Evaluating Diversity and Inclusion Programs Designed to Change Social Attitudes

Donald Fischer, Adena Young-Jones and Sequana Tolon

Missouri State University

ABSTRACT
Malleability of racial attitudes of students in courses with objectives that involve understanding and valuing diversity were assessed using both implicit and explicit measures. Support for the malleability of overt bias assessed by explicit measures was tempered with mixed support for the malleability of automatic bias assessed by implicit measures.

PRESS PARAGRAPH
Decreases in the overt expression of racist attitudes over the past half-century have been accompanied by evidence of pervasive implicit racial bias. Diversity and inclusion programs often endeavor to help participants learn about their own attitudes as a way of motivating positive change. Despite the widespread use of these programs, empirical evidence supporting their effectiveness has been lagging – one survey of colleges and universities reported that 81% used diversity training to address racial discrimination without any evaluation of their effectiveness. This study examined an intervention designed to change social attitudes and found mixed evidence in support of its effectiveness.
Diversity and inclusion initiatives are prevalent employment practices in both public and private sector organizations. It is estimated that over half of all employers in the U.S. with more than 100 employees, and over 75% of the largest of these, have diversity and inclusion programs (Hays-Thomas & Bendick, 2013). Globalization and an increasingly diverse workforce are often cited as factors that require cultural competence and inclusive climates to effectively manage human capital and leverage competitive advantage (Cascio & Aguinis, 2011; Ployhart 2012; Thomas, 2005). Diversity and inclusion initiatives address the problems of conscious and unconscious discrimination and one’s perception “that he or she is an esteemed member of the work group through experiencing treatment that satisfies his or her needs for belongingness and uniqueness” (Shore et al., 2011). Within this context an expansive list of attributes define diversity – gender and sexual orientation, race and ethnicity, age, weight, disability, religious practice, family status, educational background, personality traits, work styles, and functional specialties are all pertinent when defining the construct of diversity in the workplace (Hays-Thomas & Bendick, 2013).

Interventions seeking to create climates which respect diversity and value inclusion typically focus upon changing social attitudes (Wah, 2012; Wise, 2011). Likewise, the curriculum in higher education often includes courses that target social attitudes in an effort to develop egalitarian values and respect for diversity (Blaine, 2011). The pedagogy used in college courses frequently reflects what is used by organizational change agents in diversity and inclusion programs. These efforts often begin with the proposition that psychological health and personal effectiveness are fostered by experiential openness and self-awareness: positive change begins with accurate self-knowledge. Added to this proposition is the corollary that few of us have accurate self-knowledge when it comes to our social attitudes and how these influence our
behavior – many, if not most of us, do not accurately perceive the impact our behavior has upon others. Indeed, it is hard to find a chapter in an adjustment text in Psychology that does not include a self-assessment activity which encourages students to learn more about their current status in order to identify the direction and nature of steps they might take to achieve positive change (Nevid & Rathus, 2010; Santrock, 2006; Watson & Tharp, 2002). Similar to these initiatives in the educational setting, when the focus is upon understanding and changing social attitudes in the workplace, interventions often help participants confront accurate assessments of the beliefs and feelings they associate with target-group members, along with the behavior these beliefs and feelings foster (Wah, 2012; Wise 2011).

A decrease in the overt expression of racist attitudes over the past half-century has been accompanied by evidence of pervasive implicit racial bias (Gaetner & Dovidio, 2005; Nosek et al., 2007). In his engaging best seller *Thinking Fast and Slow*, Daniel Kahneman (2011) nicely illustrates a well-established tenet of contemporary cognitive science: both automatic or intuitive cognitive processes of which we are often unaware and effortful or deliberate cognitive processes of which we are often painfully aware, influence the way we think, feel and act. Tony Greenwald and others have illustrated how these automatic and effortful dynamics influence our social attitudes and the ways we interact with others (Dovidio et al., 1997; Greenwald et al., 2002). Corresponding to the automatic and effortful cognitive processes that Kahneman and Greenwald describe, behavioral scientists distinguish between implicit and explicit measures, where the latter are typically self-reports that provide an opportunity to reflect and deliberate before responding and the former attempt to reduce the effects that such reflection and deliberation might produce. In particular, Greenwald, McGhee, and Swartz (1998) suggest that the Implicit Association Test (IAT), which is based upon reaction times in classification tasks, is
resistant to self-presentation artifacts and independent of a person’s introspective ability or self-
knowledge. People who espouse egalitarian values and describe themselves as not the kind who
judge people according to their race, but rather according to individual merit on a case-by-case
basis, often have IAT scores that indicate a strong tendency to associate negative attributes with
Blacks (Greenwald, McGhee, & Swartz, 1998). Consequently, implicit measures like the IAT are
useful when assessing racial/ethnic attitudes, particularly in settings where demand
characteristics (e.g., referent group norms) motivate people to respond in socially acceptable
ways (Fazio & Olsen, 2003).

