

2006

Assessment of implicit bias and self-reported multicultural counseling competencies among counseling trainees

Guy A. Boysen
Iowa State University

Follow this and additional works at: <https://lib.dr.iastate.edu/rtd>

 Part of the [Bilingual, Multilingual, and Multicultural Education Commons](#), and the [Student Counseling and Personnel Services Commons](#)

Recommended Citation

Boysen, Guy A., "Assessment of implicit bias and self-reported multicultural counseling competencies among counseling trainees " (2006). *Retrospective Theses and Dissertations*. 1801.
<https://lib.dr.iastate.edu/rtd/1801>

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.

**Assessment of implicit bias and self-reported multicultural counseling
competencies among counseling trainees**

by

Guy A. Boysen

**A dissertation submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of
DOCTOR OF PHILOSOPHY**

Major: Psychology

**Program of Study Committee:
David Vogel (Major Professor)
Nancy Evans
Stephanie Madon
Norman A. Scott
Meifen Wei**

Iowa State University

Ames, Iowa

2006

UMI Number: 3217256

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI[®]

UMI Microform 3217256

Copyright 2006 by ProQuest Information and Learning Company.

All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

Graduate College
Iowa State University

This is to certify that the doctoral dissertation of

Guy Adam Boysen

has met the thesis requirements of Iowa State University

Signature was redacted for privacy.

Major Professor

Signature was redacted for privacy.

For the Major Program

TABLE OF CONTENTS	
ABSTRACT	iv
INTRODUCTION	1
LITERATURE REVIEW	10
Implicit Attitudes	10
The Implicit Association Test	11
Multicultural Training and Assessment	17
Literature Review Summary	20
STUDY 1	22
Introduction	22
Method	23
Results	29
Discussion	39
STUDY 2	42
Introduction	42
Method	42
Results and Discussion	46
GENERAL DISCUSSION	57
Relation to Past Research	58
Implications	60
Limitations	62
Future Research	64
REFERENCES	66
APPENDIX A. SAMPLE PEN AND PAPER IAT	75
APPENDIX B. INDEX OF HOMOPHOBIA	82
APPENDIX C. MODERN RACISM SCALE	83
APPENDIX D. SOCIAL DESIRABILITY SCALE	84
APPENDIX E. SELF-MONITORING SCALE	85
APPENDIX F. CROSS-CULTURAL COUNSELING INVENTORY-REVISED	86
APPENDIX G. IRB APPROVAL STATEMENT	88

ABSTRACT

Research has shown that multicultural training increases competency among counselor trainees. However, these results have been obtained using self-report measures affected by social desirability. In contrast, implicit measures of bias have the advantage of being less susceptible to social desirability. Study 1 explored the reliability and validity of a measure of implicit bias. Study 2 examined self-reported multicultural counseling competency and implicit bias toward lesbian and gay men and African Americans at the start and end of a semester in a sample of counselor trainees ($n = 124$) enrolled in multicultural counseling or unrelated comparison courses. Results indicated that multicultural competency increased as a factor of time and that significant implicit bias was present at all assessments. Discovering implicit bias among counselor trainees who have completed multicultural training indicates that the topic of implicit bias should be included in multicultural course work.

INTRODUCTION

A danger in counseling is that bias present in society will negatively affect the treatment of minority clients. Examples of bias in counseling and the mental health field include the history of biased treatment and diagnosis of African Americans (Adebimpe, 1981; Neighbors, Jackson, Campell, & Williams, 1998) and the overrepresentation of African Americans undergoing psychiatric hospitalizations compared to Whites even when considering group socioeconomic differences (Snowden & Cheung, 1990). Unfortunately, the negative effects of bias seem to have remained steady in recent years (Snowden, 2003). In the case of sexual minorities, for example, research shows that counselors with bias toward homosexuality perceive gay and lesbian clients differently than heterosexual clients (Barrett & McWhirter, 2002; Boysen, Vogel, Madon, & Wester, in press), tend to be less comfortable with gay male clients (Crawford, Humflet, Ribordy, Ho, & Vickers, 1991; Hayes & Gelso, 1993; Hayes & Erkis, 2000), and are more likely to refer gay clients (Crawford et al., 1991). As a result, there has been a growing emphasis on providing multicultural training to counselors in the hope of reducing biased attitudes and behaviors. However, assessments of the effectiveness of training in actually reducing bias have been limited in both number and methodology. To improve understanding of the outcomes of multicultural training, the current study will assess implicit bias toward lesbians and gay men and African Americans and self-reported multicultural competency before and after a multicultural counseling course.

Multicultural Training and Outcomes

With the high potential for negative outcomes when counselors' biases influence their practice, the current emphasis of counseling programs on providing multicultural training (American Psychological Association [APA], 2002; Sue & Sue, 2003) comes as no surprise. Sue and Sue (2003) have clearly delineated the three desired outcomes for

multicultural training. First, culturally competent counselors should be aware of any biases they possess that could influence their practice with clients. Second, an increase in knowledge should occur. Counselors must have some understanding of worldviews that are different from their own and the ways in which these worldviews affect the lives of their clients. Third, the ability to use culturally appropriate intervention skills must be developed. These three competencies have gained acceptance in the field of counseling and they have even been implemented into the broader APA training requirements and cultural competencies (APA, 2002). Furthermore, they have formed the structure of most efforts to assess the outcomes of multicultural training.

A small but consistent set of studies has illustrated the effectiveness of multicultural training in improving counseling competencies. For instance, in two of the best designed studies, self-reported multicultural competency (i.e., knowledge, awareness, and skill) was assessed among counselor trainees before and after taking multicultural courses or unrelated comparison courses (D'Andrea, Daniels, & Heck, 1991; Wang, 1989). In both studies counselor trainees undergoing multicultural training had significantly improved self-assessed multicultural competency while the comparison group underwent little change. In addition, counselor trainees in the multicultural course had significantly higher self-assessed multicultural competency after training than those who did not undergo multicultural training. Research using designs with less control have found similar increases in self-reported knowledge, awareness, and skill after multicultural training (Diaz-Lazaro & Cohen, 2001; Manese, Wu, & Nepomuceno, 2001; Neville, Heppner, Thompson, Brooks, & Baker, 1996; Tomlinson-Clarck, 2000). With the exception of one study that found time alone lead to significant improvements (Brooks & Kahn, 1990), assessments of multicultural training have consistently found increased self-reported competency.

Problems with Multicultural Assessment

Despite the positive results of past attempts to assess the effectiveness of multicultural counselor training, criticism of the research has emerged. On the most basic level, the number of studies using control groups and pre-tests is unfortunately small. Only five studies in the counseling literature have used pre-post designs with control groups to assess multicultural training effectiveness (Brooks & Kahn, 1990; D'Andrea et al., 1991; Parker, Moore, & Neimeyer, 1998; Rudolph, 1989; Wang, 1998). The majority of researchers rely on descriptive, post-intervention designs that preclude drawing strong conclusions from the results.

Another issue of concern is that researchers have not adequately examined the actual reduction of biased attitudes among counselors in training (Kiselica, Maben, & Locke, 1999). Ostensibly, reduction of biases and not just the awareness that they exist is the most important aspect of training. However, despite the need to determine if bias is reduced during training, accepted measures of bias have generally not been used (Kiselica et al., 1999), possibly due to the difficulty in finding scales that are not subject to floor effects when given to a counseling population (Bieschke & Matthews, 1996).

A related problem with previous attempts to assess the impact of multicultural training has been the near total reliance on self-report measures. Over-reliance on self-reports of trainees' multicultural competency is a multifaceted problem. For example, counselor trainees may systematically over- or underestimate their competency. Additionally, those without multicultural knowledge may be unable to accurately assess their competency and those with some knowledge may inflate their assessment in order to appear socially desirable (Pope-Davis, Liu, Toporek, & Brittan-Powell, 2001). Consistent with these concerns, Constantine and Landany (2000) examined the relation among four self-report multicultural competency measures and a measure of social desirability. They

found significant correlations between self-reported multicultural competency and social desirability. Thus, not only is there a concern that trainees may be unable to accurately assess their own competencies, there is actual evidence that some of the multicultural competency measures may be unduly influenced by trainees' conscious or unconscious attempts to present themselves positively.

Fortunately, researchers have recently identified ways to assess biased attitudes while simultaneously reducing the effects of socially desirable responding. Specifically, they have measured not only the explicit components involved in biased attitudes but the implicit components as well. Explicit attitudes are available for conscious evaluation, and therefore can be easily controlled on self-report measures. Explicit attitudes are what have been measured thus far in assessments of the effectiveness of multicultural training among counselors. In contrast, implicit attitudes may occur outside of conscious awareness (Greenwald & Banaji, 1995) and are subtle and difficult to control (Greenwald, McGhee, & Schwartz, 1998). Thus, measuring implicit attitudes provides the potential advantage of being able to identify bias while reducing the ability of an individual to respond in a socially desirable manner (Fazio & Olson, 2003). Therefore, the previously identified problems with assessments of multicultural training can be reduced by measuring the effectiveness of training at not only the explicit level, but the implicit level as well.

Implicit Assessment of Bias

A large body of research in social psychology has illustrated that it is possible to measure attitudes, including biases, which individuals are unable or unwilling to admit consciously (Greenwald & Banaji, 1995). These implicit attitudes "are manifest as actions or judgments that are under the control of automatically activated evaluation, without the performer's awareness of that causation" (Greenwald et al., 1998, p. 1464). Greenwald and colleagues (1998) provided a dramatic illustration of the possible dissociation between

explicit and implicit attitudes. They showed that White participants endorsed positive or neutral attitudes towards African Americans on an explicit measure but exhibited negative attitudes toward African Americans on an implicit measure. Implicit bias has also been found toward Hispanics (Ottoway, Hayden, & Oakes; 2001), Turks (Neumann & Seibt, 2001), elderly people (Nosek, Banaji, & Greenwald, 2002), homosexuality (Banse, Seise, & Zerbes, 2001), and gay men (Steffens & Buchner, 2003). Furthermore, unlike explicit bias, attempts to fake or voluntarily control implicit bias are largely ineffective (Banse et al., 2001; Kim, 2003). Overall, the predominant evidence indicates that implicit bias is subtle, hard to control, and at least partially separate from self-reported explicit bias.

Emerging evidence in social psychology also indicates that implicit bias could affect some of the most essential components of the counseling process. For example, implicit bias has been linked to the basic task of impression formation (Gawronski, Geschke, & Banse, 2003; Hugenberg & Bodenhausen, 2003), can be influenced by observer and actor emotions (Hugenberg & Bodenhausen, 2004; DeSteno, Dasgupta, Bartlett, & Cajdric, 2004), affects subtle forms of behavioral bias such as avoidance (McConnell & Leibold, 2001), and can hamper real life social interactions (Neumann, Hulsenbeck, & Seibt, 2003). If implicit bias were to interfere with just one of these areas in the therapeutic process, multiculturally competent practice would be affected. Therefore, it is necessary to determine the degree to which counseling trainees have implicit bias and the degree to which training affects these biases.

Despite this necessity there has only been one study examining whether counselor trainees possess implicit bias. Abreau (1999) sampled counseling and clinical psychology graduate students, interns, and clinicians. Using a priming procedure, Abreau exposed participants to words stereotypically associated with African Americans before asking them to evaluate vignettes describing an ethnicity-undisclosed or African American client. Results

indicated that priming of the stereotype led to less favorable assessments of hostility for the ethnicity-undisclosed client: an indication of implicit bias. In contrast, when the client's ethnicity was identified as African American rating were more favorable than they had been when the ethnicity was undisclosed: an indication of social desirability. Abreau interpreted the results as indicating that the presence of implicit bias toward African Americans can unconsciously influence judgments of hostility. Yet, this has been the only study to assess implicit bias in a sample of counselor or counselor-trainees and no studies have examined the influence of multicultural training on implicit attitudes.

Multicultural Training and Implicit Bias

Assessing the efficacy of multicultural training using an implicit measure has no direct precedent in the research literature. Therefore, two basic questions must be addressed. First, do counseling trainees possess implicit bias? Second, if implicit bias exists, does multicultural training influence that bias? Some evidence indicates that bias should be expected in counselor trainees. Abreu's (1999) research indicated that implicit bias can be present among counselor trainees and large scale studies have found that implicit bias toward stigmatized groups such as lesbians, gays, and bisexuals and African Americans is typical in the general population (Nosek et al., 2002). However, there is a possibility that the counselor trainees will not be biased. Abreu used a priming measure, and if a different implicit measure was used the same results might not emerge (Olson, & Fazio, 2003). Furthermore, Abreu's study is several years old and may not be representative of current counselor trainees. Evidence from public opinion polls (Yang, 1997) and surveys (Liddle, 1999) indicate that certain biases such as those toward lesbians and gay men, although still present, may be declining. Therefore, while counseling trainees are likely to reflect the larger biases still present in the culture, the assumption must be tested.

