

The Bogus Pipeline and Motivations to Respond Without Prejudice: Revisiting the Fading and Faking of Racial Prejudice

E. Ashby Plant

Florida State University

Patricia G. Devine and

Paige C. Brazy

University of Wisconsin—Madison

The current work draws upon recent developments in attitude and prejudice theory as well as the assessment of people's motivation to respond without prejudice in order to reexamine the classic bogus pipeline technique. Specifically, the present study examines the joint impact of individual differences in the source of motivation to respond without prejudice and the social context in which racial beliefs are reported (i.e. private, public, pipeline) on stereotype endorsement scores. Results revealed that stereotype endorsement scores in the various reporting conditions were moderated by participants' source of motivation to respond without prejudice. The findings are consistent with the argument that responses under bogus pipeline conditions assess people's awareness of their experience of bias, whereas private reports assess people's personal beliefs about the appropriateness of the stereotype of Black people.

keywords bogus pipeline, motivation, prejudice, racial beliefs, stereotype

AGAINST a backdrop of legislative and normative changes discouraging the expression of prejudice toward Blacks in the United States, discerning the motivation underlying non-prejudiced responses has proven difficult for social perceivers and social scientists. Indeed, the existence of compelling normative (e.g. external) reasons to refrain from expressing prejudice has cast doubt on the veracity of self-reported nonprejudiced attitudes (e.g. Crosby,

Bromley, & Saxe, 1980; Dovidio & Fazio, 1992; Sigall & Page, 1971). Specifically, the concern was that people would be strategic, rather than

Author's note

Address correspondence to E. Ashby Plant, Department of Psychology, Florida State University, Tallahassee, FL 32306-1270, USA [email: plant@psy.fsu.edu]

truthful, in reporting their racial attitudes so as to convey a nonprejudiced image to others. As a result, attempts were made to develop measures of racial beliefs that would bypass external motivation to respond without prejudice and, hence, the likelihood of eliciting overly positive self-presentations (e.g. Fazio, Jackson, Dunton, & Williams, 1995; Greenwald, McGhee, & Schwartz, 1998; McConahay, 1986; Sigall & Page, 1971).

One such measure is the now famous bogus pipeline technique (Jones & Sigall, 1971; Sigall & Page, 1971; see Roese & Jamieson, 1993, for a review). The goal of the bogus pipeline technique is to circumvent strategic self-presentations by convincing participants that it would be impossible, and possibly embarrassing, to conceal their true beliefs. Sigall and Page (1971) examined the extent to which participants indicated that various traits were characteristic of Blacks under a bogus pipeline and comparison reporting condition. Their results indicated that White participants were more likely to attribute negative traits to Blacks under the pipeline than the comparison condition. They concluded that these findings supported the effectiveness of the bogus pipeline to impel people's disclosure of undesirable information about themselves.

Although extensive effort has gone into bypassing normative pressures in the assessment of racial attitudes, until recently, little systematic effort explored the extent to which people are sensitive to external forces in the reporting of their racial attitudes. The prevailing assumption was that most, if not all, Whites would respond in line with normative expectations by reporting untruthful low-prejudice beliefs. Recent work, however, suggests that there is variability in people's sensitivity to external forces proscribing prejudice and that assessing racial attitudes is more complex than may have been appreciated in the past. For example, Plant and Devine (1998) argued that people vary in the extent to which they are motivated to respond without prejudice for internal (personal) and external (social) reasons. They demonstrated that these sources of motivation are not only conceptually, but also empirically,

distinct. They further argued that it may be imprudent to discount internal motivation simply because external motivation exists in many situations. Instead, they maintained that both sources of motivation are likely to affect people's prejudice-related reactions, though to varying degrees for different people.

By taking into account the source of people's motivation to respond without prejudice, we may be in a position to move beyond simply bypassing concerns about normative expectations in the assessment of racial attitudes and instead focus on the meaning underlying people's reported racial beliefs in different contexts. The current work examines how a consideration of the reasons why people are motivated to respond without prejudice may provide insights into what is revealed by classic strategies designed to measure people's 'true' racial beliefs. Specifically, we revisit the bogus pipeline technique (Jones & Sigall, 1971; Sigall & Page, 1971) because it has been widely accepted as an effective strategy for revealing bias that people otherwise would choose to hide. Ultimately, our goal is to illustrate how synthesizing new theory with the classic bogus pipeline methodology may help to reveal the meaning underlying responses to this assessment tool.

Individual differences in motivation to respond without prejudice

In an effort to capture the reasons (both internal and external) underlying people's efforts to respond without prejudice, Plant and Devine (1998) developed and validated the Internal and External Motivation to Respond without Prejudice toward Black people Scales (IMS and EMS respectively). Internal motivation arises from personally important beliefs whereas external motivation derives from a desire to avoid negative reactions from presumably nonprejudiced others. Sample items from the IMS include 'I attempt to act in nonprejudiced ways because it is personally important to me' and 'Being nonprejudiced toward Black people is important to my self-concept'. Sample items from the EMS include 'I attempt to appear nonprejudiced

toward Black people in order to avoid disapproval from others' and 'I try to act non-prejudiced toward Blacks because of pressure from others'. Plant and Devine demonstrated that the scales were reliable and provided compelling evidence regarding their convergent, discriminant, and predictive validity. Perhaps most importantly for establishing the discriminant validity, both the IMS and EMS are unrelated to general social desirability measures.¹ They also found that the IMS and EMS were largely independent (average r in 4 samples = $-.14$). Thus, individuals can be motivated to respond without prejudice primarily for internal reasons, primarily for external reasons, for both internal and external reasons, or for neither type of reason.