Considerable empirical evidence links explicit and implicit measures of negative social
attitudes with dysfunctional and harmful behavior, especially when the attitudes involve racial
and ethnic minorities. Gordon Allport’s classic The Nature of Prejudice (Allport, 1954), along
with Thomas Pettigrew’s half-century of work (Campbell & Pettigrew, 1959; Pettigrew &
Martin, 1979; Pettigrew, 2011), vividly illustrate how negative social attitudes can impede
effectiveness. Recent interest has focused upon the incremental contributions that implicit
measures can make regarding the prediction of individual conduct in social settings (Greenwald,
Poehlman, Uhlmann, & Banaji, 2009). For example, implicit measures of racial/ethnic attitudes
predict differences in police officers’ decisions to shoot or not shoot a suspect pulling an object
from his pocket (a wallet or weapon) in a simulation exercise that varies the race/ethnicity of the
suspect, better than explicit measures of racial/ethnic attitudes (Eberhardt, 2007). Perhaps of
more relevance to diversity and inclusion initiatives in organizational contexts are studies that
link IAT measures of racial attitudes with interpersonal space and collaboration initiatives with
Blacks (Amodio & Devine, 2006), discomfort and negative interactions with Blacks (Dovidio,
Kawakami & Gaertner, 2002), and negative evaluations of Black job applicants (Rooth, 2010;
These studies provide a rationale for organizations, including universities, to be concerned about the nature and prevalence of negative racial attitudes among members, and they provide an incentive for these organizations to use implicit measures when assessing their endeavors to change these attitudes.

While it is easy to find empirical evidence linking social attitudes with interpersonal behavior, it is not as easy to locate evidence for the effectiveness of diversity and inclusion programs (including university courses) designed to change negative social attitudes, particularly when the evidence includes implicit measures of relevant attitudes (Gregg, Seibt & Banaji 2006; Joy-Gaba & Nosek, 2010). Many have lamented the lack of solid, evidence-based information to guide practitioners and many have called for more program evaluation research (Kalev, Dobbin & Kelly, 2006; Hite & McDonald, 2006). There are, however, some notable efforts being made in this regard. For example, Lai et al. (2012) report a meta-analytic investigation of 18 interventions designed to reduce implicit racial bias. Interventions were organized into six descriptive categories according to the most distinctive feature of each intervention’s design and interventions in three of the six categories were found to be effective in reducing implicit racial bias, with involvement (the degree to which an attitude is linked with the self) being a common feature among the effective interventions. Phillips et al. (2011) reported that interventions using implicit “approach training” procedures to strengthen associations between the self and Blacks were effective in reducing racial bias among non-Blacks. Another study reported by Rudman, Ashmore, and Gary (2001) compared changes in the attitudes of students in an upper-division prejudice and conflict seminar taught by a Black instructor, where the course content involved relevant topics, with those in a research methods course taught by a White instructor, where the course content did not involve relevant topics. They found support for the malleability of
students’ attitudes on both explicit and implicit measures for students in the prejudice and conflict course but not for students in the experimental methods course. However, Tolon, Fischer, and Young-Jones (2013) were only partially successful when they attempted to replicate and extend the findings of Rudman, Ashmore, and Gary (2001). They used multiple implicit and explicit measures of racial attitudes in a repeated measures design to assess the effectiveness of activities that were embedded in two undergraduate psychology courses – an adjustment course taught by a Black instructor and a diversity course taught by a White instructor. A mid-term assessment of racial attitudes toward Blacks was used to introduce a unit that examined racial and ethnic attitudes in both courses. Assessment results were provided to students during a feedback session designed to help students understand what their scores mean and how they might use this information. A second assessment administered during the final week of classes indicated some support for the effectiveness of the activities, but only as reflected in one of the explicit measures. Tolon, Fischer, and Young-Jones (2013) identified several problems with their design, including limited statistical power, technical glitches with the procedures used to assess students’ attitudes, and the exclusion of relevant intervention activities from the pre-posttest interval caused by the midterm pretest assessments. The present study attempts to address these problems and, like previous efforts, help identify effective methods for developing attitudes that respect diversity and contribute to inclusive organizational climates.