Although pervasive and largely unconscious, implicit bias may be affected by training (Blair, 2002). Therefore, the second question that needs to be addressed is whether multicultural training specifically impacts the implicit bias that may be present among counselor trainees. Abreu (2001) has been forceful in emphasizing the importance of educating counselors about the existence and influence of implicit bias as part of multicultural training. While course content may be changing to include the topic of implicit attitudes, the assessment of course outcomes has not followed suit. Nonetheless, some limited evidence does exist that implicit bias can be “unlearned” through multicultural training. One study, for example, examined implicit bias among undergraduate students enrolled in a 15-week seminar on prejudice and conflict (Rudman, Ashmore, & Gary, 2001). Pre-and post-assessments of implicit bias toward African Americans were conducted. Results indicated that White students underwent a significant reduction in implicit bias toward African Americans over the course of a semester compared to students in a research methods course. The relation of these results to counselor training is tenuous, however. The previous study used a course that was intensely focused on topics related to bias while multicultural courses tend to cover a number of topics other than bias (e.g., cultural worldviews, racial identity theory, White privilege, specific cross-cultural counseling skills). Also, about one third of the participants in the undergraduate prejudice and conflict seminar were African American which is a level of diversity not guaranteed most in counseling programs. Furthermore, the prejudice and conflict seminar was an elective course while multicultural training is generally required in counseling programs. Thus, there may have been self-selection effects. The differences between multicultural counseling courses and the course used in Rudman et al.’s study warrant a conceptual replication with counselor trainees. Despite the differences, however, a reduction in implicit bias can be expected in multicultural counseling courses.

The Current Studies

Measuring bias among counselors using only explicit measures is problematic and implicit measures represent a desirable alternative (Burkard, Medler, & Boticki, 2001). So, why has only one published study examined the implicit bias of a sample including counselors and counselor trainees? The reason may be that implicit measures are generally difficult to administer to groups because they typically necessitate the use of computers. For example, the most popular and well studied implicit measure (Fazio & Olson, 2003) is the implicit association test (IAT; Greenwald et al., 1998). The IAT determines the associative strength of concepts by recording participants' reaction times to combinations of attitude consistent and inconsistent images or words. Thus, a computer is needed to collect data and obtaining a large sample of counselors is difficult and impractical. In awareness of this issue, Burkard and colleagues (2001), in the *Handbook of Multicultural Counseling*, stated that: "measures of implicit prejudice that are conceptually and empirically valid and have practical utility are needed" (p. 473). This call may have been answered with the introduction of a pen and paper version of the computerized IAT that can be administered in groups (Lowery, Hardin, & Sinclair, 2001).

The pen and paper IAT maintains the basic logic and procedure of the computerized IAT because it measures the associative strength of concepts using responses to attitude consistent and inconsistent images or words. However, entire multicultural classes could be assessed in one testing session using the pen and paper format. The practical utility of the pen and paper IAT was deemed a benefit to this study leading to its selection as a dependent measure. Unfortunately, as the result of its recent development and the popularity of the computerized IAT, the literature on the pen and paper IAT is limited to one published study (Lowery et al., 2001) and one unpublished dissertation (Lemm, 2001). Therefore, the psychometric properties of the pen and paper IAT are relatively unknown,

and thus before using this measure in the assessment of bias among a group of counselors (Study 2), a study of the reliability and validity of the pen and paper measure was conducted (Study 1)¹. Specifically, Study 1 consisted of IATs measuring bias toward lesbians and gay men and African Americans completed twice by participants with either a 10-minute or a 2-week interval between administrations in order to assess their reliability. To assess the validity of the pen and paper IAT, participants completed measures of explicit bias toward homosexuality and African Americans and measures of social desirability and self-monitoring.

Study 2 consisted of a multi-method assessment of the outcomes of multicultural counseling courses. Counseling trainees were assessed before and after a multicultural course using both explicit and implicit measures. To replicate past research, the explicit measure consisted of a self-report multicultural competency scale (Cross Cultural Counseling Inventory-Revised; LaFromboise, Coleman, & Hernandez, 1991). Implicit bias toward lesbians and gay men and African Americans was assessed using the pen and paper IAT (Lowery et al., 2001). It was hypothesized that counseling trainees enrolled in a multicultural course would (a) exhibit implicit bias at the start of the course, (b) self-report significant increases in multicultural counseling competency after the course, and (c) exhibit a significant reduction in implicit bias. All of the hypotheses are consistent with the previous literature. If supported, these results would indicate that although implicit bias is present among counselor trainees, multicultural training is an effective intervention to attenuate that bias.

¹ Study 1 and Study 2 were conducted with the approval of the Iowa State University IRB (see Appendix G).

LITERATURE REVIEW

The literature review is divided into three sections. The first section introduces the concept of implicit attitudes. The second section reviews topics related to the IAT including construction, validation, applications, and the influence of learning. Finally, the third section examines past efforts to evaluate the outcomes of multicultural training for counselors.

Implicit Attitudes

Attitudes are commonly conceptualized as the thoughts, affects, and behaviors surrounding a specific topic. However, the ability of individuals to consciously access specific attitudes varies. This finding has led to a distinction between attitudes that are explicit and those that are implicit. Explicit attitudes are available for conscious thought, evaluation, and manipulation. The self-report measures that are central to a large proportion of psychological research are based upon explicit attitudes. Unfortunately, individuals are often unable or unwilling to be accurate when self-reporting explicit attitudes. This limitation has long been recognized and researchers have attempted to find methods to eliminate the bias that is seemingly inherent in explicit attitudes (e.g., Sigall & Page, 1971). In contrast, implicit attitudes are “introspectively unidentified (or inaccurately identified)” but nonetheless affect behavior (Greenwald & Banaji, 1995, p. 5). There are many examples of research that illustrate the existence of attitudes individuals are not consciously aware of but that still reliably influence behavior. For example, individuals’ preference for people similar to themselves profoundly affects behavior but this is not an intentional process generally available to conscious introspection (Greenwald & Banaji). Although the influence of processes that are unconscious or implicit has long been recognized, the ability to measure these processes has only recently emerged. The most prominent method used to measure implicit attitudes is the implicit association test (Fazio & Olson, 2003)

The Implicit Association Test

In 1995, Greenwald and Banaji reviewed the literature on implicit attitudes and called for a measure of them that could assess individual differences. The challenge foreshadowed Greenwald and colleagues' (1998) development of the IAT, which has produced a large research literature since its emergence. The following sections will summarize that literature by describing the IAT's construction, validation process, applications, and the influence of learning on IAT results.

Test Construction

The IAT measures the associative strength between two concepts. For example, the IAT has been most commonly used to measure bias toward African Americans by determining the strength of associations between the concepts Good and Bad and the concepts of White people and Black² people. During the process of taking an IAT, participants have the task of categorizing stimuli into two groups. There are two such critical sections in any IAT. In one section stimuli representing the concepts of Good and White are categorized together and the concepts of Bad and Black are categorized together. In the other section stimuli representing the concepts of Good and Black are categorized together and stimuli representing Bad and White are categorized together. In some IATs, names associated with White culture would be categorized with pleasant words such as love while names traditionally associated with Black culture would be categorized with unpleasant words such as maggot and vice versa. Analysis of the IAT is conducted by determining which categorization task was performed more efficiently. Implicit bias, or an IAT effect as it is called, is considered to exist if one of the categorization tasks is significantly more

² IATs measuring bias toward African Americans are referred to as Black IATs because the terms Black and White are used as descriptors for the ethnic groups on the tests, and it is the term utilized in the IAT literature. However, African Americans will still be the term utilized when referring to individuals.

efficient. Efficiency is traditionally measured in time latency of response, number of errors, or items categorized per second. The common finding is that for White Americans it is easier to associate the concept of Bad with Black and Good with White than the reverse.

Validation of the IAT

Greenwald and colleagues (1998) provided three studies in their initial validation of the test that are superb illustrations of the logic and validity of the IAT. The first study simply illustrated the IAT's ability to differentiate between concepts with obvious positive and negative connotations so as to validate the IAT effect. Results showed an IAT effect with implicit bias for flower names versus insect names and musical instruments versus weapons. The second study then took a sample of Japanese and Korean students and used Japanese and Korean names as categories to be associated with positive and negative words. Citing the past cultural difficulties between the Japanese and Korean people as the reason for bias, Japanese individuals were slower when associating Korean names with positive words than when associating Japanese names with positive words and Korean individuals were slower when associating Japanese names with positive words than when associating Korean names with positive words. The second study further validated the IAT by discovering that the amount individuals were imbued in their respective cultures moderated their IAT effect. So, more Americanized Japanese and Korean participants, with ostensibly reduced cultural bias, had significantly smaller IAT effects, and subsequently less implicit bias toward the other group. Finally, the third study looked at Whites' bias toward African Americans. The researchers wanted to determine if individuals who rejected overt bias against African Americans would still show an IAT effect. Indeed, although only one person showed a pro-Black IAT effect, nearly all of the participants explicitly assented to a neutral or positive attitude toward African Americans. Greenwald and colleagues concluded that they had designed a simple procedure using categorization tasks that was able to meet

the challenge to develop an individual differences test of implicit attitudes (Greenwald & Banaji, 1995).

A large amount of research on the validity of the IAT has been published since its introduction. Early concern emerged about the possibility that the IAT was simply measuring familiarity with stimuli (Brendl, Markman, & Messner, 2001); however, by varying the familiarity of the stimuli used in the test, the possibility of a familiarity confound was eliminated (Dasgupta et al., 2000; Ottoway et al., 2000). Furthermore, evidence exists that IAT effects may be produced even when completely novel in-group and out-groups are established. For example, creating groups using the minimal group procedure will lead to measurable implicit bias in the laboratory (Ashburn-Nardo, Voils, & Menteith, 2001). The robustness of the IAT to self-presentation effects is one of its greatest potential advantages (Greenwald et al., 1998). Consistent with this potential advantage, efforts to have participants fake an implicit attitude have been largely unsuccessful (Banse et al., 2001; Kim, 2003). Despite an apparent dissociation between explicit attitudes and implicit attitudes, a source of convergent validity has emerged from the small but significant correlation that IAT results often have with explicit measures of bias (Banse et al., 2001; Steffens & Buchner, 2003). It should be noted, however, this implicit/explicit correlation has not been found universally (Greenwald et al., 1998; Lowery et al., 2001). Convergent validity has also emerged through the creation of a pen and paper IAT that measures the same construct with a method amenable to group testing (Lemm, 2001; Lowery et al., 2001). The ability to administer the pen and paper IAT as a group test makes it attractive despite a limited research literature. More research needs to be done before the pen and paper IAT can be considered to possess the validity of its computerized counterpart and that will be the function of Study 1.

The reliability of the IAT has been less convincingly established than the measures' validity, casting some doubt on claims that the IAT is capable of measuring individual differences. For example, in a study where an IAT measuring bias toward African Americans was administered to participants 4 times with 2 weeks between each test, correlations among the IATs ranged from .16 to .50 (Cunningham, Preacher, & Banaji, 2001). Other test-retest correlations of IATs measuring bias have ranged from .52 to .65 (Banse et al., 2001; Dasgupta & Greenwald, 2001; Steffens & Buchner, 2003). Relatively speaking, however, the reliability of the IAT is a vast improvement over other implicit measures that may have test-retest correlations below .10 and even negative test-retest correlations (Bosson, Swann, & Pennebaker, 2000). This improvement in reliability may account for the popularity of the measure (Nosek et al., in press). Overall, although valid for determining bias at the group level (Steffens & Buchner), the reliability of individual versions of the IAT should be determined before they are used to make diagnostic decisions at the individual level. As such, Study 1 determined the reliability of the specific pen and paper IATs being used to assess multicultural counseling training in Study 2.

Applications of the IAT

The most common use of the IAT has been as a measure of implicit attitudes about an out-group. As mentioned previously, even minimal grouping of participants in the laboratory is sufficient to produce an IAT effect. However, the more useful application is to assess attitudes about social groups that are overtly accepted but about whom implicit bias may remain. For example, in Greenwald and colleagues' (1998) original study, the majority of participants were neutral or pro-Black in their explicit attitudes while only one was pro-Black in his or her implicit attitudes. Other groups that have been examined and found to produce IAT effects include Whites toward Hispanics (Ottoway et al., 2001), Germans toward Turks (Neumann & Seibt, 2001), old and young people toward old people (Nosek et

al., 2002), and heterosexuals toward homosexuals (Banse et al., 2001). The Banse et al. study performed particularly well in establishing the IAT's validity with homosexuality because it sampled populations of both opposite sex oriented and same sex oriented individuals. The heterosexual participants were found to have IAT effects negative toward homosexuality and the same sex oriented individuals were found to have IAT effects positive toward homosexuality. In summary, these results show that the IAT effect has been established in many situations where bias would be expected, but most importantly for the present study, IATs examining implicit bias toward homosexuality and African Americans have been among the most frequently applied.

Another application of the IAT is as a predictor of biased behavior. IAT effects have been implicated as a predictor of both impression formation and interpersonal behavior. Several studies illustrate how implicit bias might affect impression formation. High levels of implicit bias have been associated with perceiving ambiguous behavior in a negative light. In one study, Germans with high levels of implicit bias toward Turks saw behaviors as negative when a Turk engaged in them but neutral when they were enacted by a German (Gawronski, et al., 2003). Impressions of facial stimuli are also related to implicit bias. Using a computer-simulated face shaded to look either White or Black, Whites with high implicit bias toward African Americans tended to perceive anger in the expression of the Black face while a White face making the same expression was perceived in a neutral way (Hugenberg & Bodenhausen, 2003). Similarly, individuals with high scores on an IAT tended to perceive racially ambiguous hostile faces as Black rather than White (Hugenberg & Bodenhausen, 2004). Taken together, these studies indicate that high IAT effects are related to the tendency to systematically misinterpret ambiguous information about a target of bias in a negative manner.

