Training and Education in Professional Psychology
2008, Vol. 2, No. 2, 103–110

1931-3918/08/$12.00 DOI: 10.1037/1931-3918.2.2.103

The Relationship Between Level of Training, Implicit Bias, and Multicultural Competency Among Counselor Trainees

Guy A. Boysen
SUNY Fredonia

David L. Vogel
Iowa State University

The assessment of attitudes toward diversity among counselor trainees has relied on self-report measures. Implicit measures might offer a valuable addition to self-report because they assess biased attitudes indirectly, do not rely on conscious introspection, and often demonstrate bias that contradicts self-reported attitudes. A sample ($N = 105$) of counselor trainees was assessed with measures of implicit bias toward African Americans and lesbians and gay men and a measure of self-reported multicultural competency. Implicit bias was present among counselor trainees despite high self-reported multicultural competency. In addition, self-reported multicultural competency varied by training level, but implicit bias did not. The results suggest that implicit bias can add to the understanding, assessment, and training of multicultural counselor competency.

Keywords: multicultural training, implicit bias, implicit attitudes, multicultural competency

A danger in counseling is that biased attitudes will influence the treatment of diverse clients. Because of the potential for negative outcomes, the current emphasis on providing multicultural training is necessary (American Counseling Association, 2005; American Psychological Association [APA], 2002). Fortunately, research has indicated multicultural courses are effective in improving counselor trainees' competency (e.g., Diaz-Lazaro & Cohen, 2001). However, this research has only used self-reports and has not used validated measures of specific types of bias (Kiselica, Maben, & Locke, 1999). Therefore, the current study will assess not only self-reported multicultural competency, but also counselor trainees' bias toward underrepresented groups using implicit measures, which assess attitudes indirectly and do not rely on conscious introspection.

Multicultural Training and Outcomes Assessment

A small but consistent set of studies has illustrated the effectiveness of counselor training courses aimed at increasing multicultural the competencies of knowledge, awareness, and skill (e.g., Diaz-Lazaro & Cohen, 2001; Neville, Heppner, Thompson, Brooks, & Baker, 1996; Wang, 1998). These assessments have included self-report measures of multicultural competency (MCC) such as the Multicultural Counseling Awareness Scale (MCAS; Ponterotto et al., 1996) or the Multicultural Awareness/Knowledge/Skills Survey (MAKSS; D'Andrea, Daniels, & Heck., 1991). The MCAS and MAKSS are part of the family of MCC instruments that also includes the Cross Cultural Counseling Inventory-Revised (CCCI-R; LaFromboise, Coleman, & Hernandez, 1991), and the Multicultural Counseling Inventory (MCI; Sodowsky, Taffe, Gutkin, & Wise, 1994). MCC measurement has also expanded to include specific groups as illustrated by the Sexual Orientation Counselor Competency Scale (SOCCS; Bidell, 2005).

The MCC measures seem to provide reliable and valid assessments of multicultural training outcomes. A recent review and analysis indicated that the MCC measures all have adequate internal consistency reliability (Dunn, Smith, & Montoya, 2006). Criterion validity has also been established by relating the measures with the tendency to minimize the effects of racism (negative relationship; MCAS; Neville et al., 2006), the ability to think in a flexible manner (MKASS; Kim, Cartwright, Asay, D'Andrea, 2003), and competency in responding to a fictional minority client (CCCI-R; Worthington, Mobley, Franks, & Tan, 2000). In addition, the measures tend to be positively correlated with each other, which illustrates their convergent validity (Constantine & Ladany, 2000).

Researchers have used MCC measures to document the effectiveness of training; however, further investigation is necessary for several reasons. First, researchers have not used validated measures of bias (Kiselica et al., 1999). Although researchers have documented change using MCC scales that correlate with bias, no study has documented change in bias toward a specific group with a validated measure, such as the Modern Racism Scale (McConahay, Hardee, & Batts, 1981) or the Index of Homophobia (Hudson & Rickets, 1980). Second, researchers have not assessed multicultural training using outcomes that do not rely on self-report. Self-reported MCC may not be significantly related to MCC rated by others (Worthington et al., 2000). In addition, counselor trainees may alter their self-reported attitudes in order to appear socially acceptable (Pope-Davis, Liu, Toporek, & Brittan-Powell, 2001). In fact, researchers have illustrated that social desirability correlates with MCC (e.g., Constantine &

GUY A. BOYSEN received his PhD in psychology from Iowa State University in 2006. He is currently an assistant professor of psychology at SUNY Fredonia in Fredonia, NY. Dr. Boysen's research interests include the impact of education on implicit and explicit bias and the definition of mental illness.

DAVID L. VOGEL is an associate professor in the Department of Psychology at Iowa State University. He received his PhD in counseling psychology from the University of Florida in 2000. His professional interests focus on stereotyping and stigma in the counseling process, counselor training, and the decision to seek professional help.

CORRESPONDENCE CONCERNING THIS ARTICLE should be addressed to Guy A. Boysen, Department of Psychology, W357 Thompson Hall, SUNY Fredonia, Fredonia, NY 14063. E-mail: guy.boysen@fredonia.edu

Ladany, 2000). Third, researchers have not explored how multicultural training affects implicit attitudes. Research has demonstrated the presence of both explicit and implicit components of attitudes (Greenwald & Banaji, 1995). Explicit attitudes—conscious attitudes available for self-report—and implicit attitudes—attitudes that do not rely on conscious self-report—often diverge (Nosek, 2005). Therefore, the effect of training on implicit attitudes is unknown.