In establishing the scales' predictive validity, Plant and Devine (1998) showed that non-Black² participants' self-reported level of stereotype endorsement varied as a function of both the source of their motivation to respond without prejudice and whether they made their reports privately (i.e. anonymously) or publicly (i.e. directly to the experimenter). Specifically, participants motivated to respond without prejudice primarily for external reasons (i.e. low IMS, high EMS) reported lower stereotype endorsement scores in public than in private. Plant and Devine argued that for these participants, making their responses publicly cued the potential to be evaluated by the external audience prompting them to strategically alter their responses to avoid social sanction. Those who were internally motivated, regardless of their level of EMS, responded with low stereotype endorsement scores in both reporting conditions and those scoring low on both the IMS and EMS had fairly high stereotype endorsement scores in both reporting conditions.

These findings strongly suggest that the public/private difference among the low IMS, high EMS participants reflects a strategic effort to conceal their prejudiced beliefs from others. However, it is less immediately clear how best to interpret the low-prejudice beliefs reported by the highly internally motivated participants. It is possible, for example, that these individuals' responses also reflect mere compliance with

normative expectations, but their compliance has generalized to include private as well as public reports of their beliefs. Casting doubt on this possibility is evidence showing that when high IMS people violate their prejudice-relevant personal standards, they report elevated levels of guilt and self-criticism (Plant & Devine, 1998). This type of affective response is theoretically consistent with self-punishment resulting from violating an internalized moral standard (Higgins, 1987). In contrast, low IMS participants did not show elevated guilt and self-criticism following violations of a similar magnitude.

Interestingly, Plant and Devine (1998) also found that the magnitude of violations from personal standards differed as a function of respondents' external motivation to respond without prejudice. Most important for the present set of issues, among high IMS participants, those who also scored high on the EMS reported being more likely to violate their personal standards (i.e. respond with prejudice) compared with their high IMS, low EMS counterparts. In addition, Plant and Devine (2000) found that high IMS, high EMS White people reported that they were more likely to have stereotypes come to mind during interactions with Black people than high IMS, low EMS White people. Confirming these expectations, Devine and colleagues (Devine, Plant, Amodio, Harmon-Jones, & Vance, 2002) demonstrated across several studies that high IMS, high EMS people actually respond with higher levels of implicit (theoretically uncontrollable) bias than high IMS, low EMS people.

The fact that some non-Black people who embrace nonprejudiced standards continue to experience negative responses to Blacks is consistent with recent theoretical approaches, which argue that people's responses to out-group members contain multiple and sometimes conflicting components (e.g. Devine, 1989; Gaertner & Dovidio, 1986; Greenwald & Banaji, 1995; Wilson, Lindsey, & Schooler, 2000). For example, Devine's dissociation model (1989) argues that it is important to distinguish between stereotypes and people's personal beliefs. According to the model,

stereotypes are defined as knowledge of the attributes that are stereotypically associated with a particular group, whereas personal beliefs are defined as the level of endorsement of the content of the cultural stereotype. Because stereotypes typically have a long socialization history and have been frequently activated, they can be automatically activated, providing a 'default' response to outgroup members. In contrast, because personal beliefs are typically developed later in time (see Allport, 1954; Proshansky, 1966), they tend to be less accessible than stereotypes. As a result, even those who personally embrace low-prejudice beliefs may sometimes have a stereotype-based response to outgroup members (e.g. Banaji & Greenwald, 1995; Devine, 1989). Devine and colleagues (Devine & Monteith, 1993; Devine, Monteith, Zuwerink, & Elliot, 1991) argued that those with internalized nonprejudiced beliefs must learn to inhibit their automatic prejudiced responses.

Plant and Devine's (1998, 2000; Devine et al., 2002) findings concerning non-Black people's awareness of their experience of bias and actual level of implicit bias toward Black people suggest that high IMS, low EMS non-Blacks are further along in developing the ability to respond without prejudice than their high IMS, high EMS counterparts. As a result, high IMS, low EMS non-Blacks are likely to respond to Black people in a manner consistent with their personal beliefs as opposed to the stereotype. In contrast, high IMS, high EMS non-Blacks report that they still have the stereotype automatically activated when reacting to Black people and, thus their responses toward Black people are likely to be influenced by the stereotype. This line of reasoning is consistent with theories of internalization, which argue that the more internalized a value, the less it is experienced as compelled by external forces and the more effective people are at responding consistently with the value (e.g. Ryan & Connell, 1989; Ryan, Sheldon, Kasser, & Deci, 1996; see Devine et al., 2002, for a detailed description of this theoretical argument).

Together these findings suggest that by knowing the source of people's motivation to

respond without prejudice we are better able to predict when and why people are likely to respond with or without intended and unintended bias. These new insights suggest that it may be prudent to reexamine the bogus pipeline technique and the assumptions concerning the meaning of racial attitudes supplied under bogus pipeline conditions. In the wake of Sigall and Page's (1971; Jones & Sigall, 1971) bogus pipeline research, with a few notable exceptions (Brigham, Bloom, Gunn, & Torok, 1974; Ostrom, 1973), the bogus pipeline was widely accepted as a useful (if cumbersome) way to reduce normative concerns presumed to contaminate traditional paper-and-pencil measures of racial attitudes (see Roese & Jamieson, 1993).³

Sigall and Page's (1971) findings are typically interpreted as evidence that responses when attached to the bogus pipeline reflect people's true attitudes and that people's self-reported beliefs are contaminated by concerns regarding normative standards.⁴ However, rather than concluding that one type of response is the *true* response, we believe that both types of responses, if collected properly, may be valid and reflect different components of non-Black people's responses to Blacks (Devine, 1989; Wilson et al., 2000). We argue that whereas self-reported beliefs reflect people's intentions and desired responses (i.e. personal beliefs), their responses when attached to the bogus pipeline reflect their perceptions regarding the likelihood of the stereotype becoming activated (Devine, 1989; Plant & Devine, 2000). Jones and Sigall (1971), raising a similar possibility, proposed that it was possible that questionnaire responses may reflect 'the way a person wants to see himself and the way in which he tries to behave' (p. 363), whereas the bogus pipeline may assess 'reservations about how his autonomic nervous system feels about Negroes' (p. 363).