Research also shows that IAT scores are able to predict interpersonal behavior in the laboratory. Individuals with high implicit bias toward people with AIDS were quicker to move a computer mouse away from their body (avoidance behavior) when seeing a picture of a person with AIDS than a picture of a person who did not have AIDS (Neumann et al., 2003). In another study that examined actual interracial interactions, White participants' IAT effects were related to trained observer's judgments of the positiveness of their interaction with an African American experimenter (McConnell & Leibold, 2001). Although the number of studies linking IAT scores with behavior is small, the findings have been confirmatory and fit in with the larger literature indicating the automaticity of behavior (Bargh & Chartrand, 1999).

Learning and Implicit Bias

Automatic bias is often assumed to be a rigid characteristic that is impenetrable to outside influence, but Blair (2002) reviewed the implicit attitude research and concluded that implicit bias is malleable. Implicit bias is influenced by factors such as attention, the stimuli used in the measure, and motivation. More germane to the present study, implicit bias seems to be affected by learning as well. Several studies have illustrated the potential to change individuals' IAT scores through learning experiences. Learning can occur through the formation of new associations by repeatedly pairing concepts that participants normally show bias toward with positive concepts. After repeated pairings, implicit bias is reduced (Dasgupta & Greenwald, 2001; Karpinski & Hilton, 2001) and the effects remain 24 hours later (Dasgupta & Greenwald, 2001). In a more applied setting, Rudman et al. (2001) used the IAT in a pre-post assessment of an undergraduate prejudice and conflict seminar. The researchers found that the course significantly reduced implicit bias among White students while students in a comparison course experienced no change. The current study will replicate and extend this result to counselors undergoing multicultural training.

Multicultural Training Assessment

There have been 13 published attempts to evaluate the effectiveness of multicultural training in reducing bias among counselors and counselor trainees. Following the structure used by Kiselica et al. (1999), this review section is organized by the research methods used in the studies. Four methods have been used: quasi-experimental pre-post control, descriptive pre-post, descriptive post, and qualitative. Racial identity development is not the focus of this review; however, when it is part of the studies included in the review, outcomes are reported.

Quasi-Experimental Pre-Post Control

Five studies have used pre and post measurements with a control group. Unfortunately, no researcher has attempted random assignment into a multicultural training experience; thus, all of the studies are quasi-experimental. Rudolph (1989) examined self-reported bias toward homosexuality and the effectiveness of therapeutic responses to lesbian, gay, and bisexual (LGB) client vignettes among practitioners and trainees before and after a three-day, 20-hour workshop on LGB counseling. Participants in the workshop improved significantly in both areas compared to a group of counseling graduate trainees who did not participate. Brooks and Kahn (1990) measured ethnic awareness by rating participants' impressions of ethnically oriented vignettes before and after a 13-week multicultural course. Self-reported attitudes toward women and traditional sex role behaviors were also assessed. In the only null result found in this review, no changes were found in ethnic awareness, and time alone rather than treatment emerged as a significant predictor of change. In two very similar studies both using the Multicultural Awareness Knowledge & Skills Survey (MAKSS; D'Andrea et al., 1991) as the dependent measure, D'Andrea et al. assessed a cross-cultural counseling course offered across several different formats and Wang (1998) assessed a 10-hour, 10-day multicultural intervention for genetic counselors.

Both studies found significant improvements in the treatment group post-intervention and significant differences between the experimental and control groups. Finally, Parker et al. (1998) examined a 15-week multicultural counseling course using measures of White racial identity and interracial comfort. The experimental group experienced significant change in their racial identity and also was found to have significant increases in interracial comfort when pre-course scores were statistically controlled. In summary, one study found time to be a significant indicator of change; however, two studies implicated training as a significant factor in self-reported change in multicultural counseling competency. Two studies also illustrated that multicultural training is related to changes in self-reported attitudes about specific multicultural groups and one study found change in racial identity.

Pre-Post Descriptive

Four studies examined multicultural interventions pre-post with no comparison group. Paradis (1981) conducted an eight-week training program for a counseling center staff and interns. Participants were asked to answer the question "Who am I?" before and after the training and completed evaluations of the experience. More people identified their gender, ethnicity, and sexuality in response to the "Who am I?" item after the training, and all reported increased self-awareness and growth. Neville, Heppner, Thompson, Brooks, and Baker (1996) examined White racial identity development and MAKSS scores before and one year after a 15-week multicultural counseling course. Some subscales of the racial identity measure showed significant change as did scores on the MAKSS. Diaz-Lazaro and Cohen (2001) also used the MAKSS to assess the effectiveness of their 13-week multicultural course and found significant increases in MAKSS scores. Finally, using a measure similar to the MAKSS, Manese, Wu, and Nepomuceno (2001) combined data from nine years of pre-doctoral interns and found that only the knowledge and awareness subscales of the measure increased 10 months after an 11-week multicultural seminar.

These pre-post descriptive studies provide further evidence of increases in self-reported multicultural competency with three studies using two different measures all indicating significant change. Additionally, one study found changes in racial identity.

Post Descriptive

Three studies have used post-intervention assessment only. Mio (1989) looked at the value of one specific aspect of a multicultural course. Experts rated trainees' final course essays on how enriching the experience was for them. Trainees who had chosen to participate in a program that matched them with a minority student for intercultural activities were rated as having a more enriching experience. In another study, Tomlinson-Clarke (2000) had trainees in a 15-week multicultural counseling course complete a checklist of multicultural competencies and an open-ended evaluation of the course as a whole. Trainees reported that 16 of 22 multicultural competencies were met by the course and that they emerged from the course with new skills and knowledge. Similarly, Burnett, Hamel, and Long (2002) used course evaluations and journal responses to assess change in a 6-week multicultural course. Trainees rated the course highly with regard to helping them examine their own bias and learn about multicultural competencies. To summarize, two of these studies indicated that after a multicultural counseling course, trainees feel they have benefited and one study indicated that experts may be able to recognize benefits from specific aspects of a multicultural course in trainees' work.

Descriptive Qualitative

One study examined a multicultural course using a qualitative method. Heppner and O'Brian (1994) had trainees rate what course activities were helpful and not helpful after each multicultural class over 15 weeks. Although the main purpose of the study was to identify effective and ineffective course content, trainee responses also included the perceived benefits of the course in general. Perceived benefits reported by the trainees

included thinking more about multicultural issues and increases in awareness, openness, interest, and knowledge.

Summary and Evaluation

From the 13 studies assessing multicultural counseling training and its ability to reduce bias, one clear implication has emerged. Multicultural courses lead to increases in various self-reported multicultural competencies that are assumed to be related to low bias. Using different validated measures, four of the best designed studies found significant increases in variables related to multicultural competency after training. In addition, some convergent evidence is provided by the descriptive and qualitative studies where trainees consistently reported increases in awareness, knowledge, and skills which are the factors that make up multicultural competency measures.

Despite the trend indicated by previous assessments of multicultural training, there are limitations to the research. For example, the general failure to use comparison groups reduces confidence in the findings, and one of the well-designed studies found that time was the only significant predictor of change. Perhaps most importantly, the studies are limited by the reliance on self-report measures. The possibility that counseling trainees are unable to accurately rate themselves on multicultural competency is strong. Finally, only one study directly measured bias. Even if trainees improve their knowledge, awareness, and skills, they may still possess biased attitudes that are implicit, and therefore outside of their conscious awareness. Taking all these limitations into consideration, there is a need for well-designed assessments of multicultural training effectiveness that measure bias using methods other than self-report.

Literature Review Summary

There is a potential for bias to negatively influence the mental health services minorities receive from counselors. Concern over this potential has led to the

implementation of multicultural training in counseling programs so trainees can gain awareness of their own bias, knowledge about other cultures, and the therapeutic skills to work cross-culturally. Imbedded in this multicultural counseling competency effort is the notion that bias should be reduced as a result of multicultural training. However, efforts to assess the outcomes of multicultural training, while successful in that counselor trainees self-report increased competency, have failed to use measures of bias that are not influenced by social desirability. Fortunately, psychological research has established methods to assess bias implicitly, and thus, attenuate social desirability influences. The IAT is the most prominent and well-researched implicit measure of bias. In fact, the IAT has been linked to interpersonal perceptions and behaviors that are important to the counseling process. Therefore, by using an IAT measuring bias as an outcome measure for a multicultural counseling course some of limitations of the research can be addressed.

STUDY 1

Research on the pen and paper IAT is limited to one published study (Lowery et al., 2001) and one unpublished dissertation (Lemm, 2001). Thus, there is little data on the psychometric properties of the measure. Temporal stability is especially important in the evaluation of multicultural training due to the interval between assessments. Furthermore, the validity of dependent measures is always a primary research concern. Thus, Study 1 was conducted in order to determine the reliability and validity of the pen and paper IAT.

In Study 1, the reliability of the pen and paper IAT was assessed by having participants complete IATs measuring bias toward (a) lesbians and gay men and (b) African Americans twice with either a 10-minute interval between administrations or a 2-week interval between administrations. Additionally, the frequency with which IAT effects categorized individuals as biased or not biased at both Time 1 and Time 2 was determined. Validity was examined through the intercorrelations of the two versions of the IAT, explicit measures of bias, social desirability and self-monitoring, and the number of close relationships with the targets of bias.

The pen and paper IAT is not expected to perform at the same level psychometrically as the typical explicit measure. The computerized version of the IAT usually has correlations with explicit measures in the .24 to .37 range and its test-retest correlations are low to moderate in size (Nosek et al., in press). There is no reason to believe the pen and paper IAT will perform differently. One reason for the psychometric difference between implicit and explicit measures is that implicit bias can fluctuate significantly based on variations in the testing condition (Blair, 2002). Another reason for the difference is that, as an implicit measure, the IAT excludes the consistency provided by test takers remembering previous responses and repeating them based on that recall. Despite these differences, the validity and reliability of the pen and paper IAT should be supported. It is expected that the IATs will

be positively correlated with each other at Time 1 and Time 2 illustrating reliability. Reliability should also be evident through the ability of the IAT to similarly categorize individuals as biased or nonbiased at the Time 1 and Time 2 administrations. Convergent validity for the IATs should be demonstrated by positive correlations between the two types of IATs, positive correlations with the explicit measures of bias, and negative correlations with the number of close relationships with targets of bias. Evidence for divergent validity is expected from the lack of correlations between the IATs and explicit measures of social desirability and self-monitoring.

Method

Sample

A total of 163 undergraduates at a large midwestern university completed two IATs each and were compensated with credit in their psychology courses in exchange for participation. Participants were 34% male and 65% female. The ethnic breakdown of the sample was 86% White, 6% Asian, 3% African American, 1% Latino/a, 3% Native American, and 1% did not indicate their ethnicity. The average age of the group was 19.3 years with a range of 17 to 26, and they had completed an average of 2.2 semesters of college.

Measures

Implicit measure. Implicit bias was assessed using a pen and paper version of the IAT (Lemm, 2001; Lowery et al., 2001; see Appendix A). The IAT measures the relative association of two concepts. Implicit bias is measured by determining the ease with which individuals can associate concepts such as gay or straight with the concepts of good and bad. The pen and paper IAT consisted of pages with columns of words and pictures that were categorized as belonging to one concept or another. IATs measuring attitudes toward lesbians and gay men and African Americans were utilized. The lesbian and gay man (LG) IAT consisted of pictures of male/female, female/female, and male/male couples. The Black

IAT consisted of pictures of either African American or White faces. Both IATs also included words representing the categories 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. Words were taken from the initial validation of the IAT (Greenwald et al., 1998).

Participants taking the pen and paper IAT are given 20 seconds to categorize the words and pictures by marking a circle on the left or right that corresponds to the different categories. Test takers first go through two practice blocks consisting of categorizing flower and insect names and good and bad words. The critical tasks consist of four pages of materials. In the attitude consistent blocks on the LG IAT the concepts of Straight and Good are categorized together and Gay and Bad are categorized together. In the attitude inconsistent blocks on the LG IAT the combinations are reversed so that Straight and Bad are categorized together and Gay and Good are categorized together. Each block consists of one page of 44 individual stimulus items arranged in 2 columns of 22 items each. The Black IAT follows the same procedure alternately combining White and Good/Black and Bad and White and Bad/Black and Good. The insect/flower IAT practice blocks were always conducted first but the order of the LG and Black IAT materials and consistent and inconsistent blocks were counterbalanced. Implicit bias is said to be present if more items are correctly categorized in the consistent rather than inconsistent blocks.

The validity of the computerized version of the IAT is well established. Illustrations of validity include convergence with related self-reported attitudes, divergence from unrelated self-reported attitudes, prediction of behavior, and robustness to faking (Nosek et al., in press). Reliability of the computerized IAT is not as well established. Measures of implicit attitudes are often found to have low test-retest reliability. For example, in one study using seven implicit measures of self-esteem, test-retest correlations ranged from -.05 to .69 with

only three of the measures having reliability above .5 (Bosson, Swann, & Pennebaker, 2000). In the few studies where it has been examined, test-retest correlations for the computerized IAT have been as low as .16 (Cunningham et al., 2001) and as high as .69 (Bosson et al.). The longest retest interval reported was one year for a computerized anxiety IAT ($r = .47$; Egloff, Schwerdtfeger, & Schmukle, 2005) and the shortest was an immediate retest with an approximate ten-minute interval ($r = .52$, Banse et al., 2001). However, because all of these studies utilized computerized versions of the IAT, little can be inferred from them about the reliability of the pen and paper IAT.