**Hypothesis:** Effectiveness of the assessment and intervention activities designed to change racial attitudes will be evident in a pre-post change on both the explicit and implicit measures of students’ attitudes. More specifically, White students will explicitly express less resentment and hostility toward Blacks and greater desire to control their prejudicial reactions toward Blacks on their posttest measures, relative to their pretest measures. Similarly, White
students will implicitly express a stronger association of Blacks with positive evaluations and mental strength on the posttest measures, relative to their pretest measures.

If the assessments and subsequent feedback activities are effective in helping students confront undesirable aspects of their racial attitudes, their enhanced self-awareness should motivate a desire to change, which should be reflected in the posttest assessments. We expect other activities in the diversity course to further enhance students’ motivation to change the way they consciously think about Blacks and try to control their expression of prejudicial behavior toward Blacks. However, we expect these gains to be more evident in the conscious, deliberative thinking of these students rather than in the implicit associations and automatic processes that are experientially conditioned from childhood to the present. We expect this because of limited exposure to positive, prosocial experiences with racially diverse others which a course with few (at most two) Blacks per section provides (Pettigrew, 2011; Zajonc, 1968).

METHOD

Participants

Participants (N = 109) were recruited from a diversity course (The Psychology of Diverse Populations) taught at a public university in the Midwest that has relatively few racial and ethnic minority students (about 15%). Although all students were taught by the same instructor, they were enrolled in five different sections distributed across three consecutive semesters.

Measures

Explicit Racial Prejudice. This measure combined the seven items of the Modern Racism Scale (McConahay, 1986) with the four items of the Explicit Racial Resentment Scale (Wilson & Davis, 2008). Higher scores indicate more negative attitudes toward Blacks.
Motivation to Control Prejudiced Reactions Scale. This measure was the 17 item scale developed by Dunton and Fazio (1997) to assess concern about acting prejudiced and the desire to avoid dispute. Higher scores indicate greater motivation to control prejudicial behavior.

Racial Preference IAT. An IAT modeled after those used by Greenwald and others was used to assess the degree of students’ implicit evaluative race bias. The seven block procedure and scoring algorithm (D measure) recommended by Greenwald et al. (2003) was used with images (cropped pictures of 12 faces: six Black male and female faces, and six White male and female faces) and 16 word stimuli (see Table 1). The more negative the IAT effect, the stronger the implicit association between negative valence and Blacks (relative to a positive valence with Whites).

Racial Stereotyping IAT. An IAT modeled after that developed by Amodio and Devine (2006) was used to assess the degree of students’ implicit racial stereotyping. The same seven block procedure, scoring algorithm, and face stimuli described above were used with 12 word stimuli to obtain this measure (see Table 1). The more negative the IAT effect, the stronger the implicit association of mental strength with Whites (and physical strength with Blacks).

Procedure

The racial preference and stereotyping IATs were administered through a website hosted by Millisecond, Inc., and InQsit software (Fortriede, 2005) was used to create a web-based questionnaire for the explicit measures. All measures were self-administered by students in a computer lab during the first week of the semester. A feedback session modeled after that used by Tolon, Fischer, and Young-Jones (2013) was conducted in conjunction with the unit that addressed racial and ethnic minorities. This activity provided a brief review of dual process attitude theory, in addition to interpretive suggestions about what scores within various ranges
might mean and how students might use this information. The same measures were administered a second time during the final week of classes.

RESULTS

Descriptive statistics for the study variables are presented in Table 2. These results suggest that the measures provide adequate variance and internal consistency to support additional analyses. Table 3 contains zero-order correlations for study variables. These results provide evidence of convergent validity for both the explicit and implicit measures in that measures within each category were significantly correlated on both the pretest and posttest. These results also provide evidence of discriminant validity and dissociation between the two types of measures in that only two of the twelve correlations involving an implicit and an explicit measure were significant. With respect to sample demographics, median age was 22 years, 74% were female and 85% were White.

Tests of Hypothesis

Only non-Black students (N = 100) were included in the sample used to test the hypothesis because our predictions and measures target attitudes toward Blacks. Consistent with our predictions, a multivariate analysis of variance for a repeated measures design produced a significant within-subjects main effect for the pre-posttest factor (Wilks’ Lambda = .74, F (4, 282) = 8.16, p < .001, partial eta-squared = .26). Univariate analyses of the four dependent measures produced significant within-subject main effects for pre-posttest differences on only the Racial Prejudice measure (F (1, 95) = 28.48, p < .001, partial eta-squared = .23) and the Motivation to Control measure (F (1, 95) = 7.08, p < .01, partial eta-squared = .07). An examination of cell means (see Table 4) indicated that the pre-posttest difference was in the predicted direction for both measures. Overall, students expressed less hostility toward Blacks
and greater desire to control their prejudicial behavior on the posttest. In addition to these main effects, there was a significant pre-posttest-by-section interaction effect (Wilks’ Lambda = .69, F (16, 282) = 2.29, p < .01, partial eta-squared = .09). An examination of these effects revealed that the pre-posttest differences were consistently more positive on the explicit measures compared to the implicit measures (see Figures 1 – 4). This was most evident on the stereotype IAT measure where two sections had smaller posttest means and two sections had larger posttest means, relative to the pretest (F(4, 95) = 2.91, p < .05, partial eta-squared = .11).