Consider the experience of people attached to the bogus pipeline machine. They are led to believe that the machine will assess their physiological responses and determine whether they have stereotypic reactions. To the extent that people believe they ever have stereotypic

responses, it is only sensible for them to predict that the machine will indicate some degree of biased responding (i.e. in this case indicated by stereotype endorsement). In contrast, when simply asked to report their beliefs, people's responses are likely to reflect their personal endorsement of the stereotype. According to this logic, people who believe they sometimes respond in an unintended stereotypic manner (high IMS, high EMS) are likely to report higher stereotypic endorsement scores when attached to the bogus pipeline than when in self-report conditions.

The present study

We believe that Plant and Devine's (1998) IMS and EMS may prove useful in reexamining attitude assessment in the context of the bogus pipeline procedure. To this end, we examined participants' stereotype endorsement scores across three reporting conditions, public, private, and bogus pipeline, as a function of participants' source of their motivation to respond without prejudice. This design allowed us to replicate Sigall and Page's (1971) procedure and enabled comparisons between private and bogus pipeline responses as well as a replication of Plant and Devine's (1998) findings comparing private and public responses. We anticipated that the degree to which participants showed consistency or dissociation across the reporting conditions would vary as a function of their source of motivation to respond without prejudice. First, considering responses in the public and private conditions, we expected to replicate Plant and Devine. Specifically, participants who were primarily externally motivated to respond without prejudice (low IMS, high EMS) were expected to respond with lower stereotype endorsement scores in public than in private. All other participants were expected to respond with similar stereotype endorsement scores in public and private.

In addition, as reviewed previously, we propose that participants' stereotype endorsement scores in the bogus pipeline condition reflect expectations regarding the activation of

the stereotype of Blacks. In contrast, we argue that stereotype endorsement scores in private reflect personal beliefs about the appropriateness of the stereotype of Blacks. In line with this reasoning, we expected that high IMS, high EMS participants would report low stereotype endorsement scores in private, but higher stereotype endorsement scores in the bogus pipeline condition (i.e. reflecting their awareness that they sometimes respond with unintended bias; Devine et al., 2002; Plant & Devine, 1998, 2000). All other participants were expected to respond with similar stereotype endorsement scores in the bogus pipeline and private condition because their activation of the stereotype tends to be consistent with their personal beliefs (Devine et al., 2002).

In summary, participants were expected to show consistency in their stereotype endorsement scores across reporting conditions with two notable exceptions. First, primarily externally motivated people were expected to respond with lower stereotype endorsement scores in public than in private (and likely pipeline). Second, because participants high in both internal and external motivation to respond without prejudice are aware that they sometimes respond with unintended bias, they were expected to respond with higher stereotype endorsement scores in the bogus pipeline condition than in the private reporting condition.

In the current study, in addition to assessing stereotype endorsement scores, an affect measure was included to explore participants' affective reactions to their stereotype endorsement scores. To the degree that high IMS, high EMS participants' stereotype endorsement scores in the bogus pipeline condition reflect a personal failure (i.e. their awareness of unintended stereotype activation), it should be accompanied by feelings of guilt and self-criticism (e.g. Higgins, 1987; Plant & Devine, 1998). As a result, we expected high IMS, high EMS participants' stereotype endorsement scores in the bogus pipeline condition to be associated with guilt and self-criticism.

The predicted responses would be quite different based on the typical interpretation of

the bogus pipeline. According to this interpretation, because people generally fall prey to concerns about normative expectations when reporting their racial beliefs, all participants should respond with higher stereotype endorsement scores in the bogus pipeline condition than in either the public or private conditions. Alternatively, one could imagine that only those who report being sensitive to normative expectations (i.e. the externally motivated) would respond with higher stereotype endorsement scores under bogus pipeline than public or private reporting conditions. Such a finding would be consistent with the traditional interpretation of the bogus pipeline. Further, if low stereotype endorsement scores in public and private merely reflect ‘faking’ in order to comply with the nonprejudiced normative pressure, then responding with bias in the bogus pipeline condition should not result in guilt and self-criticism for any of the participants.

Finally, in the current study, a traditional attitude measure of prejudice (Attitude toward Blacks Scale or ATB; Brigham, 1993) was obtained. By assessing prejudice with the ATB, we could determine whether responses across the three reporting conditions varied as a function of people’s attitudes toward Blacks and whether the pattern of stereotype endorsement scores could be more parsimoniously accounted for by people’s attitudes toward Black people.

Method

Participants and design

Respondents were 189 introductory psychology students (59% female, 94% White) who participated individually in exchange for course extra-credit. The IMS ($\alpha = .83$), EMS ($\alpha = .81$), and ATB ($\alpha = .91$) were completed as a part of a mass testing session early in the semester. Participants indicated their agreement with the IMS and EMS items on 9-point scales ranging from 1 (*strongly disagree*) to 9 (*strongly agree*). Participants’ scores on the IMS and EMS items were averaged such that they ranged from 1 to 9 with higher scores reflecting higher levels of the relevant motivation. Participants indicated

their agreement with the ATB items on a 7-point scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). Responses to the ATB items were averaged such that scores ranged from 1 to 7 with higher scores indicating more positive attitudes toward Blacks. The correlations between these three measures are presented in Table 1.

Participants were considered eligible for the study if their responses fell into the top or bottom 30% of the IMS (high IMS, $M = 8.92$; low IMS, $M = 5.92$) and EMS (high EMS, $M = 6.49$; low EMS, $M = 2.17$) distributions.⁵ Participants who met the selection criteria were contacted and invited to participate (IMS, $M = 7.45$, $SD = 1.69$, skewness = $-.84$; EMS, $M = 4.23$, $SD = 2.34$, skewness = $.17$). Participants from each of the four IMS/EMS groups were randomly assigned to the pipeline, public, or private reporting condition. Thus, the experimental design was a 2 (IMS: low vs. high) \times 2 (EMS: low vs. high) \times 3 (Reporting Condition: pipeline vs. public vs. private) between-subjects factorial. Nine participants in the pipeline condition reported suspicion about whether the bogus pipeline machine could measure their true attitudes. These nine participants were roughly distributed across the four IMS/EMS groups and their data were dropped from the analyses.