Some limited data on the reliability and validity of the pen and paper IAT does exist (Lemm, 2001; Lowery et al., 2001). Lemm found that two versions of an IAT measuring bias toward homosexuality administered consecutively correlated .47. With regard to validity data, Lemm found that the pen and paper IAT was significantly correlated ($r = .38$) with a measure of explicit bias, but Lowery and colleagues found no correlations above .09 between an IAT measuring bias toward African Americans and measures of explicit bias and motivation to control explicit bias. Generally, computerized IATs measuring bias have small but significant correlations with explicit bias (Banse et al., 2001; Steffens & Buchner, 2003), but there are exceptions (Greenwald et al., 1998), which make the correlations reported by Lemm and Lowery et al. difficult to interpret with regard to validity. However, a clear indicator of the pen and paper IAT's validity is that both studies found significant implicit bias.

Explicit measures. The Index of Homophobia (Hudson & Rickets, 1980; see Appendix B) was used to measure explicit bias toward homosexuality. The Index of Homophobia is a 25-item scale containing statements such as "I would feel comfortable working closely with a male homosexual" that are rated on a 5-point scale from 1 (*strongly agree*) to 5 (*strongly disagree*). The test's authors point to convergent validity in the form of the Index of Homophobia's significant correlation ($r = .53$) with a measure of conservative

sexual attitudes. Subsequent researchers have also demonstrated the validity of the Index of Homophobia by showing its relation to attitudes towards AIDS patients (Young, Henderson, & Marx, 1990), comfort in working with gay men (Crawford et al., 1991), and number of social relationships with gays and lesbians (Barrett & McWhirter, 2002). The Index of Homophobia demonstrated high internal consistency in the past (.90; Hudson & Rickets, 1980) and showed high internal consistency in the current sample (.90).

The Modern Racism Scale (McConahay, Hardee, & Batts, 1981; see Appendix C) measured explicit bias toward African Americans. The Modern Racism Scale was designed to measure subtle bias as opposed to old fashioned, blatant racism that centers on overt dislike and degradation. The Modern Racism Scale is a 17-item measure containing such statements as "If a Black were put in charge of me, I would not mind taking advice and direction from him or her" that are rated on a scale from 1 (*strongly agree*) to 7 (*strongly disagree*). The Modern Racism Scale is one of the most commonly used measures of bias toward African Americans. Scores on the Modern Racism Scale are positively correlated with other subtle measures of bias toward African Americans (Saucier & Miller, 2003), negatively correlated to motivation to be non-prejudiced (Lowery et al., 2001; Plant & Devine, 1998), and has been used to predict divergent reactions to African Americans and Whites (Vescio & Biernat, 1999). The test-retest correlations for the scale were reported by the authors to be .93 (McConahay et al., 1981), and other studies have found test-retest correlations ranging from .76 to .86 (Cunningham et al., 2001). Although previous studies reported internal consistency for the Modern Racism Scale to be as low as .55 (Lowery et al., 2001), the internal consistency was .89 in this sample.

A shortened version of the Marlowe-Crowne Social Desirability Scale (Strahan & Gerbasi, 1972; see Appendix D) was used as a measure of social desirability. Social desirability is the tendency to portray oneself in the best way possible even if it is not

representative of reality. The measure is a 10-item version of the original scale (Crowne & Marlowe, 1960). Items such as "I never hesitate to go out of my way to help someone in trouble" are rated by test takers as true or false. The validity of the shortened version is demonstrated by its correlation with the original longer version (Fischer & Fick, 1993). Kuder-Richardson formula 20 reliability coefficients have ranged from 0.59 to 0.88 for the scale (Strahan & Gerbasi, 1972; Fischer & Fick, 1993). In the current study, K-R 20 reliability fell slightly below the range of previous studies (.51).

A subscale of the Revised Self-Monitoring Scale (Lennox & Wolfe, 1984; see Appendix E) was used to measure self-monitoring. Self-monitoring is the tendency to be sensitive to social cues and to adapt self-presentation based on those cues. Items on the Revised Self-Monitoring Scale can be broken down into 2 subscales: (a) Ability to Modify Self-Presentation and (b) Sensitivity and the Expressive Behavior of Others. Because this study was designed to show that the IAT has advantages over explicit measures that are sometimes related to social desirability, only the Ability to Modify Self-Presentation subscale was used. The Ability to Modify Self-Presentation subscale consists of 7 items. Statements such as "In social situations, I have the ability to alter my behavior if I feel that something else is called for" are rated on a scale from 0 (*strongly agree*) to 5 (*strongly disagree*). Factor analyses have supported a 2-factor interpretation of the scale (O'Cass, 2000). The Ability to Modify Self-Presentation subscale has shown adequate internal consistency in the past ranging from .77 (Lennox & Wolfe) to .86 (O'Cass). In addition, test-retest reliability over a 2-year interval was adequate ($r = .54$; Lennox & Wolfe). Lennox and Wolfe also provided some evidence of validity by showing that the subscale was negatively correlated with social anxiety. Internal consistency in the current sample was .83.

Procedure

After providing their informed consent, all participants were told they would be taking several measures of attitudes toward different social groups. Participants in the 10-minute interval group ($n = 94$) completed the LG and Black IATs. They then completed the Index of Homophobia, Modern Racism Scale, and a demographic survey. Finally, they completed a second LG and Black IAT approximately 10 minutes after completing the first IATs. Implicit measures were always administered first because of the constraints of timing; however, the order of the explicit measures was counterbalanced. In the 2-week interval group ($n = 69$), participants completed the LG and Black IATs and then completed the Index of Homophobia, Modern Racism Scale, social desirability scale, self-monitoring scale, and demographic survey. Participants returned 2 weeks later to complete the LG and Black IATs a second time. Once again, the explicit measures were counterbalanced. For both groups, the second administration of the IAT was followed by a debriefing.

Scoring of the pen and paper IAT is conducted by creating a ratio of items categorized and seconds allowed for categorization (Lowery et al., 2001). First, seconds per categorization were computed by dividing time (20 seconds) by the number of correct items categorized on each page. Seconds per categorization in the incongruent categories was then subtracted from the seconds per categorization in the congruent categories and the resulting score is referred to as the IAT effect. IAT effects were computed separately for the LG IAT and the Black IAT. Implicit bias, or an IAT effect, is said to exist if a significantly greater number of items are categorized in the congruent blocks than the incongruent blocks of the IAT.

Error rates were computed in order to ensure that participants who did not understand the task or who engaged in random response patterns were not included in the analyses. Individuals with 30% or more errors on either IAT were not included in the

analysis. In addition, individuals who completed less than 6 items total on a block were eliminated because such slow responding is more likely to assess explicit attitudes than implicit attitudes. Thirteen participants were removed from the 10-minute interval group and 11 participants were removed from the 2-week interval group.

Analysis

Analyses designed to assess reliability of the pen and paper IAT included examination for practice effects from Time 1 to Time 2, correlations between Time 1 and Time 2, and reliability of IAT categorizations of implicit bias. Analyses designed to assess validity included computation and testing for significance of the IAT effect, correlation of the LG IAT and Black IATs, and correlation of the IATs with the Index of Homophobia, Modern Racism Scale, social desirability, self-monitoring, and number of close relationships with the targets of bias.

Results

Reliability Analyses

Practice effects. First, the means of the Time 1 and Time 2 IATs were examined for practice effects which could indicate instability in the measures if individuals benefited differently from repeated administrations (Steffens & Buchner, 2003). Comparisons of Time 1 and Time 2 IAT effect means using paired samples *t*-tests yielded a significant change in both the LG IAT and the Black IAT effects from Time 1 to Time 2 in the 10-minute interval group, but no significant differences in the 2-week interval group (see Table 1 for *t*-test results). The significant difference in the 10-minute group resulted from a reduction in IAT effects. Practice effects are a known aspect of the IAT (Greenwald et al., 1998); thus, a change in IAT effects with repeated administration is to be expected. Fortunately, only the 10-minute interval group had a significant practice effect. This could mean the IAT effects are more stable over longer retest intervals. Alternatively, it could mean that completing the

Table 1

Repeated Measures t-Test Results for IAT Effects and Test-Retest Differences in the Ten Minute and Two Week Test-Retest Groups.

Comparison	<i>t</i>	<i>p</i>
Ten Minute Interval (<i>n</i> = 81)		
Lesbian and Gay Men IAT		
Time 1 vs Time 2	3.83	.001
Time 1 Congruent Block vs Incongruent Block	12.87	.001
Time 2 Congruent Block vs Incongruent Block	8.14	.001
Black IAT		
Time 1 vs. Time 2	5.28	.001
Time 1 Congruent Block vs Incongruent Block	9.23	.001
Time 2 Congruent Block vs Incongruent Block	10.07	.001
Two Week Interval (<i>n</i> = 58)		
Lesbian and Gay Men IAT		
Time 1 vs Time 2	1.65	.11
Time 1 Congruent Block vs Incongruent Block	7.11	.001
Time 2 Congruent Block vs Incongruent Block	8.89	.001
Black IAT		
Time 1 vs Time 2	.53	.599
Time 1 Congruent Block vs Incongruent Block	4.96	.001
Time 2 Congruent Block vs Incongruent Block	9.76	.001

explicit measures immediately before the IATs reduced scores. In either case, the procedure used in Study 2 resembles the 2-week interval and should attenuate the practice effect.

Reliability of congruent and incongruent IAT blocks. Reliability of the pen and paper IAT was examined by correlating the individual congruent and incongruent blocks of the IAT at Time 1 and Time 2. Results can be seen in Table 2. Several patterns emerged from the analyses. First, test-retest correlations between Time 1 and Time 2 congruent blocks and Time 1 and Time 2 incongruent blocks were all significant. Correlations ranged from .46 to .76. Second, the congruent and incongruent blocks were also significantly and positively correlated. The relation can be explained by the overall ability individuals have in the categorization task. People who are efficient at the task tend to categorize many stimuli on both the congruent and incongruent blocks and the opposite is true for those who are inefficient at the task. Third, examination of the correlation matrices illustrates that the congruent blocks at Time 1 have higher correlations to congruent blocks at Time 2 than to incongruent blocks at Time 1 or 2. Thus, although the blocks of trials are all generally correlated, higher correlations occur between blocks of trials where the same type of categorization is occurring (i.e., both congruent, both incongruent). Overall, the expected positive correlations between individual blocks of the IAT were significant. Furthermore, some of the correlations, while modest in comparison with explicit standards, were actually larger than have been seen in analyses of the reliability of the IAT in other studies (Egloff et al., 2005; Steffens & Buchner, 2003).

Reliability of IAT effects. Correlations between Time 1 and Time 2 LG and Black IAT effects were also analyzed as an indicator of reliability. The IAT effect is a difference score calculated from the congruent and incongruent blocks of trials. Determining the reliability of the IAT effect is necessary because it is the score that will be analyzed to determine if changes occur in implicit bias due to multicultural training. The use of the IAT as an

Table 2

Correlations Between Time 1 and Time 2 IATs and Time 1 and Time 2 Congruent and Incongruent Blocks of IAT Trials for the Ten Minute and Two Week Interval Groups.

	IAT1	IAT2	C Block1	C Block2	I Block1	I Block2
Ten Minute Interval ($n = 81$)						
IAT1		.19	.03	-.10	-.80**	-.26
IAT2	.43**		.04	.07	-.17	-.70**
C Block1	.08	-.01		.76**	.45	.58**
C Block2	-.03	.04	.75**		.40**	.56**
I Block1	-.68**	-.36	.45**	.46**		.50**
I Block2	-.29**	-.69	.48**	.57**	.57**	
Two Week Interval ($n = 58$)						
IAT1		.29*	.29*	-.08	-.64**	-.28*
IAT2	.21		-.10	-.07	-.48**	-.84**
C Block1	.06	-.17		.46**	.37**	.29**
C Block2	-.04	.04	.64**		.38**	.52**
I Block1	-.81**	-.34**	.27*	.22		.59**
I Block2	-.26	-.83**	.40**	.42**	.45**	

Note. The bottom half of the correlation matrices represents the Lesbian and Gay Men IATs and the top half represents the Black IATs. C Block = Items completed in the congruent block of the IAT. I Block = Items completed in the incongruent block of the IAT. Numbers denote the testing occasions. * = $p = .05$. ** = $p = .01$.

individual difference measure relies on the stability of IAT effects across time. The full correlation matrix can be seen in Table 3. In the 10-minute interval group, the Time 1 and Time 2 LG IATs were significantly correlated, $r = .43$, $p < .001$, but the Time 1 and Time 2 Black IAT were not significantly correlated, $r = .19$, $p = .09$. In the 2-week interval group the Time 1 and Time 2 LG IATs were not significantly correlated, $r = .21$, $p = .11$, but the Time 1 and Time 2 Black IATs were significantly correlated, $r = .29$, $p = .03$. Although the correlations were inconsistently significant and small to medium in size, they did fall within the range seen in the computerized version of the IAT (Egloff et al., 2005) and were larger and more consistent than some other implicit measures (Bosson et al., 2000).

Categorization of individuals as biased at Time 1 and Time 2. Finally, the IAT results were examined to determine if individuals were consistently classified as biased or nonbiased from Time 1 to Time 2. While the previous analysis determined the stability of individuals' precise IAT effects across time, the current analysis will determine if the IAT effect categorizes individuals as biased or nonbiased in a stable way across time. In order to conduct these analyses, IAT effects were coded as categorical variables. Positive IAT effects were coded as 1 (biased) and null and negative IAT effects were coded as -1 (not biased). Consistency of coding between Time 1 and Time 2 was then analyzed (see Table 4). Although a kappa statistic would be the typical analysis for such data, the unequal distribution of scores in the cells allowed only descriptive statistics to be utilized. In the 10-minute interval group the LG IAT yielded the same classification for 95% of cases from Time 1 to Time 2 and the Black IAT yielded the same classification in 85% of cases from Time 1 to Time 2. In the 2-week interval group the LG IAT yielded the same classification for 91% of cases from Time 1 to Time 2 and the Black IAT yielded the same classification in 81% of cases from Time 1 to Time 2. In all, approximately 90% and 80% of individuals were categorized as biased during either occasion of the LG and Black IATs, respectively, and

Table 3

Means, Standard Deviations, and Correlations Among Study 1 Test Variables.