DISCUSSION

This study used both implicit and explicit measures to investigate the malleability of racial attitudes in the context of a university course. Overall, the results indicated a significant change in students’ explicit racial prejudice and motivation to control their prejudice without a corresponding significant change in their implicit attitudes. As such, the results only partially replicate the findings of Rudman, Ashmore, and Gary (2001). Nevertheless, the strength of the explicit findings represents evidence of the intervention’s effectiveness. Being aware of one’s prejudices and motivated to regulate these beliefs is a necessary first step in the direction of positive change for those seeking to create sustainable organizational climates that are inclusive and value diversity. Oftentimes, at the onset of interventions to change social attitudes, participants do not believe they are prejudiced against those who are different. Results from this study provide evidence that, as a consequence of the intervention activities, students’ awareness of their prejudicial attitudes was heightened. Thus, the explicit and conscious manifestation of prejudice was addressed.

Although this study did not find consistent support for the malleability of automatic bias assessed by implicit measures, the implicit assessments encouraged students to confront the
possibility they are at risk of being prejudiced in ways they are not consciously aware. The idea that one may be influenced in ways he or she is not aware may lead to greater concern about monitoring one’s conduct and a heightened explicit desire to control prejudicial reactions. On the other hand, this finding begs the question why the effects assessed by the implicit measures were so variable across the five sections: three of the five sections showed positive change on the racial preference IAT and two did not; two of the five sections showed positive change on the stereotype IAT and three did not (see Figures 3 and 4). Perhaps, in accord with the findings of Lai et al. (2012), the number of involved students (i.e., those self-identified with an attitude) in the sections with positive change reached critical mass. Future research with interventions that include confrontational components should assess students’ involvement so that its mediating effects might be studied. Future studies might also profit from investigating how the “approach training” interventions described by Phills et al. (2011) can be combined with intervention activities like those used here as a way to leverage student involvement.

Regardless of their involvement, students expressed recognition of implicit beliefs throughout the course, which they did not originally know existed; they also revealed motivation to work through those implicit beliefs. This alone supports the effectiveness of the intervention in terms of helping individuals understand that unconscious ideas exist and are important. When people learn about their own implicit beliefs and the resulting impact it can have on behavior, this can inspire them to try and deal with their deeply ingrained mindsets.
Notes

1. Size limitations prevent a detailed description of the intervention activities within the text of this paper; more information (PowerPoint presentation and course syllabus) is available from the authors upon request.

2. The authors want to acknowledge the contributions of research team members Jacque Byrket, Lacy Gillard, Teanna Lucas and Louis Oberdeiar, and express our gratitude to them for making this research possible.
References


Table 1.

*Concept Labels (in italics) and Word Stimuli for the Implicit Association Tests*

<table>
<thead>
<tr>
<th>Race</th>
<th>Evaluation</th>
<th>Stereotype</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Black</strong></td>
<td>joy</td>
<td>math</td>
</tr>
<tr>
<td>6 faces</td>
<td>terrible</td>
<td>athletic</td>
</tr>
<tr>
<td><strong>White</strong></td>
<td></td>
<td>run</td>
</tr>
<tr>
<td>6 faces</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Positive</strong></td>
<td>glorious</td>
<td>educated</td>
</tr>
<tr>
<td></td>
<td>nasty</td>
<td></td>
</tr>
<tr>
<td><strong>Negative</strong></td>
<td>wonderful</td>
<td>scientist</td>
</tr>
<tr>
<td></td>
<td>evil</td>
<td></td>
</tr>
<tr>
<td><strong>Mental Strength</strong></td>
<td>love</td>
<td>smart</td>
</tr>
<tr>
<td></td>
<td>hurt</td>
<td></td>
</tr>
<tr>
<td><strong>Physical Ability</strong></td>
<td>happy</td>
<td>college</td>
</tr>
<tr>
<td></td>
<td>horrible</td>
<td></td>
</tr>
<tr>
<td>laughter</td>
<td>failure</td>
<td>read</td>
</tr>
<tr>
<td>pleasure</td>
<td>awful</td>
<td>rhythmic</td>
</tr>
<tr>
<td>peace</td>
<td>agony</td>
<td></td>
</tr>
</tbody>
</table>
Table 2.