Materials and procedure

Overview Each participant was met in the waiting room by an experimenter who was blind to the participant’s IMS, EMS, and ATB scores. The experimenter explained that the participant would be completing questionnaires on various topics. After reading and signing the consent form, following Sigall and

Table 1. Correlations between the IMS and EMS and ATB

| | IMS | EMS | ATB |
|-----|-----|------|--------|
| IMS | - | -.05 | .69** |
| EMS | | - | -.35** |
| ATB | | | - |

** $p < .001$.

Page's procedure, the participant was given a General Attitude Questionnaire. This six-item measure asked about opinions regarding music, cars, movies, and sports and responses were given on a scale that ranged from -3 (*strongly disagree*) to +3 (*strongly agree*). All participants completed the questionnaire but its purpose was to facilitate the validation of the pipeline apparatus. After completing the questionnaire, the participant placed it in a drop box and was shown to the experimental room. The experimenter explained that the next questionnaire assessed attitudes about social groups. The participant was asked to select a piece of paper from a bowl on which was written the social group he or she would be evaluating (all slips African Americans). The participant then completed the stereotype endorsement measure in one of the three reporting conditions. Finally, the participant was debriefed and thanked.

Pipeline reporting condition The pipeline condition was modeled after Sigall and Page's (1971) bogus pipeline procedure. However, to update the procedure, we substituted a joystick for Sigall and Page's steering wheel. We reasoned that our participants would have some familiarity with joystick - computer connections and that this would enhance the believability of our cover story. Participants sat in front of the joystick, which was connected to a computer console. The computer console consisted of an IBM computer, two Oscilloscopes, a printer, a microphone noise operated relay, an electrode jack, and a dial readout. The computer and oscilloscopes were turned on during the session, but neither produced actual readings. Behind the joystick, a graphic printout of the -3 (*uncharacteristic*) to +3 (*characteristic*) rating scale was presented on which participants indicated their level of endorsement for each trait. The experimenter, donning a white lab coat and latex gloves, applied electrode gel and attached electrodes to the wrist of the participant's dominant hand. Following Sigall and Page, the experimenter explained that the machine was an adapted electromyograph or EMG that could determine people's attitudes

with a high level of accuracy. The experimenter explained how the machine measured muscle reaction tendencies and, thereby, assessed undistorted first response to each of the questionnaire items.

Next, the participant was provided with 'evidence' that the machine could accurately read his or her attitudes. During this validation phase, the experimenter asked the participant, while holding the joystick, to report his or her true attitudes and false attitudes toward the issues from the General Attitude Questionnaire they had completed previously. Because the experimenter surreptitiously had retrieved the participant's responses to the questionnaire, feedback was provided indicating that the EMG could accurately read the participant's attitude. Having firmly established the machine's truth detecting capabilities, the participant was asked to predict the machine's response to the items on the stereotype endorsement measure. The experimenter read each trait aloud, and after examining the EMG and allegedly recording its response, had the participant supply his or her prediction of the machine's response and recorded this response.

Public reporting condition The public reporting condition was modeled after Sigall and Page's (1971) comparison condition. In this condition, the pipeline apparatus was present but was explained as a computer for data analysis. Thus, rather than predicting the machine's response for each trait in the stereotype endorsement measure, the participant simply provided his or her response to each trait aloud to the experimenter and the experimenter recorded each response.

Private reporting condition The private reporting condition was modeled after traditional paper-and-pencil measures, which typically encourage a sense of anonymity. In this condition, the participant completed the stereotype endorsement measure in private and was assured that his or her response would be completely confidential. When done, the participant placed the questionnaire in a drop box filled with questionnaires.

Stereotype endorsement measure Following Sigall and Page (1971), participants rated the extent to which each of 35 traits were characteristic of Blacks on a scale from -3 (*uncharacteristic*) to +3 (*characteristic*). Although many of the traits used by Sigall and Page (1971) were included on the measure, some of the traits they used are no longer part of the stereotype of Blacks (e.g. happy-go-lucky; see Devine & Elliot, 1995) and were deleted from the list. In addition, other traits, which recent research has suggested are core components of the cultural stereotype of Blacks were added to the list (e.g. athletic, criminal, see Devine & Elliot, 1995; Wittenbrink, Judd & Park, 1997). The list also included traits that were nonstereotypic of Blacks (e.g. sensitive, kind). The stereotype index was created by summing responses to the 14 stereotype-related traits: athletic, aggressive, rude, criminal, dangerous, violent, low in intelligence, unreliable, careless, free-loaders, ignorant, streetwise, rhythmic, lazy (Cronbach's $\alpha = .86$).⁶ Scores on the stereotype index could range from -42 to +42. Scores around zero indicate that the trait was viewed as neither characteristic nor uncharacteristic of Blacks.

Affect measure In order to assess participants' affective reaction to their stereotype endorsement, participants rated the extent to which each of 33 affect items (e.g. guilty, fearful, happy, sad, angry at others) applied to how they were feeling about the task they just completed on a scale that ranged from 1 (*does not apply at all*) to 7 (*applies very much*). Affect indices were created by averaging related items based on previous work and were confirmed by factor analysis (e.g. Devine et al., 1991; Plant & Devine, 1998). The most theoretically relevant index reflected guilt and self-criticism (i.e. negself). The negself index included angry at myself, guilty, annoyed at myself, disappointed with myself, disgusted with myself, regretful, and self-critical ($\alpha = .94$).⁷

Results

Stereotype endorsement score analyses

Our first step was to examine whether we replicated Sigall and Page's (1971) findings. To this

end, we compared stereotype endorsement scores in the public and pipeline reporting conditions using an independent samples *t* test. Although, in general, the stereotype endorsement scores were quite low in the current sample, consistent with Sigall and Page, stereotype endorsement scores were higher in the bogus pipeline condition ($M = 4.51$, $SD = 9.25$) than in the public condition ($M = 0.72$, $SD = 8.72$) ($t(119) = -2.32$, $p < .03$). Next, we compared stereotype endorsement scores in the private to those in the public and pipeline conditions. These comparisons revealed that scores in the private condition were intermediate ($M = 3.07$, $SD = 10.31$) between pipeline and public scores and did not significantly differ from either ($t(118) = -.81$, $p = .42$ and $t(117) = 1.34$, $p = .18$, respectively).

