummy codes with the control condition serving as the reference group (West, Aiken, & Krull, 1996; control condition: DC1 = 0, DC2 = 0; internal framing: DC1 = 1, DC2 = 0; external framing: DC1 = 0, DC2 = 1). Thus, DC1 compared the internal framing and control conditions, and DC2 compared the external framing and control conditions. Dummy codes were entered as a set, and the increment in $R^2$ for the set was assessed for significance at each step. When a set involving dummy codes was significant, analyses were rerun with modified dummy codes that allowed comparisons between the internal and external framings (see West et al., 1996). Bias type was dummy coded as described previously.

Analyses used the following steps: 1: IMS, EMS, and bias type; 2: DC1 and DC2; 3: DC1 x Bias Type and DC2 x Bias Type; 4: DC1 x IMS and DC2 x IMS; 5: DC1 x EMS and DC2 x EMS; 6: Bias Type x IMS, Bias Type x EMS, and IMS x EMS; 7: DC1 x Bias Type x IMS and DC2 x Bias Type x IMS; 8: DC1 x Bias Type x EMS and DC2 x Bias Type x EMS; 9: Bias Type x IMS x EMS; and 10: DC1 x Bias Type x EMS x IMS, DC2 x Bias Type x EMS x IMS.

Stereotypic Responses

Preconfrontation stereotyping. On average, participants provided 1.88 ($SD = 1.02$) stereotypic responses across the preconfrontation prompts, and relatively few ($n = 29$) provided none. The hierarchical regression analysis predicting preconfrontation stereotyping revealed a significant effect for bias type, $t(241) = 9.68, B = 1.07, 95\% CI [0.85, 1.29], \beta = 0.52, p < .001$, such that women were stereotyped more ($M = 2.43, SD = 0.67$) than Black men ($M = 1.37, SD = 1.03$). Additionally, as participants’ IMS increased, their preconfrontation stereotyping decreased, $t(241) = 2.83, B = -0.11, 95\% CI [-0.19, -0.03], \beta = -0.15, p = .005$. The effect of confrontation condition was not significant ($p = .481$), nor were any other effects.

Postconfrontation stereotyping. The hierarchical regression analysis predicting postconfrontation stereotyping, now controlling for preconfrontation stereotyping, revealed a significant effect for the covariate, $t(240) = 4.63, B = 0.29, 95\% CI [0.17, 0.42], \beta = 0.27, p < .001$. The effect of bias type was significant, $t(240) = 7.36, B = 0.94, 95\% CI [0.69, 1.19], \beta = 0.43, p < .001$, with women ($M = 1.97, SD = 0.87$) being stereotyped more than Black men ($M = 0.73, SD = 0.91$). Also, as participants’ IMS increased, postconfrontation stereotyping significantly decreased, $t(240) = 2.41, B = -0.09, 95\% CI [-0.17, -0.02], \beta = -0.12, p = .017$.

Most importantly, the effect of confrontation was significant, $F(2, 238) = 20.03, \Delta R^2 = .09, p < .001$. Compared to the control condition, the internal and external framings resulted in significantly fewer stereotypic responses, $t(238) = 5.02, B = -0.51, 95\% CI [-0.86, -0.37], \beta = -0.27, p < .001$ and $t(238) = 5.88, B = -0.73, 95\% CI [-0.97, -0.48], \beta = -0.32, p < .001$, respectively. Confrontation conditions did not differ ($p = .348$; see Table 1). The effect of confrontation was not moderated by any interactions involving bias type or participants’ IMS/EMS ($ps > .119$). Thus, we again found that internal and external framings were equally powerful strategies for curbing stereotypic responding. This pattern held regardless of whether race or gender bias was confronted and the extent to which participants were internally or externally motivated to respond without prejudice.

Experiment 3

The results of Experiments 1 and 2 indicated that internal and external framings reduced stereotyping to the same extent, but they did not address whether motivational framing is critical to confrontation effectiveness. That is, perhaps people need not be concerned with framing their confrontations in terms of internal or external reasons for avoiding stereotyping, and simply confronting bias, even without motivational framing, is sufficient for its reduction. Experiment 3 investigated this issue by including a condition
in which participants were informed that their responses were stereotypic (including the explanation of what constitutes stereotypic responding with examples), and comparing it to the internal and external framing confrontations, as well as to a control condition.