	HIAT1	HIAT2	BIAT1	BIAT2	IH	MRS	SD	SM	LGBR	AAR
1		.43**	.08	.42**	.18	.07	.06	.13	-.05	-.01
2	.21		.13	.35**	.30**	.03	-.02	.27*	-.24*	.08
3	.13	.08		.19	.03	.05	.20	-.04	.00	.02
4	.21	.44**	.29*		.22*	.04	.07	.04	-.15	-.03
5	.13	.11	.17	.13		.53**	-.07	-.07	-.50**	-.10
6	.15	-.04	.24	.06	.62**		-.06	.21	-.31**	-.18
7	.13	.19	.07	.06	-.06	-.18		.11	.03	.05
8	.05	.12	.07	-.01	.04	-.23	.09		.21	.11
9	-.26	-.17	-.12	-.04	-.50**	-.29*	-.04	.05		.40**
10	.01	-.16	-.09	-.16	-.35	-.38**	-.13	.20	.39**	
<u>Ten Minute Interval (n = 81)</u>										
<i>M</i>	.69	.40	.65	.28	68.38	46.14	4.90	24.18	1.92	4.22
<i>SD</i>	.76	.28	.64	.25	19.42	15.42	1.76	4.84	2.63	5.81
<u>Two Week Interval (n = 58)</u>										
<i>M</i>	.76	.57	.48	.43	72.10	52.02	5.12	24.36	1.90	2.88
<i>SD</i>	.81	.49	.74	.34	17.48	19.88	2.15	5.30	2.38	3.63

Note. HIAT1 = Lesbian and Gay Man IAT Time 1, HIAT2 = Time 2, BIAT1 = Black IAT Time 1, BIAT2 = Time 2, IH = Index of Homophobia, MRS = Modern Racism Scale, SD = Social Desirability, SM = Self-Monitoring, LGBR = number of close relationships with LGB individuals, AAR = number of close relationships with African Americans. The bottom half of the matrix represents the 2-week interval group, and the top half represents the 10-minute interval group. * = $p = .05$. ** = $p = .01$.

Table 4

Reliability of Categorizations of Participants as Biases or Nonbiased at Time 1 and Time 2 Administrations of the IAT in the Ten Minute and Two Week Interval Groups.

		Ten Minute Interval ($n = 81$)			
		LG IAT		Black IAT	
		Time 2		Time 2	
		-1	1	-1	1
Time 1	-1	1	3	1	3
	1	2	75	9	68

		Two Week Interval ($n = 58$)			
		LG IAT		Black IAT	
		Time 2		Time 2	
		-1	1	-1	1
Time 1	-1	1	4	1	9
	1	1	52	2	49

Note. IAT effects of above 0 are represented with a 1. IAT effects of 0 and below are represented with a -1. LG IAT = Lesbian and Gay Men IAT.

this is consistent with past research (Greenwald et al., 1998; Olson & Fazio, 2003). The high percentage of consistent classifications from Time 1 to Time 2 indicates stability in the classification of people as biased or nonbiased.

Validity Analyses

Convergent and discriminant validity. Construct validity of the pen and paper IAT was assessed first through convergent and divergent validation. Convergent validity was assessed through the correlations between the IATs, explicit measures of bias, self-reported number of close relationships with lesbians, gays, and bisexuals, and self-reported number of close relationships with African Americans. Discriminant validity was assessed through correlations with social desirability and self-monitoring. Significant correlations are presented here, but the full correlation matrix can be seen in Table 3.

In past research implicit bias toward different outgroups has been positively correlated (Gawronski, 2002); thus, it was expected that the LG and Black IATs would show convergent validity through positive correlations with each other. In the 10-minute interval group the Time 1 LG IAT was significantly correlated with Time 2 Black IAT, $r = .42, p < .001$, and the Time 2 LG IAT was also significantly correlated to the Time 2 Black IAT, $r = .35, p = .002$. In the 2-week interval group the Time 1 Black IAT was significantly correlated with the Time 2 LG IAT, $r = .44, p < .001$. These significant and positive correlations between the IATs are indicative of a positive relation between the different forms of implicit bias and support the validity of the measures.

It was also expected that the LG and Black IATs would show convergent validity through positive correlations with explicit measures of bias and negative correlations with the number of close relationships with targets of bias. In the 10-minute interval group the Index of Homophobia and Modern Racism scale were significantly correlated, $r = .53, p < .001$. In addition, the Index of Homophobia was significantly correlated with the Time 2 LG

IAT, $r = .30$, $p = .006$, and the Time 2 Black IAT, $r = .22$, $p = .05$. Finally, the Time 2 LG IAT was significantly correlated with the number of close relationships with LGB individuals, $r = -.24$, $p = .04$. Similarly, in the 2-week interval group the Index of Homophobia and Modern Racism Scale were significantly correlated, $r = .62$, $p < .001$, and the LG IATs were negatively, although not significantly, correlated with the number of self-reported close relationships with LGB individuals at Time 1, $r = -.26$, $p = .05$, and Time 2, $r = -.17$, $p = .20$. The existence of positive correlations between the IATs and correlations in the expected direction with measures related to explicit bias support the validity of the IATs.

Finally, it was expected that the LG and Black IATs would show discriminant validity through their lack of relation to social desirability and self-monitoring. Consistent with expectation, the IATs were not correlated with social desirability and self-monitoring for either the LG or Black IATs, except for one significant correlation between self-monitoring and the 10-minute interval group Time 2 LG IAT, $r = .27$, $p = .04$. Because this represented the only significant correlation and because the other correlations between the IATs and social desirability and self-monitoring were small ($< .20$), the discriminant validity of the IATs seems to be supported. To summarize the convergent and discriminant validity results, the correlations with the LG and Black IATs were all in the expected directions and were generally indicative of the IATs' validity.

IAT effects. Next, validity was assessed by determining if there were significant IAT effects illustrating implicit bias. Significant IAT effects were expected and would provide some indication of validity due to the replication of past research (Banse et al., 2001; Greenwald et al., 1998; Lemm, 2001; Lowery et al., 2001; Steffens & Buchner, 2003). Paired samples *t*-tests indicated that significant differences existed between performance on the congruent and incongruent blocks of the LG and African American IATs at Time 1 and Time 2 for both the 10-minute and 2-week interval groups (see Table 5 for means and

Table 5

Means and Standard Deviations for Congruent and Incongruent Blocks of IAT Trials for the Ten Minute and Two Week Interval Groups.

IAT	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Ten Minute Interval (<i>n</i> = 81)				
	<u>Congruent Block</u>		<u>Incongruent Block</u>	
Time 1 LG IAT	1.02	.27	1.70	.76
Time 2 LG IAT	.83	.21	1.22	.32
Time 1 Black IAT	.98	.23	1.63	.65
Time 2 Black IAT	.82	.20	1.09	.31
Two Week Interval (<i>n</i> = 58)				
Time 1 Homosexuality IAT	1.01	.24	1.77	.82
Time 2 Homosexuality IAT	.87	.17	1.44	.51
Time 1 Black IAT	1.31	.44	1.61	.71
Time 2 Black IAT	.85	.15	1.29	.37

Note. All differences between congruent and incongruent block significant, $p < .001$. LG IAT = Lesbian and Gay Men IAT.

standard deviations and Table 1 for *t*-test results). Computation of the IAT effects illustrated that in every case the significant differences were indicative of bias toward lesbians and gay men and African Americans. These results provide some evidence of validity because they are consistent with past research indicating significant implicit bias toward these two groups.

Study 1 Discussion

Study 1 was designed to assess the validity and reliability of the pen and paper IAT. With regard to validity, significant IAT effects were found that replicated past research on implicit attitudes about homosexuality and gay men (Banse et al., 2001; Lemm, 2001; Steffens & Buchner, 2003) and African Americans (Greenwald et al., 1998; Ottoway et al., 2001). In addition, convergent validity was demonstrated with the positive correlations between the LG and Black IATs, positive correlations among the IATs and explicit measures of bias, and negative correlation between the LG IAT and number of close relationships with LGB individuals. Divergent validity was also demonstrated with the IATs being generally unrelated to social desirability or self-monitoring.

The reliability of the pen and paper IAT was moderately supported by Study 1. Of particular importance was the finding that the majority of individuals received the same classification as biased or nonbiased at Time 1 and Time 2. This finding suggests that the pen and paper IAT can classify people consistently over time. Less clearly supportive of the reliability of the pen and paper IAT were the test-retest estimates, although they were within the range of previous studies of the computerized IAT (Cunningham et al., 2001). The test-retest reliability of implicit measures often struggles to reach levels considered acceptable with explicit measures. For example, the highest test-retest correlation for an IAT measuring bias has been .65 (for a review see Egloff et al., 2005). Therefore, it was not surprising to find the test-retest correlations of the pen and paper IAT to be modest. The question is, why is this so and what does it mean for interpretation of the IAT?

There are two major factors that reduce the stability of IAT effects. First, implicit bias is affected by factors such as the test taker's motivation, strategies to overcome bias, and focus of attention (Blair, 2002). Therefore, variations in the test taker or in the testing situation can lead to variations in IAT effects. Second, because of their implicit nature, IATs largely remove the effects of explicit memory from the testing situation. Therefore, memory of the last testing occasion does not contribute to stability of scores. For example, on explicit measures a test taker can remember that they chose a 6 on 7-point scale the first time they took the test and respond similarly during the second administration. In contrast, that ability to be consistent due to memory of previous responses is completely removed from the IAT.

The implications of the IATs' reliability are not entirely clear. Some researchers are satisfied with the test-retest reliability because the correlations are moderate in size and are maintained over intervals of varying length (Nosek et al., in press). Other researchers are more skeptical and suggest that IAT results should only be interpreted at the group rather than individual level (Steffens & Buchner, 2003). Although the issue is far from resolved, there does appear to be consensus that the IAT can be validly used to make statements about the average bias of a group but making statements about individuals should be done with caution.

It should also be noted that in comparison to the LG IAT and Index of Homophobia, the Black IAT and Modern Racism Scale seem to have performed worse in terms of validity and reliability. Unfortunately, there is no way to fully explain these results in the current study. One possible explanation is the difficulty of the task of categorizing faces on the IAT using only small black and white pictures. However, the performance of the explicit measure was similarly poor giving some indication that assessment of bias toward African Americans itself may have been a factor.

In summary, the IATs seem to be measuring a subtle form of bias toward lesbians and gay men and African Americans. However, the stability shown across time must warrant caution in interpreting change between multiple administrations. Although the IAT may grossly label a group of individuals as biased or unbiased, the ability to measure incremental change and individual differences may be attenuated. Despite these concerns, the similarity of the pen and paper IAT to the computerized IAT in terms of reliability and validity supported its use as the implicit measure in Study 2. Furthermore, measuring bias among counselors using only explicit measures is problematic and, thus, the IAT may be the only alternative (Burkard et al., 2001). As Study 2 would be the first assessment of multicultural training using an IAT, the possible trade off in precision seemed worthwhile in order to be able to measure biased attitudes in a way relatively immune to social desirability.

STUDY 2

Study 2 assessed the outcomes of multicultural counseling courses using implicit and explicit measures. Multicultural competency was assessed using an explicit measure of self-reported multicultural knowledge, awareness, and skill. Implicit bias toward lesbians and gay men and African Americans was assessed using pen and paper versions of the IAT. It was hypothesized that counseling trainees enrolled in a multicultural course would (a) exhibit implicit bias at the start of the course, (b) self-report significant increases in multicultural counseling competency after the course, and (c) exhibit significant change in implicit bias.

Method

Sample

A sample of 124 students enrolled in graduate programs in counseling from four universities participated in the study. All of the universities were in the Midwest. Two of the institutions were large land-grant universities and two were small urban universities. Degree programs offered include masters in psychology, counseling, community counseling, rehabilitation counseling, and school counseling, and doctorates in counseling psychology. The demographic survey was not completed by 5 participants, but of those reporting there were 25 men, 93 women, and 1 person not reporting a sex. The mean age was 22 years with a range from 22 to 51. The ethnic breakdown was 84% White, 9% African American, 3% Hispanic or Lanino/a, 2% Asian American or Pacific Islander, and 2% other. A total of 6 individuals indicated that their sexual orientation was primarily homosexual and 7 reported a bisexual orientation. The highest degree obtained for 78% of participants was a Bachelors, for 16% it was a Masters, and for 2% it was a Doctorate. They averaged 4 semesters completed in their programs ($SD = 4$) and 19 had previously taken a multicultural course at the graduate level.