Implicit Measurement of Bias

Implicit measures "provide estimates of individuals' attitudes without our having to directly ask them for such information" (Fazio & Olson, 2003, p. 303). Specifically, most implicit measures assess "actions or judgments that are under the control of automatically activated evaluation, without the performer's awareness of that causation" (Greenwald, McGhee, & Schwartz, 1998, p. 1464). Therefore, implicit measures assess attitudes in a way that avoids the use of conscious introspection. For example, Greenwald and colleagues (1998) created the Implicit Associating Test (IAT). The IAT measures the efficiency with which a person can associate concepts by having the person engage in a categorization task. Greenwald and colleagues proposed that for most European Americans, African American faces are more easily associated with the concept bad and European American faces are more easily associated with the concept good. Thus, the IAT consists of categorizations that are congruent (European American and Good/African American and Bad) and incongruent (European American and Bad/African American and Good). Implicit bias, or an IAT effect, is said to exist when reaction times are significantly faster for the congruent categorizations compared to the incongruent categorizations. Quicker reaction times indicate that people tend to have more positive cognitive associations with European Americans than African Americans.

Measuring attitudes implicitly can lead to divergences from individuals' self-reported attitudes. Greenwald and colleagues (1998) provided a dramatic illustration of the possible dissociation that has been replicated among various populations and targets of bias (e.g., Nosek, Banaji, & Greenwald, 2002). They showed that the majority of European American participants endorsed positive or neutral attitudes toward African Americans on an explicit measure, but exhibited negative attitudes toward African Americans on an implicit measure.

Implicit measures provide information that is different than self-reports in two major ways. First, unlike explicitly measured attitudes, attempts to fake or voluntarily control implicitly measured attitudes are largely ineffective (Kim, 2003). This does not mean that implicit measures are immune to environmental effects (Blair, 2002), but they do offer a method of assessment that individuals are less likely to manipulate intentionally. Second, implicit measures predict some subtle behavioral biases better than explicit measures (Greenwald, Poehlman, Uhlmann, & Banaji, 2006). For example, implicit bias predicts observers' ratings of friendliness during interactions with targets of bias (e.g., McConnell & Leibold, 2001). In addition, implicit bias predicts avoidance of targets of bias (e.g., sitting farther away; Rydell & McConnell, 2006). Hugenberg and Bodenhausen (2003, 2004) showed that implicitly measured bias was related to perceiving anger more quickly in African American faces than European American faces. Finally, interpersonal ratings made by African Americans who engaged in interethnic interactions with European American participants were predicted by the European American's level of implicit bias (McConnell & Leibold, 2001). These studies were performed in the laboratory with participants from the general population; as such, the effects may not generalize to practicing counselors. However, if implicit bias were to interfere with just one similar area it could affect multiculturally competent practice.

Several literature reviews and meta-analyses have supported the validity of implicit measurement and indicated that it is useful in research (Greenwald, & Nosek, 2001; Hofmann, Gawronski, Gschwendner, Le, & Schmitt, 2005; Lane, Banaji, Nosek, & Greenwald, 2007; Nosek, Greenwald, & Banaji, 2007; Greenwald et al., in press). Nonetheless, implicit measurement is a relatively new concept with some limitations. Some researchers have asserted that bias is just one of several explanations for faster reaction times during congruent IAT categorizations as compared with incongruent IAT categorizations (i.e., the IAT effect; Brendl, Markman, & Messner, 2001). In addition, the reliability of the IAT does not reach the level associated with self-report measures. The median test-retest correlation of the IAT is .56 (Nosek et al., 2007). The lower reliability might be explained by the fact that implicit measurement is influenced by subtle variations in the testing situation and stimuli used (Blair, 2002). Because of its reliability some researchers caution that IAT results should be interpreted only at the group level and not at the individual level (Steffens & Buchner, 2003). Therefore, the IAT cannot supplant MCC scales as the main tool of multicultural education outcomes assessment, but it may offer a valuable additional perspective.

Only one researcher has examined whether counselor trainees possess implicit bias. Abreu (1999) sampled graduate students, interns, and clinicians, and measured their bias by exposing them to words related to the African American stereotype (i.e., priming), which includes the view of African Americans as more hostile than European Americans (Devine, 1989). Participants then rated the hostility of a client. The stereotype exposure led to overestimations of hostility, and Abreu asserted that such implicit bias toward African Americans could influence clinical judgments of hostility.

Abreu's (1999) valuable study should be updated and expanded. Specifically, implicit attitudes about several targets of bias should be measured using a sample of current counselor trainees. The current generation of counselor trainees may not possess implicit bias due to rapidly changing cultural views. In addition, validated measures of bias toward several groups should be utilized. Using different measures and assessing multiple biases would increase the generalizability of the results. Finally, the effect of counselor training on implicit bias should be examined. Overall, the presence of implicit bias among counselor trainees and the effect of training on these biases is an unexplored area that deserves attention.