Experiment 3 also addressed how participants perceived the internal and external framings. One could argue that similar reductions in stereotyping were observed in Experiments 1 and 2 because participants did not perceive any differences in the framings despite accurate recall of feedback content. Thus, all participants in Experiment 3 rated their feedback to the stereotype inference along two dimensions. First, because we are conceptualizing an internal motivational framing as appealing to people’s personal values in a supportive and constructive way (Deci & Ryan, 1985, 2000; Legault et al., 2007; Legault et al., 2011; Plant & Devine, 1998), we assessed the extent to which feedback was perceived as supportive and constructive. Second, because our conceptualization of an external motivational framing is that it demands conformity to norms and external standards (Legault et al., 2011; Plant & Devine, 1998), we assessed the extent to which feedback was perceived as controlling and demanding. We also assessed perceptions of the overall tone of the feedback (positive/negative), expecting that the internal framing would be perceived more positively than the external one (see Muraven, Gagné, & Rosman, 2008), so that we could control for tone when analyzing perceptions of the feedback.

Finally, Experiment 3 sought to replicate the gender-bias confrontation effects observed in Experiment 2, and to determine whether we would replicate the null findings suggesting that participants’ internal and external motivations did not interact with confrontation.

**Method**

**Participants**

Participants were recruited from MTurk in return for $0.75. Participation was restricted to people who had not completed other research from our laboratory that had similar materials or involved confrontation. Fourteen participants completed the photo-description task incorrectly, four participants disallowed their data on the postsession consent form, and one reported living in the United States for less than 5 years (and thus may have held stereotypes of women that differed from those of other participants). These data were distributed evenly across conditions and were deleted. This left 237 participants (65% women; 78.5% White, 4.0% Black, 8.0% Asian, 6.0% Hispanic, 3.2% “other,” 0.4% missing; $M_{age} = 39.03, SD = 12.30$). A sensitivity power analysis was performed with G*Power 3.1.9.2 (Faul et al., 2009), specifying a fixed model linear regression for $R^2$ increase, with $n = 237, \alpha = .05$, and 80% power. The analysis was performed for the highest order interaction (i.e., Confrontation x IMS x EMS), and it indicated that the minimum effect size that could be detected was $f^2 = .05$ (i.e., a small effect). Based on similar past research (Parker et al., 2018), Experiment 3 appeared to be sufficiently powered to detect interaction effects if present.

**Design**

A single-factor between-participants design was used (control condition, confrontation without motivation, internal framing, external framing), and IMS and EMS varied continuously.

**Procedure**

The procedure was identical to the gender-bias condition of Experiment 2, with two exceptions. First, control participants were presented with the same list of possible errors/biases as confronted participants but learned they had not demonstrated any biases. All confronted participants (i.e., confrontation without motivation, internal framing, and external framing) saw the same list with an “X” next to “Gender Bias,” followed by the explanation of their bias (as in Experiments 1 and 2). Thus, all confronted participants were informed of their stereotypic
responses and saw nonstereotypic alternatives that would have constituted nonstereotypic responses. Participants in the internal and external framing conditions then received the same motivational framings described previously.

Second, perceptions of feedback were assessed, with particular interest in whether the internal and external framings differed in line with the intended motivational focus. Participants were presented with a series of prompts beginning with, “My feedback was . . .” followed by 12 items. Embedded among fillers were four items theoretically consistent with an internal (supportive, encouraging, thoughtful, and constructive; \( M = 3.70, SD = 1.66, \alpha = .91 \)) and external framing (controlling, threatening, demanding, and bossy; \( M = 3.20, SD = 1.82, \alpha = .91 \)) rated on a 7-point scale (1 = strongly disagree, 7 = strongly agree). Next, all participants responded to a single item assessing tone of feedback (1 = negative, 7 = positive; \( M = 3.55, SD = 2.02 \)).

Participants then completed the postconfrontation stereotyping task, followed by the IMS (M = 7.10, SD = 1.90, \( \alpha = .87 \)) and EMS (M = 3.61, SD = 2.08, \( \alpha = .86 \)) scales (1 = strongly disagree, 9 = strongly agree).

Results

Preliminary Analyses

A one-way ANOVA revealed that condition did not predict IMS nor EMS (Fs < 1, ps > .75). IMS and EMS were mean centered for subsequent analyses.