At Time 1, everyone who was asked to participate did so ($N = 124$). Of these, 105 participants (85%) completed the Time 2 measures. Those who did not participate at Time 2 were absent from class or had dropped the course. Demographic characteristics were analyzed to determine if there were differences between individuals who participated at Time 1 only and those who participated at both Time 1 and Time 2. Independent samples t -tests were conducted to determine if there were differences in age, semesters of training completed, number of practica, number of LGB clients counseled, number of minority clients counseled, number of close relationships with LGB individuals, or number of close relationships with African American individuals. Chi square analyses were conducted to determine if the groups differed in terms of sex or ethnicity. Participants who participated at Time 1 only and those who participated at Time 1 and Time 2 did not differ in terms of age, semesters of training completed, number of practica, number of LGB clients counseled, number of minority clients counseled, number of close relationships with LGB individuals, or number of close relationships with African American individuals (all t 's < 1.08 , all p 's $> .29$). Chi square analyses found no differences in terms of sex, $\chi^2 = 1.54$, $p = .21$, or ethnicity, $\chi^2 = 2.17$, $p = .70$. Thus, there were no apparent differences between individuals who participated at Time 1 only and those who participated at Time 1 and Time 2. In addition, due to high error rates 40 participants were eliminated from the analyses of the LG IAT ($n = 65$) and 30 were eliminated from analyses of the Black IAT ($n = 75$).

Measures

Implicit measure. Pen and paper IATs assessing bias toward lesbians and gay men and African Americans were used in Study 2. These were the same measures used in Study 1. Similar measures of bias toward Homosexuality (Lemm, 2001) and African Americans (Lowery et al., 2001) have also been used in previous research.

Explicit measure. The multicultural counseling competency scale used was the Cross Cultural Counseling Inventory-Revised (CCCI-R; LaFromboise et al., 1991; see Appendix F). The CCCI-R was selected because it assesses the multicultural competencies of awareness, knowledge, and skill in a general way separate from specific course content. Thus, differences in the specific topics covered in the various multicultural courses would not confound the results. In addition, its brevity was seen as an advantage. Although initially designed to be a supervisor rating scale, it has been successfully used as a self-rating scale (Constantine & Ladany, 2000). The CCCI-R requires counselors to rate themselves on 20 items with a scale ranging from 1 (*strongly disagree*) to 6 (*strongly agree*). Typical items include statements such as “I am aware of my cultural heritage,” “I am willing to suggest referral when cultural differences are extensive,” and “I demonstrate knowledge about my clients’ cultures.” Constantine and Ladany (2000) demonstrated the validity of the CCCI-R by showing that it correlated with the MAKSS ($r = .71$; D’Andrea et al., 1991), Multicultural Counseling Inventory ($r = .73$; Sadowsky, Taffe, Gutkin, & Wise, 1994), and Multicultural Counseling Knowledge and Awareness Scale ($r = .63$; Ponterotto, Rieger, Barrett, Sparks, Sanchez, & Magids, 1996). In addition, it has shown adequate internal consistency with scores ranging from .88 to .95 (Constantine & Ladany, 2000; LaFromboise et al., 1991). The internal consistency in the current sample was .89.

Demographics. A demographic and educational survey was completed by participants. Age, sex, ethnicity, and sexual orientation were assessed. Items asking about the number of close relationships participants had with LGB and African American individuals were also included. If participants were African American, they responded to an item about the number of close relationships with Whites. Educational information was collected through items about highest degree attained, semesters of graduate school completed, and other multicultural experiences. Counseling experience was determined by

items about the number of practica completed and the number of LGB and African American clients counseled. Finally, one item asked about previous experience with the IAT.

Design

The design of the study was quasi-experimental pre-post control. Naturally occurring graduate courses with counseling trainees self-selected into them were sampled to form two groups: multicultural and comparison. The multicultural group consisted of counselor trainees enrolled in multicultural competency courses and the comparison group consisted of counseling trainees enrolled in courses unrelated to multicultural competency.

Procedure

Counseling trainees in the multicultural group were solicited from five courses focusing on multicultural counseling competency. Students in the comparison group consisted of students enrolled in courses on research methods, vocational behavior, professional ethics, and cognitive neuroscience. Review of the comparison course syllabi revealed no significant component of any course devoted to multicultural competency issues. The multicultural group consisted of 75 participants and the comparison group consisted of 49 participants.

At both Time 1 (the first week of class) and Time 2 (last week of class) participants first read and signed an informed consent document. Next, they read brief instructions for the pen and paper version of the LG and Black IATs and completed them using the same procedure described in the pilot study. Participants then completed the CCCI-R and demographic survey. Counterbalancing occurred for the LG and Black IATs, consistent and inconsistent blocks within the IATs, and the self-report measures. After all the materials were completed at Time 2, participants were debriefed.

Data Reduction and Analysis

Data reduction for the IAT was conducted as outlined in the pilot study. To test the primary hypotheses, repeated measures analyses of variance (ANOVA) were conducted to identify changes in either the IATs or the CCCI-R among trainees in the multicultural group or comparison group. Exploratory correlational analyses were also conducted to determine the relation of the measures to other educational and demographic characteristics.

Results and Discussion

Comparison of Counseling Trainee Groups

Random assignment was not possible; thus, the first analysis used independent samples *t*-tests to compare the characteristics of individuals in the multicultural and comparison groups who completed measures at both Time 1 and Time 2. Comparisons were made at Time 1 for age, semesters of training completed, number of practica, number of LGB clients counseled, number of minority clients counseled, number of close relationships with LGB individuals, number of close relationships with African American individuals, and CCCI-R scores. Means, standard deviations, and *t* statistics can be found in Table 6. Although the means varied, none of the differences were statistically significant. Thus, the multicultural and comparison groups were roughly equivalent at Time 1, which allows the subsequent comparisons between the groups to be made with more confidence. The hypothesized analyses also depend on the ability to collapse counselor trainees from several sites into one group. An ANOVA was conducted in order to determine if counselor trainees from the four training sites differed in terms of age, semesters of training completed, number of practicum, number of LGB clients counseled, number of minority clients counseled, number of close relationships with LGB individuals, number of close relationships with African American individuals, or CCCI-R scores. Only the number of

Table 6

Demographic Characteristics and IAT Effects of Multicultural and Comparison Groups.

	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>p</i>
	Multicultural			Comparison				
Age	59	28.92	5.89	43	31.68	9.30	1.65	.10
Semesters	53	5.36	5.21	43	3.74	.80	1.51	.14
Practicums	57	1.09	1.34	42	.62	1.54	1.58	.12
LGB Clients	57	1.09	3.09	43	.88	1.66	.42	.67
Minority Clients	56	6.63	18.98	43	3.28	9.21	1.15	.25
LGB Relationships	61	2.28	2.82	42	3.76	6.12	1.47	.15
A A Relationships	53	1.89	2.25	38	3.97	6.39	1.93	.06
CCCI-R Time 1	61	93.76	9.80	43	94.26	9.35	.26	.79
Time 1 H IAT	49	.45	1.01	27	.31	.59	.76	.45
Time 2 H IAT	39	.43	.38	24	.37	.80	.35	.73
Time 1 B IAT	49	.49	.76	27	.37	.80	.63	.54
Time 2 B IAT	39	.36	.64	24	.51	.53	1.03	.31

Note. The demographic means are based on only participants who completed both pre and post measures while the IAT means are based on all participants correctly completing the IATs. Semesters = semesters completed in current program. Practicums = number of practicums completed. LGB Clients = total number of lesbian, gay or bisexual clients counseled. Minority Clients = total number of minority clients counseled. LGB Relationships = number of close relationships with lesbians, gays, or bisexuals. A A Relationships = number of close relationships with African Americans. CCCI-R Time 1 = Cross Cultural Counseling Inventory-Revised pre semester scores. H IAT = Lesbian and Gay Men IAT. B IAT = Black IAT.

practica emerged as significantly different among the groups, $F(3, 98) = 16.64, p < .001$ (all other F 's < 1.59 , all other p 's $> .20$). Post hoc Tukey honest significant difference tests indicated that students at one of the large land-grant universities with a counseling Ph.D. program had significantly more practicum experience than all the other sites (all p 's $< .001$). No other comparisons were significantly different (all p 's $> .09$). The average number of practica for all individuals completing the Time 1 and Time 2 measures was .90 ($SD = 1.66$) and the average number completed at the site that was significantly different was 2.88 ($SD = 2.39$). The mean differences in practica between the significantly different site and the other large land-grant university with a counseling Ph.D. program was 2.04, and the difference between the significantly different site and the two small urban universities without Ph.D. programs were 2.54, and 2.83. The similarity among the groups for all other variables allowed for them to be collapsed into groups for subsequent analyses.

Implicit Measures

Before conducting the analyses using IAT effects, the standard procedure for eliminating invalid tests was implemented. Specifically, error rates were computed in order to eliminate participants who did not understand the task or engaged in random response patterns. Individuals with 30% or more errors or who completed less than 6 items total on a block were eliminated from analysis of the LG or Black IATs. A total of 40 participants were eliminated from analysis of the LG IAT and 30 participants were eliminated from analysis of the Black IAT. As such, $n = 65$ for analyses using the LG IAT, and $n = 75$ in analyses using the Black IAT.

Independent samples t -tests and chi square analyses were conducted to determine if there were differences between individuals who were excluded from the analyses and those who were included in the analyses. Independent samples t -tests were conducted to determine if there were differences in age, semesters of training completed, number of

practica, number of LGB clients counseled, number of minority clients counseled, number of close relationships with LGB individuals, number of close relationships with African American individuals, or CCCI-R scores. Chi square analyses were conducted to determine if the groups differed in terms of sex or ethnicity. One significant result did emerge from the *t*-tests. Those excluded from IAT analyses ($M = 32.78$; $SD = 8.98$) tended to be older than those who were included in IAT analyses ($M = 28.64$; $SD = 6.27$), $t(113) = 2.62$, $p = .01$ (all other *t*'s < 1.94 , all *p*'s $> .06$). Chi square analyses found no differences in terms of sex, $\chi^2 = .00$, $p = .58$, or ethnicity, $\chi^2 = 6.75$, $p = .15$. Thus, the groups were statistically similar for most of the variables of interest indicating a low likelihood of systematic differences between the groups.

IAT effects. The congruent and incongruent blocks of the IAT were compared in order to determine if a significant IAT effect was present at Time 1 and Time 2. If significantly more items are categorized on the congruent block compared to the incongruent block, implicit bias is said to exist. Therefore, in order to determine if there was significant implicit bias in the sample, the number of stimuli correctly categorized on the congruent and incongruent blocks of the LG and Black IATs were analyzed using paired samples *t*-tests. It was hypothesized that significant bias would be present on both the LG and Black IATs at Time 1. Analyses of the LG and Black IATs indicated that significant bias was present at both Time 1 and Time 2 (See Table 7 for means, standard deviations, and *t*-test results). The research hypothesis was confirmed because significant IAT effects were seen in every case, indicating that implicit bias toward lesbians and gay men and African Americans was present at both Time 1 and Time 2.

The effect of multicultural training on IAT effects. According to the research hypothesis, significant reductions in implicit bias should occur among counselor trainees enrolled in a multicultural course. Because random assignment was impossible, some of the

Table 7

Means, Standard Deviations, and t-Test Comparisons of Congruent and Incongruent Blocks of the Lesbian and Gay Men and Black IATs at Time 1 and Time 2 in the Counselor Trainee Sample.

	<i>n</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>p</i> <
		Congruent		Incongruent			
Time 1 LG IAT	84	1.37	.60	1.99	2.13	2.68	.009
Time 2 LG IAT	68	1.11	.36	1.50	.56	7.02	.001
Time 1 Black IAT	94	1.25	.39	1.74	.77	6.00	.001
Time 2 Black IAT	79	1.04	.53	1.41	.56	5.27	.001

Note. LG IAT = Lesbian and Gay Men IAT.

participants had taken a previous multicultural course (e.g., a Ph.D. student took a multicultural course as masters student at another university). Thus, in order to eliminate this potential confound, the effect of having taken a previous multicultural course was controlled for in the analysis. Thus, the hypothesis was first tested by analyzing the LG IAT effects at Time 1 and Time 2 using a 2 (Group: Multicultural vs. Comparison) x 2 (Previous Multicultural Experience: Previous Course vs. No Previous Course) x 2 (Time: Time 1 vs. Time 2) repeated measures ANOVA. Results of the ANOVA for the LG IAT indicated that the main effect of time was not significant, $F(1, 64) = 2.14, p = .15$. However, there was a significant interaction between time and previous multicultural experience, $F(1, 64) = 7.55, p = .008$. Paired samples *t*-tests indicated that the interaction emerged because counselor trainees who had taken a previous multicultural course exhibited a significant reduction in IAT effects, $t(13) = 2.51, p = .032$, from Time 1 ($M = .57, SD = .66$) to Time 2 ($M = .21, SD = .43$), but counselor trainees who had not taken a previous multicultural course did not exhibit a significant change in IAT effects, $t(53) = 1.09, p = .28$, from Time 1 ($M = .33, SD = .69$) to Time 2 ($M = .43, SD = .45$). The interactions between time and group, $F(1, 64) = 2.56, p = .11$, and the three way interaction, $F(1, 64) = .35, p = .56$, were not significant.

Next, the hypothesis was tested by analyzing the Black IAT effects at Time 1 and Time 2 using a 2 (Group: Multicultural vs. Comparison) x 2 (Previous Multicultural Experience: Previous Course vs. No Previous Course) repeated measures ANOVA. Results of the ANOVA for the Black IAT indicated that there was not a significant main effect of time, $F(1, 74) = 1.46, p = .23$. The interactions between time and multicultural group, $F(1, 74) = .73, p = .39$, time and previous multicultural experience, $F(1, 74) = 3.43, p = .07$, and the three way interaction, $F(1, 74) = .51, p = .48$, were also not significant.