The Current Study

This study consisted of an assessment of MCC and implicit bias among counselor trainees. Multicultural competencies were measured using the CCCI-R (LaFromboise et al., 1991). In order to determine if counselor trainees possess a significant amount of implicit bias they were given pen and paper IATs (Lowery, Hardin, & Sinclair, 2001) measuring bias toward African Americans and lesbians and gay men. We hypothesized that counselor trainees would report high levels of MCC and that MCC would significantly vary by level of multicultural training. However, because of

the divergence that is often found between implicitly and explicitly measured bias and the initial documentation of implicit bias among counselors by Abreu (1999), we hypothesized that counselor trainees would exhibit implicit bias, and that this bias would not vary by level of multicultural training.

Method

Sample

Trainees ($N = 105$) enrolled in graduate programs in counseling from four universities participated in the study. All of the universities were in the Midwest United States. Two of the institutions were large land-grant universities and two were small urban universities. Degrees offered included APA-accredited doctoral and masters degrees in counseling psychology ($n = 53$), as well as masters degrees in rehabilitation counseling accredited by the Counsel on Rehabilitation Education (CORE), school counseling accredited by the state, and mental health counseling, which as unaccredited ($n = 52$). The number of participants in each of these degree programs was not collected in order to assure anonymity. The students in APA accredited programs were mostly female (75%), European American (75% European American; 15% African American; 6% Asian American; 8% Hispanic or Latino/a; 2% other), and heterosexual (4% homosexual; 11% bisexual). Many had completed advanced degrees (38% masters; 2% doctorate), and on average they had completed 3.27 ($SD = 1.89$) semesters of training, 1.37 ($SD = 1.93$) practicums, and had counseled 6 ($SD = 12$) minority clients and 1 ($SD = 2$) lesbian, gay, or bisexual (LGB) client. The students in CORE accredited programs were mostly female (75%), European American (90% European American; 4% African American; 2% Asian American; 4% Hispanic or Latino/a), and heterosexual (10% homosexual). Few had completed advanced degrees (10% masters), and on average they had completed 3.70 ($SD = 5.12$) semesters of training, .47 ($SD = 1.23$) practicums, and had counseled 12 ($SD = 30$) minority clients and 3 ($SD = 9$) LGB clients.

Participants were solicited from courses focusing on multicultural counseling competency, research methods, vocational behavior, professional ethics, and cognitive neuroscience. Different instructors taught the multicultural courses. However, the courses were similarly structured. The text in all classes was Sue and Sue (2003). Topics common to all the courses included the definition of MCC, characteristics of different ethnic groups and the LGB community, White privilege, and racial identity. Students also reflected on their multicultural experiences in all of the classes.

Measures

Multicultural competency. The MCC measure used was the CCCI-R (LaFromboise et al., 1991). The CCCI-R assesses the competencies of multicultural awareness, knowledge, and skill. The CCCI-R requires counselors to rate themselves on 20 items with a scale ranging from 1 (*strongly disagree*) to 6 (*strongly agree*). Constantine and Ladany (2000) demonstrated the validity of the CCCI-R by showing that it correlated with other measures of MCC ($r = .63$ to $.73$). In addition, it has shown internal consistency ranging from .88 to .95 (Constantine & Ladany, 2000; LaFromboise et al., 1991). The internal consistency in the current sample was .89.

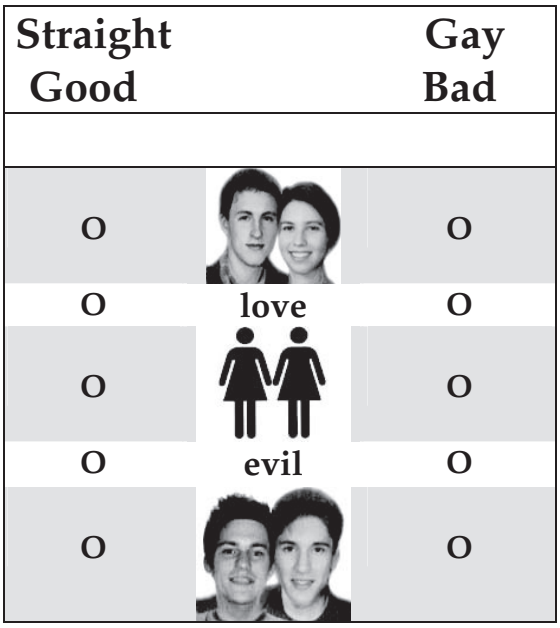
Implicit bias. The implicit measure was a pen and paper version of the IAT (Lowery et al., 2001). We utilized IATs measuring attitudes toward African Americans and lesbians and gay men. The pen and paper IAT consisted of pages with columns of words and pictures to be categorized into concepts represented on the left and the right of the columns (see Pruett & Chan, 2006 for a detailed description of a similar IAT). The African American IAT consisted of pictures of African American and European American faces. The lesbian and gay man (LG) IAT consisted of pictures of male/female, female/female, and male/male couples. Both IATs also included words representing the concepts of Good and Bad. Words with good connotations such as "friend" were used for the concept Good and words with bad connotations such as "tragic" were used for the concept Bad (see Figure 1 for an example).

Participants were asked to categorize the words and pictures by marking a circle next to the column that corresponded to the different categories. Categories were listed at the top of each page so that one was on the left side of the column and one was on the right side of the column. Participants filled out six separate pages and were given 20 seconds per page. Each page consisted of 44 individual items arranged in two columns of 22 items. Participants first went through two practice pages. The practice pages consisted of categorizing flower and insect names and good and bad words. Participants then completed the pages consisting of African American and European American faces and pictures representing same sex and opposite sex couples. The African American and LG IAT materials were counterbalanced.