If we had not included individual difference measures, we could conclude at this point that our findings generally replicated the pattern of findings reported by Sigall and Page (1971) but that unlike public reports, private stereotype endorsement scores do not differ from those obtained under bogus pipeline conditions. However, we anticipated that stereotype endorsement scores would vary as a function of the source of people's motivation to respond without prejudice as well as the reporting condition in which responses were made. Thus, stereotype endorsement scores were submitted to an IMS (low vs. high) \times EMS (low vs. high) \times Reporting Condition (pipeline vs. public vs. private) between-subjects analysis of variance (ANOVA). In the case of significant interactions, simple main effects were explored and then cell differences were compared using planned comparisons based on a priori expectations.

The analysis on the stereotype endorsement index revealed a main effect of IMS, such that high IMS participants reported lower stereotype endorsement scores ($M = 0.46$, $SD = 9.76$) than low IMS participants ($M = 5.09$, $SD = 8.73$) ($F(1, 168) = 11.75$, $p < .005$). The analysis also revealed a significant main effect of EMS ($F(1, 168) = 4.47$, $p < .04$). Overall, high EMS participants reported higher stereotype endorsement scores ($M = 4.27$, $SD = 9.55$) than their low EMS

counterparts ($M = 1.40, SD = 9.34$). In addition to the individual difference main effects, as described in the previous set of analyses, there was a marginal main effect of Reporting Condition ($F(2, 168) = 2.71, p = .07$). However, these main effects were qualified by an $IMS \times EMS \times$ Reporting Condition interaction ($F(2, 168) = 3.19, p < .05$). The means are presented in Figure 1.

To explore the nature of this three-way interaction, we examined the $IMS \times$ Reporting Condition interaction separately for low and high EMS participants. The analysis on low EMS participants only revealed a significant IMS main effect such that high IMS participants reported lower stereotype endorsement scores ($M = -.92, SD = 9.67$) than low IMS participants ($M = 3.72, SD = 8.47$) ($F(1, 88) = 5.88, p < .02$). Consistent with expectations, neither the Reporting Condition main effect nor the $IMS \times$ Reporting Condition interaction was significant, suggesting that condition did not affect the responses for low EMS participants at either level of IMS.

Among high EMS participants, there was also a significant main effect of IMS such that high IMS participants reported lower stereotype endorsement scores ($M = 1.95, SD = 9.75$) than low IMS participants ($M = 6.58, SD = 8.87$) ($F(1, 80) = 5.91, p < .02$). This main effect, however, was qualified by an $IMS \times$ Reporting Condition interaction ($F(2, 80) = 5.17, p < .01$). Planned comparisons between the reporting conditions were conducted for the low IMS, high EMS and the high IMS, high EMS participants separately. Among the high EMS participants who were low in IMS, the pattern of results replicates and extends Plant and Devine's (1998) findings. That is, these participants reported high stereotype endorsement scores in private ($M = 10.21, SD = 10.39$), but very low scores in public ($M = 1.79, SD = 7.16$) ($t(26) = 2.50, p < .02$). Also noteworthy, the stereotype endorsement scores for the low IMS, high EMS participants in the pipeline condition ($M = 7.67, SD = 7.12$) were significantly higher than the public condition ($t(27) = -2.22, p < .04$), but did not differ from the private reporting condition ($t(24) = .76$,

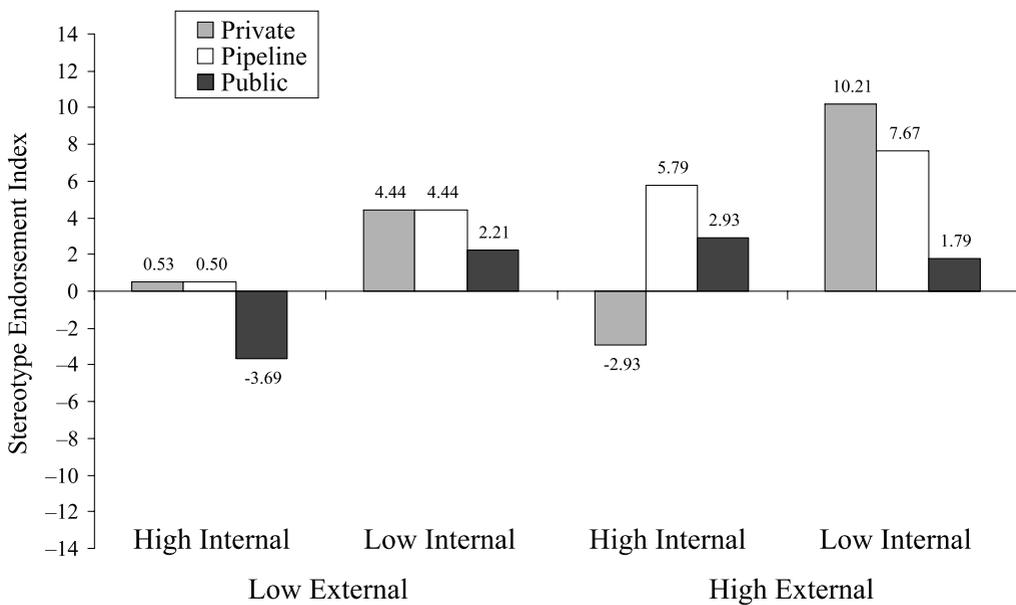


Figure 1. Stereotype endorsement as a function of IMS, EMS, and Reporting Condition.