In summary, current enrollment in a multicultural course did not lead to significant change in IAT effects. Thus, the hypothesis was not confirmed. However, enrollment in a

multicultural course in a previous semester did lead to a reduction in IAT effects regardless of current multicultural class enrollment. Unfortunately, these results must be interpreted with caution due to the low test-retest reliability found for the pen and paper IAT in Study 1 and the fact that only the LG IAT interaction reached conventional levels of significance.

Relation of IAT effects and demographic variables. Although not part of the main hypotheses, exploratory correlational analyses were conducted. In order to explore the relation between IAT effects and demographic characteristics, the LG and Black IATs, number of LGB clients counseled, number of minority clients counseled, number of close relationships with LGB individuals, and number of close relationships with African American individuals were subject to Pearson correlational analysis. Relations between IAT effects and the categorical variable of sex were also examined using Spearman nonparametric correlations. Correlations were conducted using all participants who correctly completed IATs at Time 1 or Time 2 and can be found in Table 8. Results indicated that the Time 1 and Time 2 LG IATs were significantly correlated, $r = .28, p = .03$, but the Time 1 and Time 2 Black IATs were not correlated, $r = .17, p = .18$. The LG and Black IATs also shared significant correlations, $r = .26, p = .03$. The demographic variables were generally not related to IAT effects. However, the two significant correlations were in the expected direction. The number of close relationships with LGB individuals was negatively correlated with the LG IAT at Time 1, $r = -.31, p = .008$, and the number of minority clients counseled was negatively correlated with the Black IAT at Time 2, $r = -.34, p = .007$. Although tentative, these significant correlations may suggest that low implicit bias is related to long term exposure to the targets of bias.

Explicit Measures

The effect of multicultural training on CCCI-R scores. Consistent with previous research, it was hypothesized that self-reported multicultural competency would significantly

Table 8

Correlations Among IAT Effects, Educational Variables, and Demographic Variables in the Counselor Trainee Sample.

	1	2	3	4	5	6	7	8	9
1 Time 1 LG IAT									
2 Time 2 LG IAT	.28*								
3 Time 1 B IAT	.26*	.16							
4 Time 2 B IAT	.09	.47**	.17						
5 Sex	.23	-.06	.12	.02					
6 LGB Clients	-.09	-.12	-.06	-.09	-.06				
7 Minority Clients	-.02	-.15	-.01	-.34**	-.16	.88**			
8 LGB Relations	-.31**	-.07	-.13	-.07	-.17	-.02	.02		
9 AA Relations	.01	-.04	-.10	-.02	-.10	.58**	.49**	.47**	

Note. Correlations are based counselor trainees who correctly completed the at Time 1 ($n = 76$) and Time 2 ($n = 63$). LG IAT = Lesbian and Gay Men IAT. B IAT = Black IAT. LGB Clients = total number of lesbian, gay or bisexual clients counseled. Minority Clients = total number of minority clients counseled. LGB Relations = number of close relationships with lesbians, gays, or bisexuals. African American Relations = number of close relationships with African Americans. * = $p < .05$. ** = $p < .01$.

increase in the multicultural group but not the comparison group. In order to eliminate a potential confound, the effect of having taken a previous multicultural course was controlled for in the analysis. The effect of multicultural training on self-reported multicultural competency was examined using a 2 (Group: Multicultural vs. Comparison) x 2 (Previous Multicultural Experience: Previous Course vs. No Previous Course) x 2 (Time: Time 1 vs. Time 2) repeated measures ANOVA. Three participants failed to complete a page of the survey leading to the elimination of either their Time 1 or Time 2 scores on the CCCI-R and 3 participants did not answer the item about previous multicultural courses, eliminating them from the analysis. Results of the ANOVA indicated that scores on the CCCI-R were significantly, $F(1, 95) = 6.02, p = .02$, higher at Time 2 ($M = 96.70, SD = 8.33$) than at Time 1 ($M = 93.69, SD = 9.58$). However, the interactions of time and group, $F(1, 95) = .04, p = .84$, time and previous multicultural experience, $F(1, 95) = .40, p = .53$, and the three way interaction, $F(1, 95) = .01, p = .91$, were not significant. These results indicate, contrary to several previous studies, that counselor trainees enrolled in a multicultural course do not necessarily report greater increases in multicultural competency than counselor trainees not enrolled in a multicultural course. Instead, time alone was responsible for the significant increases in CCCI-R scores. Only one other study has found in which time alone predicted change during multicultural training (Brooks & Kahn, 1990). In contrast to the many assessments that have failed to incorporate proper experimental controls, Brooks and Kahn also used a pre-post control design. Thus, replicating the result that time alone predicts increases in multicultural competency using a properly controlled study should be given some consideration for its contribution to the limited research on the topic.

Relation of demographic variables to CCCI-R scores. Although not part of the main hypotheses, exploratory correlational analyses were conducted in order to explore the relation between CCCI-R scores and demographic characteristics. The Time 1 and Time 2

CCCI-R scores, number of LGB clients counseled, number of minority clients counseled, number of close relationships with LGB individuals, and number of close relationships with African American individuals, were subject to a Pearson correlational analysis. Relations between CCCI-R scores and the categorical variable of sex were also examined using Spearman nonparametric correlations. The correlation matrix can be seen in Table 9, but only significant relations to CCCI-R scores will be presented here. Time 1 and Time 2 CCCI-R scores were significantly correlated, $r = .42, p < .001$. The number of LGB clients counseled was correlated with Time 2 CCCI-R scores, $r = .21, p = .04$, and having taken a multicultural course previous to the semester of the assessment was also significantly correlated with Time 2 CCCI-R scores, $r = .30, p = .002$. No other significant correlations emerged with the CCCI-R scores. The correlational analyses indicated that multicultural competency was relatively unrelated to demographic variables but that the long-term effects of multicultural training and practice may have some relation to competency.

Table 9

Correlations Among Cross Cultural Counseling Inventory-R Scores, Educational Variables, and Demographic Variables in the Counselor Trainee Sample.

	1	2	3	4	5	6	7	8	9	10
1 CCCI-R Time 1										
2 CCCI-R Time 2	.42**									
3 Age	.06	.12								
4 Semesters	.00	.09	.07							
5 Practicum	.10	.10	-.08	.25*						
6 LGB Clients	.08	.21*	.25*	-.09	-.01					
7 Minority Clients	.17	.18	.25*	-.06	.06	.75**				
8 LGB Relations	.05	.05	-.01	-.07	.04	.15	-.07			
9 A A Relations	.17	.10	.22*	-.08	-.07	.11	.04	.32**		
10 Sex	.20	.10	.04	-.09	-.07	.04	-.12	-.03	-.09	
<i>M</i>	93.97	96.73	30.10	3.51	.90	1.85	8.67	2.93	5.56	
<i>SD</i>	9.57	8.44	7.73	3.89	1.66	6.47	23.48	4.21	14.80	

Note: The means and correlations are based on counselor trainees who completed both pre and post measures ($n = 104$). CCCI-R Time 1 and Time 2 = Cross Cultural Counseling Inventory-R scores at the start of the semester (Time 1) and end of the semester (Time 2). Semesters = semesters completed in current program. Practicums = number of practicums completed. LGB Clients = total number of lesbian, gay or bisexual clients counseled. Minority Clients = total number of minority clients counseled. LGB Relations = number of close relationships with lesbians, gays, or bisexuals. African American Relations = number of close relationships with African Americans. * = $p = .05$. ** = $p = .01$.

GENERAL DISCUSSION

Study 1 was conducted to validate implicit measures of bias toward lesbians and gay men and African Americans. The results illustrated that pen and paper versions of the IAT (Greenwald et al., 1998) have reliability and validity similar to their computerized counterparts. Although this means that pen and paper IATs are generally valid as measures of implicit bias, their reliability largely invalidates assumptions about individual differences. Despite low reliability, the practical utility of the pen and paper IAT led to the decision to include it as a dependent measure in Study 2.

Study 2 was conducted in order to assess the effect of multicultural training on counselor trainee's implicit bias and self-reported multicultural competency. It was hypothesized that counselor trainees participating in a multicultural course would show greater improvements in self-reported multicultural competency than counselors in comparison courses. In addition, it was hypothesized that implicit bias would be present at Time 1 but that counselors enrolled in a multicultural course would exhibit a significant reduction in implicit bias over the semester. The hypotheses were partially supported by the results. As hypothesized, implicit bias was present among trainees enrolled in a multicultural course and those not enrolled in a multicultural course at the start and at the end of the semester. However, no significant changes in implicit bias occurred due to current enrollment in a multicultural course. In contrast, having taken a previous multicultural course was related to reductions in implicit bias toward lesbians and gay men regardless of current multicultural course enrollment. Contrary to the hypothesis, trainees enrolled in a multicultural course did not have greater improvements in multicultural competency than trainees in other courses. Rather, multicultural competency increased, not only in the group enrolled in a multicultural course during the semester of the assessment, but in the whole sample.

Relation to Past Research

The use of implicit measures to assess bias among counseling trainees is almost entirely novel in the research literature, but comparisons with past work can be made. Broadly speaking, discovering that counselor trainees possess implicit bias toward lesbians and gay men and African Americans was not surprising. Thousands of people have taken IATs measuring bias toward homosexuality and African Americans and implicit bias is typically found (Nosek et al., 2002). The results were also consistent with the one other study that sampled trainees and clinicians and found evidence of implicit bias (Abreu, 1999). Implicit bias toward lesbians and gay men and African Americans is pervasive and counselor trainees seem to be no exception to that rule.

It also appeared that counselor trainees enrolled in a multicultural course do not necessarily exhibit a reduction in implicit bias although past work has found that implicit bias can be modified. Implicit bias is generally malleable depending on the assessment situation (Blair, 2002), and one study has specifically found that implicit bias was reduced by a course related to multicultural issues (Rudman et al., 2001). However, Rudman and colleagues' research was conducted on an undergraduate course with different content, making direct comparisons difficult. In addition, the psychometric properties of Rudman et al.'s implicit measure are unknown and may have been superior to the implicit measures used in Study 2. As will be discussed in greater detail below, the characteristics of the implicit measure in the current study may have prevented a meaningful assessment of attitude change.

Another important finding was that self-reported multicultural competency significantly increased in the whole sample but not specifically among the counselor trainees enrolled in the multicultural course. Finding that enrollment in a multicultural course was unrelated to specific increases in self-reported multicultural competency contradicts much of the past research. Two previous studies used a pre-post control design and measures of

self-reported multicultural knowledge, awareness, and skill to assess multicultural courses (D'Andrea et al., 1991; Wang, 1998), and that makes them most comparable to Study 2. Both found that counselor trainees undergoing multicultural training had improvements that were significantly greater than trainees in other courses. Furthermore, three studies using explicit measures similar to Study 2 and pre-post designs have found that multicultural training leads to significant increases in competence (Diaz-Lazaro & Cohen, 2001; Manese et al., 2001; Neville et al., 1996). In contrast, the current study seems to indicate that, on average, counseling trainees believe they possess high multicultural competency (i.e., average ratings of about 5 on a 6 point scale) and that the competency is increasing (i.e., about 3 points) regardless of the specific training they are undergoing.

Even though the published literature seems consistent, there is some indication that the multicultural competency results of Study 2 are not as contradictory as they seem. First, only two other assessment studies using measures of multicultural knowledge, awareness, and skill also utilized a comparison group (D'Andrea et al., 1991; Wang, 1998). The majority of studies that have found significant improvement in a multicultural training group have not employed a comparison group as a means of experimental control. Second, one previous assessment of multicultural training that did use a pre-post control design also found time to be the only significant predictor of change (Brooks & Kahn, 1990). The authors of the study suggested that being enrolled in a counseling program by itself may change counselor trainees' attitudes. In summary, a total of only five published studies have used designs with control groups that allow for the inference of change, and one of those studies found time to be the only significant predictor of change; thus, the results of the current study should be given consideration in this small research literature.

Implications

Implicit bias. The major implication of this study is the identification of significant implicit bias in a large sample of counselor trainees. Implicit bias can have effects that would make effective counseling difficult. Counselors in training with implicit bias might misinterpret the emotions of their clients (Hugenberg & Bodenhausen, 2003; 2004) or negatively evaluate the behavior of clients (Gawronski et al., 2003). In addition, implicit bias might lead to subtle avoidance behaviors during counseling (Neumann et al., 2003) and could more generally inhibit the therapeutic relationship (McConnell & Leibold, 2001). Although there is evidence that implicit bias can be overcome during interpersonal contact with potential targets of bias (Shelton, Richeson, Salvatore, & Trawalter, 2005), having no implicit bias certainly fits with the ideals of multiculturalism and would better ensure competent treatment of all clients.

One of the goals of multicultural training is awareness, and it is almost certainly the case that the counselor trainees in the present study were unaware of their level of implicit bias. In fact, without testing it may be impossible to accurately self-assess for the presence of implicit bias. Abreu (2001) has argued for the addition of implicit bias into the topics covered during multicultural training. This study adds the empirical basis to his argument by providing evidence of implicit bias among counselor trainees. Many counselors in training may be unaware not only of their own level of implicit bias but that such a concept could be affecting them. By learning about implicit bias during their training, counselor trainees can more fully meet the multicultural ideal of awareness.