Researchers have explored the reliability of pen and paper IATs designed to measure bias toward homosexuality (Boysen, 2005; Lemm, 2001) and African Americans (Boysen, 2005; Lowery et al., 2001). Implicit measures tend to have lower reliability coefficients than self-report measures due to the lack of conscious control over the responses (i.e., conscious recall of previous responses on explicit measures leads to increased consistency). Without this conscious control, implicit assessment may have test-retest correlations below .10 and even negative test-retest correlations (Bosson, Swann, & Pennebaker, 2000). Relatively speaking, however, the reliability of the pen and paper IAT is a vast improvement over some other implicit measures. For example, Lemm found that two consecutively administered versions of an IAT measuring bias toward homosexuality correlated .47, and Boysen found test-retest correlations ranging from .19 to .29 for the African American IAT and .21 to .43 for the LG IAT used in this study. These reliability estimates are similar to the computerized IAT (Cunningham, Preacher, & Banaji, 2001). Although neither the computerized nor pen and paper IAT currently possesses the reliability necessary to inform diagnostic decisions, they are appropriate for research purposes.

Support exists for the validity of the pen and paper IAT as well. Lemm (2001) found that the pen and paper IAT measuring implicit bias toward gay men was significantly correlated ($r = .38$) with a measure of explicit bias toward gay men. Similarly, using the IATs from this study Boysen found a significant correlation of .30 between the LG IAT and homophobia and a significant negative correlation ($r = -.24$) between the LG IAT and the number of

A



B

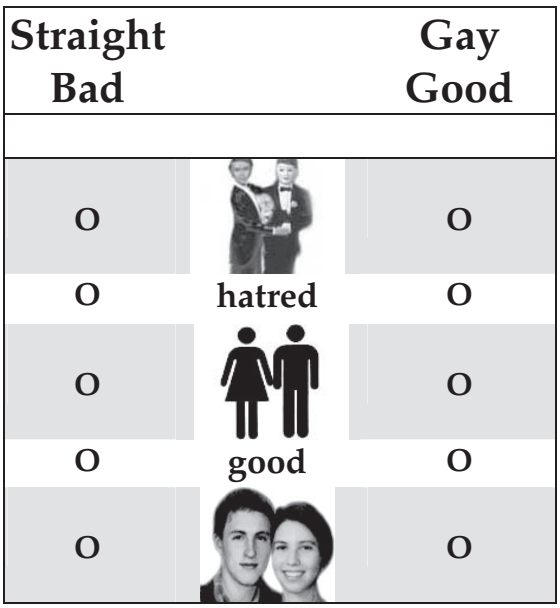


Figure 1. Example of stimuli from the congruent block of trials (A) and incongruent block of trials (B) from the lesbian and gay man IAT (stimuli from www.briannosek.com).

self-reported close relationships with LGB individuals. Perhaps most importantly all of the studies using pen and paper IATs have replicated previous computerized IAT research by identifying implicit bias.

Procedure

The assessment took place in class during the last week of the semester, so participants had just completed their courses. The first author administered all materials following the established procedure (Lowery et al., 2001). Participants first read and signed an informed consent document. Next, they read instructions for the IATs and completed it following the procedure outlined above. Finally, participants completed the CCCI-R and demographic questions (i.e., age, sex, ethnicity, sexual orientation, number of close relationships participants had with African American and LGB individuals, highest degree attained, semesters of graduate school completed, number of practicum completed, and whether the participant had taken a multicultural training class). After all the materials were completed, participants were debriefed.

Data Reduction

Scoring of the pen and paper IAT followed the standard procedure (see Lowery et al., 2001). The ratio of items completed per second in the congruent trials (e.g., Black-Bad/White-Good) was subtracted from the ratio completed in the congruent trials (e.g., White-Bad/Black-Good) to compute the IAT effect. Individuals with 30% or more errors or who completed less than 6 items total on a page were not included in the analysis (Nosek & Lane, 1999). This led to 10 participants being eliminated from analyses of the African American IAT ($n = 95$) and 19 participants being eliminated from the analyses of the LG IAT ($n = 84$). Chi-square analysis indicated that no group had a significantly different proportion of participants excluded from the analysis of the IATs due to excessive errors, $\chi^2 < .90$, $p > .64$.

Results

Preliminary Analyses

We examined the relations between the demographic (age, race, sexual orientation, number of close relationships with African American and LGB individuals) and educational (number of degrees, number of semesters of graduate school completed, specific program, practicums completed, number of African American and LGB clients) variables and the CCCI-R and IAT effects. Only the number of close relations with LGB people was significantly correlated with CCCI-R scores ($r = .21$, $p = .04$). No other correlations were significant (p's $> .08$). Similarly, exploratory analysis using the demographic and educational variables as predictors and the CCCI-R and the IATs as dependent variables yielded no significant results. Therefore, we did not include these variables in subsequent analyses of the dependent variables.