$p = .45$). Among the high EMS participants who were also high in IMS, a different pattern emerged. Replicating Plant and Devine (1998), for these participants there was no difference in stereotype endorsement scores across the private ($M = -2.93$, $SD = 10.42$) and public ($M = 2.93$, $SD = 10.96$) reporting conditions ($t(27) = -1.47$, $p = .15$). In addition, consistent with expectations, the stereotype endorsement scores in pipeline condition ($M = 5.79$, $SD = 5.16$) were significantly higher than in the private condition ($M = -2.93$, $SD = 10.42$) ($t(26) = 2.19$, $p < .01$). However, the stereotype endorsement scores in the pipeline condition did not differ from those in the public condition ($t(27) = .89$, $p = .38$). Before commenting on the meaning and implications of these data for key assumptions in the bogus pipeline literature, we first present the ATB and affect analyses.

Analyses involving ATB In order to examine whether these findings could be accounted for by participants' attitudes toward Blacks, we conducted two analyses involving participants' ATB scores. First, we examined the extent to which attitudes toward Blacks moderated the effects of reporting condition on stereotype endorsement scores. Second, we repeated the analysis involving IMS, EMS, and reporting condition with ATB as a covariate.

Eight participants had not completed the ATB in the mass survey. These eight participants came from six different cells of the of the full IMS \times EMS \times Reporting Condition design. Participants' stereotype endorsement scores were submitted to an ATB (high-prejudice vs. low-prejudice) \times Reporting Condition (private vs. public vs. pipeline) between-subjects factorial ANOVA. This analysis revealed only a significant main effect of ATB scores, such that high-prejudice participants reported higher stereotype endorsement scores ($M = 5.63$, $SD = 8.04$) than did low-prejudice participants ($M = -0.12$, $SD = 10.21$) ($F(1, 166) = 15.78$, $p < .001$). The ATB \times Reporting Condition interaction was not significant ($F < 1.00$).

We next conducted an IMS (low vs. high) \times EMS (low vs. high) \times Reporting Condition (pipeline vs. public vs. private) between-subjects

analysis of covariance with ATB as the covariate. The analysis indicated that ATB was a significant covariate ($F(1, 159) = 15.55$, $p < .001$). However, even when controlling for participants' attitudes toward Blacks, the IMS \times EMS \times Reporting Condition interaction remained significant ($F(1, 159) = 3.38$, $p < .04$). In addition, the nature of this interaction was identical in form to the findings reported above without the covariate. High IMS, high EMS participants reported higher stereotype endorsement scores in the pipeline than in the private condition ($F(1, 25) = 6.02$, $p < .03$). In addition, the low IMS, high EMS participants reported lower stereotype endorsement scores in the public reporting condition than in the pipeline and private conditions ($F(1, 40) = 5.63$, $p < .03$). As before, none of the other comparisons between conditions for any of the participants significantly differed. These findings indicate that the findings based on participants' source of motivation to respond without prejudice could not be accounted for by their attitude toward Blacks.

Affect analyses

In our analysis of the affective responses, we focused on the prediction that high IMS, high EMS participants' stereotype endorsement scores in the bogus pipeline condition would be associated with feelings of guilt and self-criticism. Therefore, we examined the correlation between guilt-related (i.e. *negself*) affect and stereotype endorsement scores for high IMS, high EMS participants in the bogus pipeline condition. The analysis revealed a moderate correlation between *negself* and stereotype endorsement ($r(12) = .50$, $p = .07$), such that higher stereotype endorsement was associated with more guilty and self-critical affect. It is worth noting, that the correlation between stereotype endorsement and *negself* was small and in some cases negative for the other groups of participants in the bogus pipeline condition (all r s $.15$ or below).⁸ In addition, high IMS, high EMS participants' stereotype endorsement scores in the public and private reporting conditions were not significantly associated with *negself* ($r(14) = .41$, $p = .13$ and $r(13) = .16$, $p = .58$, respectively).

Discussion

We began this article by noting that the changing social climate, which is increasingly intolerant of expressions of prejudice, has aroused extensive concern in the field of prejudice regarding the potential contamination of self-reported attitudes by normative concerns. The bogus pipeline (Jones & Sigall, 1971; Sigall & Page, 1971) was offered as a means of circumventing such concerns and the initial evidence provided by Sigall and Page was taken as an indication that it effectively did so. However, we argued that theoretical developments in the understanding of the complexity of people's attitudes suggest that the assessment of racial attitudes may be more complicated than was previously realized. More specifically, we proposed that differences between non-Black people's responses across bogus pipeline and private self-report conditions may reflect the fact that these techniques tap into different components of people's attitudes toward Blacks. We argued that private self-reports assess people's personal beliefs and desired responses, whereas responses in the bogus pipeline condition may tap into people's expectations regarding the likelihood of the stereotype becoming activated. In order to explore these issues, we examined stereotype endorsement scores across reporting conditions (i.e. public, private, and bogus pipeline) for people with varying sources of motivation to respond without prejudice.

Participants' responses across the reporting conditions were wholly consistent with our predictions and conceptualization of the bogus pipeline. Specifically, replicating Plant and Devine's (1998) findings, low IMS, high EMS participants reported lower stereotype endorsement scores in public than in private, suggesting that reporting their beliefs publicly aroused concerns about being evaluated by others. The other groups of participants' responses were similar in the public and private conditions. Together these findings indicate that public reporting conditions may introduce normative concerns that are not present in private reporting conditions, but such concerns are only likely

to influence people primarily motivated to respond without prejudice in order to avoid negative reactions from others.

Further, although our comparison of the responses in the public condition to responses in the bogus pipeline condition generally replicated Sigall and Page's (1971) findings, the comparison of stereotype endorsement scores in the anonymous self-report condition and the bogus pipeline condition revealed generally similar responses. These findings strengthen the argument that responses under public and private reporting conditions are not equivalent and raise doubt about the conclusion most often drawn in the literature regarding the bogus pipeline. Specifically, the bogus pipeline did not appear to draw out more truthful responses than a private paper-and-pencil measure. Instead, the findings from the current study indicate that the situation appears to be considerably more complex than previously believed.