*Descriptive Statistics for Study Variables*

<table>
<thead>
<tr>
<th>Variables</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pretest Measures</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Racial Prejudice</td>
<td>109</td>
<td>2.11</td>
<td>.59</td>
<td>.84</td>
</tr>
<tr>
<td>Motivation to Control</td>
<td>109</td>
<td>4.46</td>
<td>.69</td>
<td>.78</td>
</tr>
<tr>
<td>IAT-Preference</td>
<td>109</td>
<td>-0.442</td>
<td>.37</td>
<td>.67</td>
</tr>
<tr>
<td>IAT-Stereotype</td>
<td>109</td>
<td>-0.188</td>
<td>.37</td>
<td>.74</td>
</tr>
<tr>
<td><strong>Posttest Measures</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Racial Prejudice</td>
<td>109</td>
<td>1.91</td>
<td>.55</td>
<td>.83</td>
</tr>
<tr>
<td>Motivation to Control</td>
<td>109</td>
<td>4.61</td>
<td>.69</td>
<td>.79</td>
</tr>
<tr>
<td>IAT-Preference</td>
<td>109</td>
<td>-0.432</td>
<td>.38</td>
<td>.60</td>
</tr>
<tr>
<td>IAT-Stereotype</td>
<td>109</td>
<td>-0.190</td>
<td>.37</td>
<td>.60</td>
</tr>
</tbody>
</table>

*a* All implicit measures are IAT effects expressed as D measures (Greenwald et al., 2003).

*b* Larger negative values indicate a stronger association of Black+negative (and White+positive).

*c* Larger negative values indicate a stronger association of Black+physical (and White+mental).
Table 3.

Zero-order Correlations for Study Variables

<table>
<thead>
<tr>
<th>Variables</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pretest Measures</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Motivation to Control</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Racial Prejudice</td>
<td>-.21*</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. IAT-Preference</td>
<td>-.03</td>
<td>-.23*</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. IAT-Stereotype</td>
<td>-.08</td>
<td>-.05</td>
<td>.22**</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Posttest Measures</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Motivation to Control</td>
<td>-.70**</td>
<td>-.20*</td>
<td>-.15</td>
<td>-.09</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Racial Prejudice</td>
<td>-.23*</td>
<td>.77**</td>
<td>-.13</td>
<td>-.07</td>
<td>-.33**</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. IAT-Preference</td>
<td>-.16</td>
<td>-.15</td>
<td>.47**</td>
<td>-.18</td>
<td>-.11</td>
<td>-.21*</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>8. IAT-Stereotype</td>
<td>-.09</td>
<td>-.06</td>
<td>.36**</td>
<td>.41**</td>
<td>-.18</td>
<td>-.07</td>
<td>.28**</td>
<td>-</td>
</tr>
</tbody>
</table>

Note: N = 109 (all students with complete data)

*p < .05; **p < .01
Table 4.

*Cell Means for Study Variables Used in the Analysis of Variance.*

<table>
<thead>
<tr>
<th>Sectiona</th>
<th>Nb</th>
<th>Pretest</th>
<th>Posttest</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>RPc MCd IAT-Pf</td>
<td>IAT-Sf</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FA12-1</td>
<td>14</td>
<td>2.08</td>
<td>4.56</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FA12-2</td>
<td>17</td>
<td>2.30</td>
<td>4.50</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SP13-1</td>
<td>22</td>
<td>2.10</td>
<td>4.75</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SP13-2</td>
<td>25</td>
<td>2.09</td>
<td>4.27</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SU13-1</td>
<td>22</td>
<td>2.11</td>
<td>4.44</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*aFA12-1 = fall semester 2012, section 1; FA12-2 = fall semester 2012, section 2; SP13-1 = spring semester, section 1; SP13-2 = spring semester 2013, section 2; SU13-1 = summer semester 2013, section 1.*

bNon-Black students with complete data. cRP = Racial Prejudice; dMC = Motivation to Control; eIAT-P = IAT-Preference; fIAT-S = IAT-Stereotype.
Figure 1. Cell means for the Racial Prejudice measure.
Figure 2. Cell means for the Motivation to Control measure.
Figure 3. Cell means for the IAT-Preference measure.
Figure 4. Cell means for the IAT-Stereotype measure.