Multicultural Competency

Level of multicultural competency. We hypothesized that counselor trainees would report a high level of MCC. To test this hypothesis we examined CCCI-R scores in order to determine the attitudes counselors in training possessed about working with individuals from diverse backgrounds. The mean score on the CCCI-R was 96.73 with a standard deviation of 8.44. In a previously published sample of Masters and Doctoral level counselors ($n = 91$) and counselor trainees ($n = 44$), where 84% had under-

gone multicultural training, the average score on the CCCI-R was 95.56 (SD = 9.31; Constantine & Ladany, 2000). The average score in this study translates into a mean item rating on the CCCI-R of 4.84 (SD = .42), which corresponds roughly to a response of *agree*. In fact, the lowest score in the sample was a 73, which corresponds to an average response just below *slightly agree*. We used a one-sample t test to determine if the scores were significantly different than an average response of 3, which corresponds to *slightly disagree* on the CCCI-R. An average response of 3, it could be argued, represents a bias against working with diverse clients. The mean rating on the CCCI-R was significantly higher than 3, $t(102) = 20.11$, $p < .001$. Taken together, these results support our hypothesis that counselor trainees would report a high level of MCC.

Influence of training on multicultural competency. We hypothesized that MCC would vary by level of multicultural training. We tested the hypothesis by analyzing scores on the CCCI-R using an analysis of variance (ANOVA) with level of training serving as the independent variable. The counselor trainees consisted of a cross-sectional sample comprised of three levels of multicultural training: (1) trainees who had not yet completed a multicultural course (no course group), (2) counselor trainees who had just completed a multicultural course the semester the assessment occurred (multicultural course group), and (3) counselor trainees who had completed a multicultural course in a previous semester (multicultural course plus time group). Thus, trainees either had little exposure to multicultural material, full exposure to multicultural material, or full exposure plus subsequent time enrolled in the training program. Due to the unequal sample sizes we utilized the Welch statistic. The analyses indicated that significant differences did exist between the training groups, $F(2, 100) = 6.52$, $p = .003$. Means and standard deviations can be seen in Table 1. Post hoc Tukey's tests indicated that scores in the multicultural group were not significantly different than the no course group, $p = .77$. However, CCCI-R scores in the multicultural course plus time group were significantly higher than in the multicultural course group, $p = .004$, and were also higher than the no course group, but not significantly so, $p = .05$. These results partially support the hypothesis and suggest that a multicultural course plus subsequent time enrolled in a training program leads to increased MCC.

Implicit Bias

Level of implicit bias. We hypothesized that counselor trainees would possess significant levels of implicit bias. To test this we compared the congruent and incongruent blocks of the IAT to determine if a significant IAT effect was present among the counselor trainees. If significantly more items are categorized per second on the congruent block, as compared with the incongruent block, implicit bias is said to exist. Therefore, in order to determine if there was significant implicit bias in the sample, we analyzed the number of stimuli correctly categorized per second on the congruent and incongruent blocks of the African American and LG IATs using paired samples t tests. According to our hypothesis, significant bias should be present on both IATs. Analysis of the African American IAT indicated that categorizations were significantly faster, $t(94) = 4.28$, $p < .001$, in the congruent block (M = 1.15, SD = .88) than in the incongruent block (M = 1.59, SD = .66). Similarly, the LG IAT indicated that categorizations were significantly faster, $t(83) = 6.43$, $p < .001$, in the congruent block (M = 1.18, SD = .45) than in the incongruent block (M = 1.56, SD = .69). These differences indicate that implicit bias toward African Americans and lesbians and gay men was present among these counselor trainees, which confirmed our hypothesis.

Influence of training on implicit bias. We also hypothesized that, unlike the self-reported competencies, implicit bias would not vary by level of training. To test for differences in implicit bias based on training we analyzed the IAT effects of the African American and LG IATs using ANOVA's with level of training (no course, multicultural course, multicultural course plus time) serving as the independent variable. Due to the unequal sample sizes we utilized the Welch statistic. Although the multicultural course plus time group did have the lowest means, the ANOVA indicated that significant differences did not exist between the training groups for the African American IAT effect, $F(2, 92) = .93$, $p = .40$, or the LG IAT effect, $F(2, 81) = 2.56$, $p = .09$. Means and standard deviations of the IAT effects can be seen in Table 1. We also assessed for relations between MCC and implicit bias. No significant correlations existed between the CCCI-R and the African American ($r = -.16$, $p = .12$) or LG ($r = .10$, $p = .39$) IAT effects.

Table 1
Means and Standard Deviations of Cross Cultural Counseling Inventory – Revised Scores and IAT Effects of Multicultural Training Groups

	Training group								
	No course			Multicultural course			Course plus time		
	n	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
AA IAT	24	0.31	1.31	53	0.46	1.34	18	0.18	0.42
LG IAT	20	0.43	0.51	48	0.45	0.59	16	0.15	0.45
CCCI-R	27	96.40[b]	8.46	57	95.08[a]	8.21	19	102.16[a,b]	7.12

Note. No Course = never completed a multicultural course; Multicultural Course = completed a multicultural course the semester of the assessment; Course Plus Time = completed a multicultural course in a semester before the assessment; AA IAT = African American IAT effect; LG IAT = Lesbian and Gay Man IAT effect; CCCI-R = Cross Cultural Counseling Inventory-Revised. Means that share a superscript are different at the following significance levels: [a] $p = .004$. [b] $p = .05$.

Thus, the hypothesis was supported and unlike MCC there was no evidence that training had a significant effect on the implicit bias.