That is, only those people who are aware that they sometimes respond with more bias than their personal beliefs indicate is appropriate (i.e. the high IMS, high EMS) responded with higher stereotype endorsement scores in the bogus pipeline than the private condition. Further, consistent with our conceptualization of the meaning underlying these people's responses in the bogus pipeline condition, stereotype endorsement scores in the bogus pipeline condition were associated with guilt and self-criticism for high IMS, high EMS participants. Such an affective reaction indicates that these participants' responses in the pipeline condition reflect a personal failure (i.e. unintended stereotype activation) (e.g. Devine et al., 1991; Higgins, 1987; Plant & Devine, 1998).

All other participants responded with similar stereotype endorsement scores in the bogus pipeline and private reporting conditions, and their stereotype endorsement scores in the pipeline condition were not related to feelings of guilt and self-criticism. Importantly, participants who were primarily internally motivated (i.e. high IMS, low EMS) reported low stereotype endorsement scores across all of the

reporting conditions.⁹ Thus, even when hooked up to a 'truth detecting' machine, their responses revealed no greater stereotype endorsement than obtained in either the private or public conditions. These findings are consistent with previous work indicating that these people both anticipate responding with less bias and actually respond with less implicit prejudice toward Black people than do people representing all other combinations of IMS and EMS (Devine et al., 2002; Plant & Devine, 1998, 2000).

Participants low in internal motivation to respond without prejudice reported relatively high stereotype endorsement scores in both the private and bogus pipeline conditions. These responses are consistent with the private stereotype endorsement scores reported by Plant and Devine's (1998) low IMS participants as well as Devine and her colleagues' (2002) findings that low IMS participants tend to respond with prejudiced responses across implicit and privately reported explicit measures of prejudice.

It is also worth noting that our analyses involving participants' attitudes toward Blacks revealed that attitude alone is not sufficient to reveal the complexity of participants' responses across the reporting conditions. These findings indicate that it is important to assess the motivational forces underlying people's prejudice-relevant responses in order to gain a full understanding of what these responses mean.

Conclusion

As changes in social norms occur over time, the assessment of racial beliefs and attitudes has become an increasingly challenging endeavor. Whereas historically the challenge was to develop strategies, such as the bogus pipeline, to circumvent normative pressures in the reporting of racial beliefs, recent advances in theory and research provide new insights on meaning of responses supplied in this paradigm. Rather than the somewhat simplistic conclusion that pipeline responses are true and paper-and-pencil measures are suspect, our findings suggest that these alternative types of measures reflect different components of

people's attitudinal representations. We believe that both types of response are valid and need to be considered carefully as we explore the ongoing challenges involved in assessment of racial attitudes.

Notes

1. It may also be of interest that the IMS and traditional measures of prejudice (e.g. Attitude towards Blacks scale, MRS) are strongly correlated such that high IMS scores were associated with low-prejudice attitudes. The EMS, in contrast, is only modestly correlated with traditional prejudice measures such that high levels of external motivation are associated with high prejudice scores.
2. Although the vast majority of participants in this line of work have been White, a small percentage have been neither White nor Black (i.e. Hispanic, Asian). Therefore, in order to be inclusive, when appropriate, we use the term non-Black.
3. The bogus pipeline technique has been applied to a variety of other socially sensitive topics (e.g. expressions of dislike of others; reports of drug use). Roese and Jamieson (1993) provide a good review of the history of bogus pipeline research. In the present article, we focus only on issues related to assessment of racial stereotypes.
4. There have been alternative interpretations of the bogus pipeline technique proposed in the past (see Arkin, 1981; Roese & Jamieson, 1993), however, the normative standard explanation is the more widely accepted interpretation.
5. In the current study, we selected participants at the extremes of the IMS and EMS distributions. Although this approach limits the generalizability of the findings, because our complex set of predictions involved the experiences of people who respond with high and/or low levels of IMS and EMS, we believed this was the best initial approach to test these predictions.
6. In order to thoroughly assess the nature of the stereotype of Black people, we included positive stereotypic traits (i.e. athletic, rhythmic) in addition to the negative traits (i.e. criminal) in our stereotype endorsement index. However, it should be noted that when the positive traits were excluded from the stereotyping index the findings were similar in form and strength.
7. Other indices were identified (e.g. positive, depressed), but they were not of central theoretical interest and did not reveal systematic

differences; thus, they will not be discussed further.

8. We also examined the relationship between negself and the stereotype endorsement scores for the high IMS, high EMS participants in the pipeline condition, controlling for general discomfort (afraid, anxious, tense, uncomfortable, agitated, threatened, uneasy; $\alpha = .88$). The analysis continued to reveal a strong correlation between negself and stereotype endorsement ($r(11) = .60, p < .04$). In contrast, for the other participants in the bogus pipeline condition this correlation was very small or negative (all r s .05 or below).
9. A reviewer for this manuscript suggested that this group of participants may be composed of a combination of truly honest responders and 'effective liars'. Such a combination would result in lower mean levels of stereotype endorsement and less variable responses in the public reporting condition than in the pipeline condition. However, both the effect of reporting condition and a test of the homogeneity of variance across reporting condition for these participants resulted in F values < 1 , indicating that such an explanation of the findings was highly unlikely.

Acknowledgments

Preparation of this manuscript was partially supported by a University of Wisconsin Romnes Faculty Fellowship and by a Hilldale Undergraduate-Faculty Fellowship awarded to the second and third authors. The authors gratefully acknowledge the assistance of Marnie Hamilton, Melissa Henze, and Anna Liska in collecting and entering the data reported in this manuscript. We also thank David Amodio and Brenda Buswell for their thoughtful comments on a previous draft of this manuscript.