Plan of Analyses

Unless otherwise noted, hierarchical linear regressions were used. Condition was represented with three dummy codes with the control condition serving as the reference group (West et al., 1996; control: DC1 = 0, DC2 = 0, DC3 = 0; internal framing: DC1 = 1, DC2 = 0, DC3 = 0; external framing: DC1 = 0, DC2 = 1, DC3 = 0; confrontation without motivation: DC1 = 0, DC2 = 0, DC3 = 1). Dummy coding was modified as needed in follow-up analyses to make additional comparisons. Dummy codes were always entered as a set, and the increment in \( R^2 \) was assessed for significance.

Analyses used the following steps: 1: IMS and EMS; 2: DC1, DC2, and DC3; 3: DC1 x IMS, DC2 x IMS, and DC3 x IMS; 4: DC1 x EMS, DC2 x EMS, and DC3 x EMS; 5: IMS x EMS; 6: DC1 x IMS x EMS, DC2 x IMS x EMS, and DC3 x IMS x EMS.

Perceptions of Confrontations

Overall tone. Confrontation condition significantly predicted perceived tone, \( \Delta F(3, 231) = 23.16, \Delta R^2 = .23, p < .001 \) (see Table 2). Participants perceived the external framing as more negative than the internal framing, \( \beta(231) = 2.85, B = −0.91, 95\%\ CI [−1.54, −0.28], \beta = −0.20, p = .005 \). The internal framing was rated as more negative than the confrontation without motivation, \( \beta(231) = 2.17, B = −0.70, 95\%\ CI [−1.34, −0.07], \beta = −0.15, p = .031 \), which was rated as more negative than the control condition, \( \beta(231) = 3.14, B = −1.04, 95\%\ CI [−1.70, −0.39], \beta = −0.23, p = .002 \).

Perceived internal focus. Analyses predicting perceived internal focus, now controlling for perceived tone, revealed an effect for tone, \( \beta(233) = 15.54, B = 0.58, 95\%\ CI [0.51, 0.65], \beta = 0.70, p < .001 \) and condition, \( \Delta F(3, 230) = 3.58, \Delta R^2 = .02, p = .015 \) (see Table 2). Critically, the internal framing was perceived as more internally focused than the external framing, \( \beta(230) = 2.65, B = −0.55, 95\%\ CI [−0.96, −0.14], \beta = −0.15, p = .009 \). The internal framing was also perceived as more internally focused than the control feedback, \( \beta(230) = 2.69, B = −0.60, 95\%\ CI [−1.05, −0.16], \beta = −0.15, p = .008 \). No other differences were significant (ps = .121).

Perceived external focus. Perceived tone predicted perceived external focus, \( \beta(233) = 10.22, B = −0.50, 95\%\ CI [−0.60, −0.41], \beta = −0.56, p < .001 \). Also, as EMS scores increased, participants perceived the feedback as more externally focused,
Most importantly, the effect of condition was significant, $\Delta F(3, 230) = 15.83$, $\Delta R^2 = .12$, $p < .001$. The external framing was perceived as more externally focused than the internal framing, $t(230) = 4.18$, $B = 1.06$, 95% CI [0.56, 1.56], $\beta = 0.26$, $p < .001$, and control condition, $t(230) = 5.40$, $B = 1.57$, 95% CI [0.99, 2.15], $\beta = 0.38$, $p < .001$.

Overall, these results demonstrated that participants perceived the framings differently, above and beyond the feedback’s perceived tone. Crucially, the internal framing was perceived as more supportive and constructive, and the external framing was perceived as more controlling and demanding.6

Stereotypic Responses

Preconfrontation stereotyping. On average, participants provided 2.78 ($SD = 0.67$) stereotypic responses across the preconfrontation prompts, and no participant provided none. As participants’ IMS increased, preconfrontation stereotyping decreased, $t(234) = 2.84$, $B = −0.07$, 95% CI [−0.11, −0.02], $\beta = −0.18$, $p = .005$. Preconfrontation stereotyping did not differ across conditions ($p = .523$), nor were any other effects significant.

Postconfrontation stereotyping. Stereotypic responding decreased as IMS increased, $t(233) = 3.64$, $B = −0.12$, 95% CI [−0.19, 0.06], $\beta = −0.23$, $p < .001$. Replicating our previous experiments, the effect of condition was significant, $\Delta F(3, 227) = 4.53$, $\Delta R^2 = .05$, $p = .004$. Unlike Experiment 2, however, this effect was qualified by an interaction between IMS and condition, $\Delta F(3, 227) = 3.21$, $\Delta R^2 = .04$, $p = .024$. This interaction (see Figure 2) was probed using Hayes’s (2013) PROCESS Model 1.