Discussion

Using a sample of counselor trainees from four different programs, we discovered significant levels of implicit bias toward African Americans and lesbians and gay men. At the same time, the counselor trainees reported high levels of MCC overall, and counselor trainees who were furthest along in their program reported the highest level of competency. In contrast, implicit bias did not vary significantly based on level of training. Together, these results suggest divergence between counselor trainees' explicitly and implicitly measured attitudes.

Relation to Past Research

The use of implicit measures to assess bias among counseling trainees is new, but some comparisons with past work can be made. Broadly speaking, discovering that counselor trainees possess implicit bias is consistent with past research examining attitudes of noncounselor populations. Thousands of people have taken IATs measuring bias toward African Americans and lesbians and gay men and bias is typically found (Nosek et al., 2002). The results were also consistent with the one other study that sampled trainees and found evidence of implicit bias toward African Americans (Abreu, 1999). Implicit bias toward African Americans and lesbians and gay men is prevalent in the United States and counselor trainees seem to be no exception.

Individuals in the current study who had completed a multicultural counseling course in a previous semester self-reported a higher level of MCC than those who had never completed or just completed a multicultural course. In one sense, this finding is consistent with the studies that have shown the effectiveness of multicultural training (e.g., Diaz-Lazaro & Cohen, 2001). However, our finding that counselor trainees who had at least a semester of additional training after their multicultural course exhibited significantly higher competency than even those counselor trainees who had just completed a multicultural course seems to contradict some previous results. The few studies that have presented longitudinal data indicate that the effects of multicultural training remain constant after the initial postcourse increase (Neville Heppner, Thompson, Brooks, & Baker, 1996; Wang, 1998). However, with so few studies having examined long-term effects of multicultural training and the cross-sectional design of this study, it is unclear at this point if attitudes should be expected to remain constant or continue developing.

The lack of a significant difference between trainees who did not take a multicultural course and those who just completed a course is also difficult to interpret. This could be considered an accurate assessment of the effectiveness of counseling programs in promoting multiculturalism across the curriculum. It may be that calls for affirmation of multicultural principles at the organizational level have been answered (Phillips, 2000). In contrast, it may be that counselor trainees who have not taken a multicultural course simply lack the knowledge necessary for accurate self-assessment, and thus, systematically overestimated their competency (Pope-Davis et al., 2001). Because the average competency ratings were near the ceiling of the assessment tool, the results suggest that the counselor trainees' were responding to the measure in a socially desirable manner. In all, the difficulty in interpreting the self-report results with confidence illustrates the value in adding other types of measures to broaden the information gathered during multicultural assessments.

Implications

The major finding in this study is the identification of significant implicit bias in a sample of counselor trainees; unfortunately, this leads to more questions than answers. From a measurement standpoint, the IAT does not have the reliability associated with self-report measures. As such, interpreting implicit bias as a stable trait might be questioned. One reason for the low reliability might be the malleability of implicit attitudes. IAT scores can vary based on the testing conditions or the stimuli utilized (Blair, 2002). Theoretically, some aspect of the testing situation may have raised or lowered trainees' IAT scores. Of course, self-reports are also affected by seemingly subtle testing variations (e.g., Schwarz, 1999). Another consideration is that implicit bias may or may not affect the counseling process. Although implicit bias predicts unintentional forms of interpersonal bias in laboratory studies of the general population (e.g., Hugenberg & Bodenhausen, 2003, 2004), counseling trainees may be able to control biased behaviors when working with clients.

Explicit and implicit bias are an important topic to address in multicultural education (Abreu, 2001), but fostering awareness of implicit bias is much more complicated than fostering awareness of its explicit counterpart. Implicitly measuring attitudes by definition does not involve self-report, and that makes the use of traditional scales impossible. Educators interested in developing awareness of implicit bias among their counseling trainees do have several options, however. The first option is to administer pen and paper IATs like those used in the present study, which, even for large groups, can be administered, scored, and interpreted quickly, easily, and inexpensively. However, as will be discussed in the limitations section, error rates are higher with the pen and paper IAT than with other methods. The second option is to construct a computerized IAT. Unfortunately, this requires not only computers but some programming expertise. The third option is to send counselor trainees to the Project Implicit website (https://implicit.harvard.edu) where they can take and receive immediate feedback on any number of demonstration IATs. Although this method requires access to a computer and the Internet, students could be assigned to take the IAT outside of class. It is important to note that interpretations of IAT scores, especially those offered through the Project Implicit website, have been subject to some contention (Blanton & Jaccard, 2006). In addition, the reliability of the IAT compared to self-report prevents its use for diagnostic or decision making purposes; nonetheless, the IAT is a valuable research and educational tool (Nosek et al., 2007). Overall, we would recommend the computerized IAT as a first choice for research and classroom assessment because of its reliability. However, the pen and paper IAT is superior for demonstration due to its simplicity and immediacy.

Limitations

The major limitation of this study was the high proportion of participants who had to be excluded from analyses because they

did not correctly complete the IAT. There were specific types of mistakes that lead to high error rates and removal from analyses. In general, participants with high error rates (a) did not follow directions, (b) engaged in random responding, (c) attempted to use an unallowable response strategy such as skipping items, (d) responded as quickly as possible at the cost of accuracy, or (e) failed to complete a page of the IAT. While the pen and paper IAT appears to have reliability and validity comparable to the computerized version (Boysen, 2005), the number of critical trials completed on the typical pen and paper IAT is less than one fifth of those completed on the computerized version, making mistakes proportionally more detrimental and removal of participants more frequent.