References

- Allport, G. W. (1954). *The nature of prejudice*. Reading, MA: Addison-Wesley.
- Arkin, R. M. (1981). Self-presentational styles. In J. T. Tedeschi (Ed.), *Impression management theory and social psychological research* (pp. 311–333). San Diego, CA: Academic Press.
- Banaji, M. R., & Greenwald, A. G. (1995). Gender stereotyping in judgments of fame. *Journal of Personality and Social Psychology, 68*, 181–198.
- Brigham, J. C. (1993). College students' racial attitudes. *Journal of Applied and Social Psychology, 23*, 1933–1967.
- Brigham, J. C., Bloom, L. M., Gunn, S. P., & Torok, T. (1974). Attitude measurement via the bogus pipeline: A dry well? *Representative Research in Social Psychology, 5*, 97–114.
- Crosby, F., Bromley, S., & Saxe, L. (1980). Recent unobtrusive studies of blacks and white discrimination and prejudice: A literature review. *Psychological Bulletin, 87*, 546–563.
- Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology, 56*, 5–18.
- Devine, P. G., & Elliot, A. J. (1995). Are racial stereotypes really fading? The Princeton trilogy revisited. *Personality and Social Psychology Bulletin, 21*, 1139–1150.
- Devine, P. G., & Monteith, M. J. (1993). The role of discrepancy-associated affect in prejudice reduction. In D. M. Mackie & D. L. Hamilton (Eds.), *Affect, cognition, and stereotyping: Interactive processes in group perception* (pp. 317–344). San Diego, CA: Academic.
- Devine, P. G., Monteith, M. J., Zuwerink, J. R., & Elliot, A. J. (1991). Prejudice with and without compunction. *Journal of Personality and Social Psychology, 60*, 817–830.
- Devine, P. G., Plant, E. A., Amodio, D. M., Harmon-Jones, E., & Vance, S. (2002). The regulation of explicit and implicit race bias: The role of motivations to respond without prejudice. *Journal of Personality and Social Psychology, 82*, 835–848.
- Dovidio, J. F., & Fazio, R. H. (1992). New technologies for direct and indirect assessment of attitudes. In J. M. Tanur (Ed.), *Questions about questions: Inquiries into the cognitive bases of surveys* (pp. 203–237). New York: Russell Sage Foundation.
- Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology, 69*, 1013–1027.
- Gaertner, S. L., & Dovidio, J. F. (1986). The aversive form of racism. In J. F. Dovidio & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 61–89). New York: Academic Press.
- Greenwald, A., & Banaji, M. (1995). Implicit social cognition: Attitudes, self-esteem and stereotypes. *Psychological Review, 102*, 4–27.
- Greenwald, A. G., McGhee, D. E., & Schwartz, L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test.

Journal of Personality and Social Psychology, 74, 1464–1480.

- Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review*, 94, 319–340.
- Jones, E. E., & Sigall, H. (1971). The bogus pipeline: A new paradigm for measuring affect and attitude. *Psychological Bulletin*, 76, 349–364.
- Lambert, A. J., Cronen, S., Chasteen, A. L., & Lickel, B. (1996). Private and public expressions of racial prejudice. *Journal of Experimental Social Psychology*, 32, 437–459.
- LaPiere, R. T. (1934). Attitude vs. paradigm. *Social Forces*, 13, 230–237.
- McConahay, J. B. (1986). Modern racism, ambivalence, and the modern racism scale. In J. F. Dovidio & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 91–126). San Diego, CA: Academic Press.
- Monteith, M. J., Deneen, N. E., & Tooman, G. D. (1996). The effect of social norm activation on the expression of opinions concerning gay men and blacks. *Basic and Applied Social Psychology*, 18, 267–288.
- Ostrom, T. M. (1973). The bogus pipeline: A new ignis fatuus? *Psychological Bulletin*, 79, 252–259.
- Pavlos, A. J. (1972). Racial attitude and stereotype change with bogus pipeline paradigm. *Proceedings of the 80th Annual Convention of the American Psychological Association*, 7, 292.
- Pavlos, A. J. (1973). Acute self-esteem effects on racial attitudes measured by rating scale and bogus pipeline. *Proceedings of the 81st Annual Convention of the American Psychological Association*, 7, 165–166.
- 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. (2000). *Approach and avoidance regulatory concerns for interracial interactions*. Unpublished doctoral dissertation, University of Wisconsin—Madison.
- Proshansky, H. M. (1966). The development of intergroup attitudes. In L. W. Hoffman & M. L. Hoffman (Eds.), *Review of child development research* (Vol. 2, pp. 311–371). New York: Russell Sage Foundation.
- Roese, N. J., & Jamieson, D. W. (1993). Twenty years of bogus pipeline research: A critical review and meta-analysis. *Psychological Bulletin*, 114, 363–375.
- Ryan, R. M., & Connell, J. P. (1989). Perceived locus of causality and internalization: Examining reasons for acting in two domains. *Journal of Personality and Social Psychology*, 57, 749–761.
- Ryan, R. M., Sheldon, K. M., Kasser, T., & Deci, E. L. (1996). All goals are not created equal: An organismic perspective on the nature of goals and their regulation. In P. M. Gollwitzer & J. A. Bargh (Eds.), *The psychology of action: Linking cognition and motivation to behavior* (pp. 7–26). New York: Guilford.
- Sigall, H., & Page, R. A., (1971). Current stereotypes: A little fading, a little faking. *Journal of Personality and Social Psychology*, 16, 252–258.
- Wilson, T. D., Lindsey, S., & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review*, 107, 101–126.
- Wittenbrink, B., Judd, C. M., & Park, B. (1997). Evidence for racial prejudice at the implicit level and its relationship with questionnaire measures. *Journal of Personality and Social Psychology*, 72, 262–274.

Paper received 1 June 2001; revised version accepted 25 June 2002.

Biographical notes

E. ASHBY PLANT is an assistant professor of psychology at Florida State University. Her primary research interests involve interracial interactions and the prejudice reduction process.