Participants relatively low on IMS (−1 $SD$) tended to provide fewer stereotypic responses in the three confrontation conditions relative to the control condition. The comparison between the control and external framing conditions reached significance, $t(227) = 2.36$, $b = −0.57$, 95% CI [−1.05, −0.09], $se = 0.24$, $p = .019$. However, the internal framing and control conditions did not differ, $t(227) = 1.17$, $b = −0.30$, 95% CI [−0.79, 0.20], $se = 0.25$, $p = .245$, and only a trend was found when comparing the control and confrontation without motivation conditions.
that some individuals are highly motivated to respond without bias because of their egalitarian personal standards and values (i.e., high internal motivation), and some individuals are highly motivated to respond without bias because of societal demands and pressures to be politically correct (i.e., high external motivation). In contrast, the present research investigated whether confrontations of bias that were framed in terms of internal or external reasons to avoid bias were differentially effective at curbing stereotypic responding.

Across three experiments, we considered the possibility that confrontations with an internal motivational framing would be more effective at reducing expressions of bias relative to an externally framed confrontation because external pressures may result in backlash and/or reactance (e.g., Legault et al., 2011). We also considered the possibility that internally and externally framed confrontations would reduce biased responses equally given prior research findings that even external sources of motivation reduce expressions of bias (e.g., Blanchard et al., 1994; Monteith et al., 1996). Experiment 1 found that, relative to a control condition, internally and externally framed confrontations were equally effective at reducing stereotyping of Black men, both when postconfrontation stereotyping was assessed immediately and after 2–3 days. Experiment 2 found that internal and external framings were equally effective at reducing expressions of race and gender bias, and confrontation effects were not qualified by participants’ internal and external motivations to respond without bias. Experiment 3 also found a significant effect of confrontation condition, but this effect was qualified by participants’ preexisting internal motivation to avoid prejudiced responses. Relative to a control condition, an internally framed confrontation reduced stereotyping among participants scoring relatively high on internal motivation to respond without bias. In contrast, the externally framed confrontation reduced stereotyping among participants scoring relatively low on internal motivation. Finally, the use of motivational framing in the confrontation appeared

General Discussion

Prior research investigating bias and motivation has primarily focused on enduring individual differences in internal and external motivations to respond without prejudice (Dunton & Fazio, 1997; Plant & Devine, 1998). This research demonstrates
necessary for reducing stereotyping, because a confrontation that pointed out and explained participants’ bias but did not include a motivational framing did not reduce stereotyping, relative to a control condition.

The weak support for the idea that matching participants’ chronic motivation for responding without prejudice with the confrontation’s motivational framing is at odds with the persuasion literature supporting the message–person matching hypothesis (e.g., Cesario et al., 2008; Hirsh et al., 2012). The only support we found for this idea was greater stereotyping reduction among highly internally motivated participants in the internal framing condition in Experiment 3. Perhaps if the aim of a confrontation is to change attitudes rather than to reduce stereotypic responding, the matching hypothesis would find greater support in the context of confrontation.

Our results can be contrasted with Legault et al.’s (2011) finding that participants reported more prejudiced attitudes following an externally framed egalitarian message relative to a control condition, whereas an internally framed message reduced prejudiced attitudes. We believe the conflicting motivational framing effects between our research and Legault et al.’s can be explained by the different nature of the research questions. When urging people to adopt egalitarian attitudes, it is unsurprising that people reacted against being told they must hold certain attitudes to fit in with social norms (see Brehm & Brehm, 1981). However, such reactance apparently is not produced when people are confronted about stereotypic responses that they actually just generated. In this case there is evidence of biased behavior, and given most people neither want to appear prejudiced to themselves nor to others (Pearson, Dovidio, & Gaertner, 2009), they do not react against an externally framed confrontation urging the reduction of the biased behavior. Thus, the present research is more consistent with prior studies showing that people can be prompted to regulate expressions of bias by external sources of motivation (Blanchard et al., 1994; Monteith et al., 1996).

Beyond reducing stereotyping, potential confronters may be concerned with interpersonal relations (Ashburn-Nardo et al., 2008), and motivational framing is likely to have implications for such outcomes. Experiment 3 demonstrated that participants perceived the external framing’s tone as more negative than the internal framing’s, and confrontations marked by a negative tone are likely to have negative interpersonal consequences (e.g., decreased liking; Czopp et al., 2006). Thus, confronters motivated to stop biased expressions, as was our focus, need not worry about confrontation tone. However, confronters worried about interpersonal relations may wish to confront by focusing on internal reasons to avoid prejudice.