Another limitation was the sample of counselor trainees used in the study. Although four different universities of two distinct sizes were used, they were all located in the Midwest. As a result, the sample was not very diverse in terms of ethnicity and may not have had access to the same diversity experiences as would be available in training programs in other parts of the country. Also, differences may exist between the multicultural courses and counseling programs; therefore, the counselor trainees may not have received the same intervention. Finally, the quasi-experimental nature of the study should be noted. Random assignment into various levels of training was not possible, and the groups may have differed on variables other than training level that were not measured as part of this study. It is possible that cohort, history, or selection effects may have accounted for the results of this study; assumptions about causality should be taken with great caution due to the cross-sectional design.

Conclusion

Counselor trainees are made aware of their values and biases through the process of multicultural training. However, these conscious attitudes toward diversity may not reflect subtle, hard-to-control bias. The counselor trainees in the present study illustrated that the presence of implicit bias and a strong belief in personal competency when working with diverse clients can occur together. Counseling educators and researchers should recognize the potential of assessing attitudes without self-report and further integrate the concept of implicitly measured attitudes into their work.

References

Abreu, J. M. (1999). Conscious and nonconscious African American stereotypes: Impact on first impression and diagnostic ratings by therapists. *Journal of Consulting and Clinical Psychology, 67,* 387–393.

Abreu, J. M. (2001). Theory and research on stereotypes and perceptual bias: A didactic resource for multicultural counseling trainers. *The Counseling Psychologist, 29,* 487–512.

American Counseling Association (2005). *ACA code of ethics.* Alexandria, VA: American Counseling Association.

American Psychological Association. (2002). *Guidelines on multicultural education, training, research, practice, and organizational change for psychologists.* Washington DC: American Psychological Association.

Bidell, M. P. (2005). The sexual orientation counselor competency scale: Assessing attitudes, skills, and knowledge of counselors working with lesbian, gay, and bisexual clients. *Counselor Education & Supervision, 44,* 267–279.

Blair, I. V. (2002). The malleability of automatic stereotypes and prejudice. *Personality and Social Psychology Review, 6,* 242–261.

Blanton, H., & Jaccard, J. (2006). Arbitrary metrics in psychology. *American Psychologist, 61,* 27–41.

Bosson, J. K., Swann, W. B., & Pennebaker, J. W. (2000). Stalking the perfect measure of implicit self-esteem: The blind men and the elephant revisited? *Journal of Personality and Social Psychology, 79,* 631–643.

Boysen, G. A. (2005). *Assessment of implicit bias and self-reported multicultural counseling competencies among counseling trainees.* Unpublished doctoral dissertation, Iowa State University.

Brendl, C. M., Markman, A. B., Messner, C. (2001). How do indirect measures of evaluation work? Evaluating the inference of prejudice in the Implicit Association Test. *Journal of Personality and Social Psychology, 81,* 760–773.

Constantine, M. G., & Ladany, N. (2000). Self-reported multicultural counseling competence scales: Their relation to social desirability attitudes and multicultural case conceptualization ability. *Journal of Counseling Psychology, 47,* 155–164.

Cunningham, W. A., Preacher, K. J., & Banaji, M. R. (2001). Implicit attitude measures: Consistency, stability, and convergent validity. *Psychological Science, 12,* 163–170.

D'Andrea, M., Daniels, J., & Heck, R. (1991). Evaluating the impact of multicultural counseling training. *Journal of Counseling and Development, 70,* 143–150.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology, 56,* 5–18.

Diaz-Lazaro, C. M., & Cohen, B. B. (2001). Cross-cultural counseling training. *Journal of Multicultural Counseling and Development, 29,* 41–56.

Dunn, T. W., Smith, T. B., & Montoya, J. A. (2006). Multicultural competency instrumentation: A review and analysis of reliability generalization. *Journal of Counseling and Development, 84,* 471–482.

Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and use. *Annual Review of Psychology, 54,* 297–327.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102,* 4–27.

Greenwald, A. G., & Nosek, B. A. (2001). Health of the Implicit Association Tests at age 3. *Zeitschrift fur Experimentelle Psychologie, 48,* 85–93.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology, 74,* 1464–1480.

Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (in press). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology.*

Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A meta-analysis on the correlation between the Implicit Association Test and explicit self-report measures. *Personality and Social Psychology Bulletin, 31,* 1369–1385.

Hudson, W. W., & Rickets, W. A. (1980). A strategy for the measurement of homophobia. *Journal of Homosexuality, 5,* 357–372.

Hugenberg, K., & Bodenhausen, G. V. (2003). Facing prejudice: Implicit prejudice on the perception of facial threat. *Psychological Science, 14,* 640–643.

Hugenberg, K., & Bodenhausen, G. V. (2004). Ambiguity in social categorization: The role of prejudice and facial affect in race categorization. *Psychological Science, 15,* 342–345.

Kim, B. S. K., Cartwright, B. Y., Asay, P. A., & D'Andrea, M. (2003). A revision of the Multicultural Awareness, Knowledge, and Skills Survey-Counselor Edition. *Measurement and Evaluation in Counseling and Development, 36,* 161–180.