Whereas some prior research has found that gender-bias confrontations are dismissed relative to race-bias confrontations (Czopp & Monteith, 2003; Gulker et al., 2013), the present research is consistent with the recent finding that evidence-based confrontations against gender bias can be effective. Specifically, Parker et al. (2018) found that participants confronted for biased evaluations of a female job applicant reported more guilt when the confrontation contained evidence of bias and underscored its discriminatory implications, relative to a confrontation simply pointing out bias. Participants’ guilt, in turn, predicted greater personal concern about responding in gender-biased ways in the future. The confrontations used in the present research similarly provided recipients with evidence of bias by targeting specific stereotypic responses and noting nonstereotypic alternatives. Thus, the present work extends upon prior research by demonstrating that evidence-based confrontations of gender bias are effective when they also highlight motivational reasons for curbing bias.

Limitations and Future Directions

An aspect of the feedback that deserves comment relates to participants receiving “slightly above average” feedback to the photo-description task. Given that people typically believe they are above average (e.g., Alicke et al., 1995), we intended
the feedback to be neutral. Nevertheless, we cannot be certain whether “slightly” may have been perceived negatively by at least some participants and induced a negative mood. Much past research has reduced bias responding without inducing a negative mood prior to confrontation (e.g., Czopp et al., 2006; Gulker et al., 2013; Mallett & Wagner, 2011). Nevertheless, future research is needed to systematically investigate whether pre-confrontation mood (e.g., based on task feedback) influences people’s responsiveness to confrontations and as a function of motivational framings.

Generalizability of the confrontation effects we observed is also important to consider. Specifically, to what extent does confrontation prompt the self-regulation of biases that are different from those involved in the initial confrontation? The present research purposefully used a postconfrontation task that mirrored the context of the original confrontation because such a situation often occurs in real life. That is, confrontation is often aimed at putting a stop to specific biased responses. However, future research is needed to address whether internal and external framings are differentially effective for curbing bias in more generalized ways (e.g., with novel responses, and when made publicly vs. privately).

Despite the similarities between the pre- and postconfrontation stereotyping tasks, we suspect that demand characteristics were not operating. First, participants responded to novel postconfrontation items, and they were informed that responses would remain anonymous and there would be no feedback. Second, Experiment 1 incorporated a delay and a change of context for assessing postconfrontation responding. Finally, the confrontation without a motivational framing failed to reduce stereotyping, which seems an unlikely outcome if demand characteristics were operating.

Finally, participants in the present research were confronted with computer-generated feedback that was part of an experimental procedure. Did confrontation cause regulation in part because of the source and its presumed characteristics (e.g., status or power)? To date, confrontation research has not examined reactions as a function of source role (e.g., peer, subordinate, supervisor), and this is a pressing issue for future research. Nonetheless, the present work provides valuable insight for online confrontations, and our confrontation procedure could easily be implemented as training tools in workplace or educational settings. Importantly, we included both nonstudent and student samples, with similar findings across them.

**Conclusion**

People often do not confront bias in part because they are unsure how to confront (Ashburn-Nardo et al., 2008). The present research indicated that whether confrontations emphasize internal or external reasons to avoid stereotyping, using a motivational framing is critical for effectively prompting people to reduce their biased responses.

**Acknowledgements**

We would like to thank Deeba Aminuddin, Haley Brower, Nina Shah, and Lauren Zeithammer for their help with conducting this research.

**Funding**

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

**Notes**

1. In all experiments, gender was initially included in analyses but later excluded as it was not associated with significant effects, except one in Experiment 2: Women scored higher on the IMS ($M = 5.77$, $SD = 0.33$) than men ($M = 5.07$, $SD = 1.53$), $F(1, 233) = 12.73, p < .001, \eta^2_p = .05$. Furthermore, the pattern and significance of the findings across all experiments do not change when gender is entered as a covariate.

2. In all experiments, participants also completed 18 affect items prior to the postconfrontation stereotype measure (see supplemental materials for details).

3. Across all three experiments, exclusion of the preconfrontation stereotyping covariate did not
alter results notably. Given our theoretical interest in whether the confrontation resulted in less biased responding, over and above participants’ typical level of biased responding, we retained the covariate.

4. Beyond the reported sample sizes, 22 Black participants in the racial bias condition were excluded as this sample was insufficient for testing whether target group status affected confrontation outcomes.


6. The interaction between IMS and EMS was also significant for external focus but is not discussed because it is not of theoretical interest.

References