Kim, D. (2003). Voluntary controllability of the implicit association test (IAT). *Social Psychology Quarterly, 66,* 83–96.

Kiselica, M. S., Maben, P., & Locke, D. C. (1999). Do multicultural education and diversity appreciation training reduce prejudice among counseling trainees? *Journal of Mental Health Counseling, 21,* 240–254.

LaFromboise, T. D., Coleman, H. L., & Hernandez, A. (1991). Development and factor structure of the Cross-Cultural Counseling Inventory-Revised. *Professional Psychology: Research and Practice, 22,* 380–388.

Lane, K. A., Banaji, M. R., Nosek, B. A., & Greenwald, A. G. (2007). Understanding and using the Implicit Association Test: IV. What we know (so far) (pp. 59–102). In B. Wittenbrink & N. S. Schwarz (Eds.). *Implicit measures of attitudes: Procedures and controversies.* New York: Guilford Press.

Lemm, K. (2001). *Personal and social motivation to respond without prejudice: Implications for implicit and explicit attitude and behavior.* Unpublished doctoral dissertation, Yale University.

Lowery, B. S., Hardin, C. D., & Sinclair, S. (2001). Social influence effects on automatic racial prejudice. *Journal of Personality and Social Psychology, 81,* 842–855.

McConahay, J. B., Hardee, B. B., & Batts, V. (1981). Has racism declined in America? It depends on who is asking and what is asked. *Journal of Conflict Resolution, 25,* 563–579.

McConnell, A. R., & Leibold, J. M. (2001). Relations among the implicit association test, discriminatory behavior, and explicit measures of racial attitudes. *Journal of Experimental Social Psychology, 37,* 435–442.

Neville, H., Spanierman, L., & Doan, B. (2006). Exploring the association between color-blind racial identity and multicultural counseling competencies. *Cultural Diversity & Ethnic Minority Psychology, 12,* 275–290.

Neville, H. A., Heppner, M. J., Thompson, C. E., Brooks, L., & Baker, C. E. (1996). The impact of multicultural training on White racial identity attitudes and therapy competencies. *Professional Psychology: Research and Practice, 27,* 83–89.

Nosek, B. A. (2005). Moderators of the relationship between implicit and explicit evaluation. *Journal of Experimental Psychology: General, 134,* 565–584.

Nosek, B., Banaji, M. R., & Greenwald, A. G. (2002). Harvesting implicit group attitudes and beliefs from a demonstration web site. *Group Dynamics: Theory, Research, and Practice, 6,* 101–115.

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2007). The Implicit Association Test at age 7: A methodological and conceptual review. (pp. 265–292). In J. A. Bargh (Ed.), *Automatic Processes in Social Thinking and Behavior.* Psychology Press.

Nosek, B., & Lane, K. (1999). *Analyzing the paper-pencil IAT data.* Unpublished laboratory report, Yale University.

Phillips, J. C. (2000). Training issues and considerations. In R. M. Perez, K. A. DeBord, & K. J. Bieschke (Eds.), *Handbook of counseling and psychotherapy with lesbian, gay, and bisexual clients* (pp. 337–358). Washington, DC: American Psychological Association.

Ponterotto, J. G., Rieger, B. P., Barrett, A., Sparks, R., Sanchez, C. M., & Magids, D. (1996). Development and initial validation of the Multicultural Counseling Awareness Scale. In G. Sodowsky, & J. Impara (Eds.), *Multicultural assessment in counseling and clinical psychology* (pp. 247–282). Lincoln, NE: Buros Institute of Mental Measurements.

Pope-Davis, D. B., Liu, W. M., Toporek, R. L., & Brittan-Powell, C. S. (2001). What's missing from multicultural competency research: Review, introspection, and recommendations. *Cultural Diversity & Ethnic Minority Psychology, 7,* 121–138.

Pruett, S. R., & Chan, F. (2006). The development and psychometric validation of the disability attitude implicit attitude scale. *Rehabilitation Psychology, 51,* 202–213.

Rydell, R. J., & McConnell, A. R. (2006). Understanding implicit and explicit attitude change: A systems of reasoning analysis. *Journal of Personality and Social Psychology, 91,* 995–1008.

Schwarz, N. (1999). Self-reports: How the questions shape the answers. *American Psychologist, 54,* 93–105.

Sodowsky, G. R., Taffe, R. C., Gutkin, T. B., & Wise, S. L. (1994). Development of the Multicultural Counseling Inventory: A self-report measure of multicultural competencies. *Journal of Counseling Psychology, 41,* 137–148.

Steffens, M. C., & Buchner, A. (2003). Implicit Association Test: Separating transsituationally stable and variable components of attitudes toward gay men. *Experimental Psychology, 50,* 33–48.

Sue, D. W., & Sue, D. (2003). *Counseling the culturally diverse: Theory and practice.* New York: Wiley.

Wang, V. O. (1998). Curriculum evaluation and assessment of multicultural genetic counselor education. *Journal of Genetic Counseling, 7,* 87–111.

Worthington, R. L., Mobley, M., Franks, R. P., & Tan, J. A. (2000). Multicultural counseling competencies: Verbal content, counselor attributions, and social desirability. *Journal of Counseling Psychology, 47,* 460–468.

Received August 28, 2006
Revision received December 17, 2007
Accepted December 27, 2007 ■