Awareness of implicit bias among counselor trainees may be as important as awareness of explicit bias. Counselor trainees tend to have low levels of self-reported bias (Bieschke & Matthews, 1996; Boyesen et al., in press), and this self-assessment of low bias could actually impede multicultural training. Counselor trainees who believe that they

possess little explicit bias might have lower motivation to invest in their multicultural training. Believing that they have little to gain from the course, they might not challenge their beliefs or seek out diverse experiences and this could leave their implicit bias unchanged. In contrast, counselor trainees who know they have a low level of explicit bias but become aware of the potential for implicit bias, or even their own implicit bias, may experience increased motivation to invest in their multicultural training so as to ensure reduction of implicit bias. Although these scenarios are purely speculative at this point, it is probably safe to assume that awareness of implicit bias is important in much the same way as awareness of explicit bias. Therefore, both implicit and explicit bias should be addressed as part of multicultural training curricula.

Multicultural competency. The results showing that time alone accounted for increases in multicultural counseling competency may also have implications for training. At least three possible interpretations of the data exist, and each would have a different meaning for training. First, the lack of a significantly greater change in competency among the participants undergoing multicultural training could indicate ineffectiveness of the training methods. Because so few studies using a control group have been published to counter this claim, this explanation is plausible. However, this is the least likely explanation because of the uniform scores across the four institutions sampled. Although it is likely that ineffective multicultural counseling courses do exist, the chance of sampling four of them seems unlikely.

Second, the scores on the multicultural competency scale could be a result of inaccurate self-assessments (Pope-Davis et al., 2001). Individuals who have not engaged in multicultural training may be unable to accurately assess their level of competency. Thus, scores of trainees in the multicultural group may have been inflated at the start of the semester leading to a ceiling effect. In fact, as training progressed, the self-assessments of

counselor trainees in the multicultural course may have even fluctuated below their pre course estimates as they were exposed to new information and realized their initial overestimation. As such, the change represented among the counselor trainees undergoing multicultural training may not reflect the actual gains they accrued during the semester due to Time 1 overestimations of competency.

Third, the generally high self-assessments of multicultural competency and significant increase in the whole sample may be an indication of the effectiveness of the infusion of multicultural issues at the program level. On an anecdotal level, several participants responded to the item asking about previous multicultural courses by volunteering that, although they had not taken the specific course yet, they received multicultural content in many other courses. The improvement of the whole sample seen in the data, therefore, could be an accurate assessment of a counseling profession answering calls for affirmation of multicultural principles at the organizational level (Phillips, 2000). Each of the three preceding explanations for the trend in multicultural competency scores is plausible. Unfortunately, the data provide no explanation for change or lack of change, and several factors or interactions of factors may account for the results.

Limitations

The greatest limitation of the current study is the relatively poor reliability of the implicit measure. In effect, the reliability found in Study 1 prevents valid interpretation of change scores in Study 2. Although the IAT was designed as a measure of individual differences (Greenwald et al., 1998), some researches have argued that its reliability prevents that use (Steffens & Buchner, 2003). The pen and paper IAT seems to have similar psychometric issues. Furthermore, there are several sources of error variance that enter into the administration of the pen and paper IAT. Some sources of error might be the group testing format, timing, and the lack of actual control over when individuals start and stop

responding. An additional factor was the practice effect, which could be problematic for reliability if individuals benefited from it at different levels (Steffens & Buchner, 2003). Finally, 20 seconds per block may not provide enough time for reliable trends to emerge in responses.

Despite these limitations, an argument can be made that the pen and paper IAT is a valid measure of implicit bias, albeit not one stable enough to allow assumptions about individual differences. Reliability is often characterized as a necessary but not sufficient component of validity. In other words, an unreliable measure cannot be valid. However, the pen and paper IAT does not need to be conceptualized as a measure of individual differences. Rather, it can be conceptualized as an experimental manipulation that identifies bias at the group level. Despite the low reliability it must be recalled that all conditions are counterbalanced and equal in the administration of the IAT; therefore, a cognitive bias among counselor trainees on average is the best explanation for the reduction in categorization efficiency on the inconsistent blocks of trials compared to the consistent blocks of trials. As such, although the interpretation of change over the semester may have been hampered by the reliability of the IAT, the presence of significant implicit bias can still be considered valid.

Another limitation of this study is that a large number of counselor trainees were excluded from the analyses because they did not correctly complete the IAT. Individuals who were excluded tended to be older than those who were included. However, because of the distinction between implicit and explicit attitudes, there is no way to determine if excluded individuals would have possessed significantly different implicit bias on the IAT they erred on than those who were included. Furthermore, there were several specific types of mistakes that lead to high error rates and removal from analyses. In general, participants with high error rates (a) did not understand the task, (b) engaged in random responding, (c)

attempted to use a 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. Unfortunately, the pen and paper IAT is a difficult task leading to high error rates, and there is no way to eliminate the possibility that counselor trainees who had high error rates were somehow different from those with lower rates in terms of implicit attitudes.

The quasi-experimental design of the Study 2 harmed the ability to make assumptions about the outcomes. The multicultural counseling training literature has suffered from the lack of random assignment into training conditions (Kiselica et al., 1999). Individuals in the multicultural and comparison groups typically also differ in age, courses taken, and counseling experience because of the sequence of courses in training programs. Although there were not significant differences found between the groups in this study, equality of groups could not be assured. In addition, although having taken a previous multicultural course was controlled for statistically, an experimental design would have lead to comparisons only of counselor trainees with and without multicultural training, which would be superior.

A final 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 did not have access to the same diversity experiences as would be available in training programs in other parts of the country. The lack of diversity may have limited the potential for gains in competency among the multicultural training groups.

Future Research

The obvious next step for research would be to replicate the current study using a more reliable measure of implicit bias. Although the reliability of other implicit measures (i.e., the computerized IAT) can also be quite low (Cunningham et al., 2001), in some studies

they have reached acceptable levels (Bosson et al., 2000). A measure of implicit bias that allowed the accurate measurement of change would provide a better test of the hypotheses of the current study. A second direction for future research would be to conduct multiple assessments over an extended period of time. Fluctuation in multicultural competency and implicit bias may occur as the semester progresses, and only through repeated testing will these changes be identified. Moreover, if the assessments continued to occur for several semesters after the completion of training, delayed increases in competency might be identified. The interaction between time and having taken a previous multicultural course seen in Study 2 provides some evidence that the effects of multicultural training may take time to be more fully realized. Finally, now that implicit bias has been found among a large sample of counselors in training, the next step might be to determine what effect implicit bias has on the therapeutic relationship. All of the implications of implicit bias for counseling are derived from laboratory studies, and therefore, may not apply to the actual practice of counseling. Counselor trainees may be able to fully control their implicit bias and prevent it from affecting their behavior (Shelton et al., 2005). On the other hand, implicit bias may be doing vast damage to therapeutic alliances in cross-cultural counseling situations. Therefore, future research should measure implicit bias and effectiveness of cross-cultural counseling to determine if there is a relation.

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.
- Adebimpe, V. R. (1981). Overview: White norms and psychiatric diagnosis of black patients. *American Journal of Psychiatry, 138*, 279-285.
- American Psychological Association (2002). Guidelines on multicultural education, training, research, practice, and organizational change for psychologists. Washington DC: American Psychological Association.
- Ashburn-Nardo, L., Voils, C. I., & Monteith, M. J. (2001). Implicit associations as the seed of intergroup bias: How easily do they take root? *Journal of Personality and Social Psychology, 81*, 789-799.
- Banse, R., Seise, J., & Zerbes, N. (2001). Implicit attitudes towards homosexuality: Reliability, validity, and controllability of the IAT. *Zeitschrift für Experimentelle Psychologie, 48*, 15-169.
- Bargh, J. A., & Chartrand, T. L. (1999). The unbearable automaticity of being. *American Psychologist, 54*, 462-479.
- Barrett, K. A. & McWhirter, B. T. (2002). Counselor trainees' perceptions of clients based on client sexual orientation. *Counselor Education & Supervision, 41*, 219-232.
- Bieschke, K. J., & Matthews C. (1996). Career counselor attitudes and behaviors toward gay, lesbian, and bisexual clients. *Journal of Vocational Counseling, 48*, 243-255.

- Blair, I. V. (2002). The malleability of automatic stereotypes and prejudice. *Personality and Social Psychology Review*, 6, 242-261.
- Bosson, J. K., Swann, W. B., & Pennebacker, 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., Vogel, D. L., Madon, S., & Wester, S. R. (in press). Mental health stereotypes about gay men. Manuscript in press at *Sex Roles: A Journal of Research*.
- 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.
- Brooks, G. S., & Kahn, S. E. (1990). Evaluation of a course in gender and cultural issues. *Counselor Education and Supervision*, 30, 66-76.
- Burkard, A. W., Medler, B. R., & Boticki, M. A. (2001). Prejudice and racism: Challenge and progress in measurement. In J. G. Ponterotto, J. M. Casas, L. A. Suzuki, & C. M. Alexander (Eds.), *Handbook of Multicultural Counseling* (2nd ed, pp. 457-481). Thousand Oaks, CA: Sage.
- Burnett, J. A., Hamel, D., & Long, L. L. (2004). Service learning in graduate counseling education: Developing multicultural counseling competency. *Journal of Multicultural Counseling and Development*, 32, 181-191.
- 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.

- Crawford, I., Humflet, G., Ribordy, S. C., Ho, F. C., & Vickers, V. L. (1991). Stigmatization of AIDS patients by mental health professionals. *Professional Psychology: Research and Practice, 22*, 357-361.
- Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology, 24*, 349-354.
- 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.
- Dasgupta, N., & Greenwald, A. G. (2001). On the malleability of automatic attitudes: Combating prejudice with images of admired and disliked individuals. *Journal of Personality and Social Psychology, 81*, 800-814.
- Dasgupta, N., McGhee, D. E., Greenwald, A. G., & Banaji, M. (2000). Automatic preferences for white Americans: Eliminating the familiarity explanation. *Journal of Experimental Social Psychology, 36*, 316-328.
- DeSteno, D., Dasgupta, N., Bartlett, M. Y., & Caidric, A. (2004). Prejudice from thin air: The effect of emotion on automatic intergroup attitudes. *Psychological Science, 15*, 319-324.
- Diaz-Lazaro, C. M., & Cohen, B. B. (2001). Cross-cultural counseling training. *Journal of Multicultural Counseling and Development, 29*, 41-56.
- Egloff, B., Schwerdtfeger, A., & Schmukle, S. C. (2005). Temporal stability of the Implicit Association Test – anxiety. *Journal of Personality Assessment, 84*, 82-88.
- Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and use. *Annual Review of Psychology, 54*, 297-327.

- Fischer, D. G., & Fick, C. (1993). Measuring social desirability: Short forms of the Marlowe–Crowne Social Desirability Scale. *Educational and Psychological Measurement, 53*, 417–424.
- Gawronski, B. (2002). What does the implicit association test measure? A test of the convergent and discriminant validity of prejudice-related IATs. *Experimental Psychology, 49*, 171-180.
- Gawronski, B., Geschke, D., & Banse, R. (2003). Implicit bias in impression formation: Associations influence the construal of individuating information. *European Journal of Social Psychology, 33*, 573-589.
- Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102*, 4-27.
- 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., & Nosek, B. A. (2001). Health of the Implicit Association Tests at age 3. *Zeitschrift fur Experimentelle Psychologie, 48*, 85-93.
- Hayes, J. A., & Erkis, A. J. (2000). Therapist homophobia, client sexual orientation, an source of client HIV infection as predictors of the therapist reactions to clients with HIV. *Journal of Counseling Psychology, 47*, 71-78.
- Hayes, J. A. & Gelso, C. J. (1993). Male counselors discomfort with gay and HIV-infected clients. *Journal of Counseling Psychology, 40*, 86-93.
- Heppner, M. J. & O'Brian, K. M. (1994). Multicultural counselor training: Students' perceptions of helpful and hindering events. *Counselor Education and Supervision, 34*, 4-18.
- Hudson, W. W., & Ricketts, 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.
- Karpinski, A., & Hilton, J. L. (2001). Attitudes and the implicit association test. *Journal of Personality and Social Psychology*, 81, 774-788.
- 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.
- Lemm, K. (2001). Personal and social motivation to respond without prejudice: Implications for Implicit and Explicit Attitude and Behavior. Unpublished doctoral dissertation, Yale University.
- Lennox, R. D., & Wolfe, R. N. (1984). Revision of the Self-Monitoring Scale. *Journal of Personality & Social Psychology*, 46, 1349-1364.
- Liddle, B. J. (1999). Recent improvements in mental health services to lesbian and gay clients. *Journal of Homosexuality*, 37, 127-137.
- 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.

- Manese, J. E., Wu, J. T., & Nepomuceno, C. A. (2001). The effect of training on multicultural counseling competencies: An exploratory study over a ten-year period. *Journal of Multicultural Counseling and Development, 29*, 31-40.
- 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.
- Mio, J. S. (1989). Experiential involvement as an adjunct to teaching cultural sensitivity. *Journal of Multicultural Counseling and Development, 17*, 38-46.
- Neighbors, H. W., Jackson, J. S., Campell, L., & Williams, D. (1989). The influence of racial factors on psychiatric diagnosis: A review and suggestions for research. *Community Mental Health Journal, 25*, 301-311.
- Neumann, R., & Seibt, B. (2001). The structure of prejudice: Associative strength as a determinant of stereotype endorsement. *European Journal of Social Psychology, 31*, 609-620.
- Neumann, R., Hulsenbeck, K., & Seibt, B. (2003). Attitudes towards people with AIDS and avoidance behavior: Automatic and reflective bases of behavior. *Journal of Experimental Social Psychology, 40*, 543-550.
- 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., 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. (in press). The Implicit Association Test at age 7: A methodological and conceptual review. In J. A. Bargh (Ed.), *Automatic Processes in Social Thinking and Behavior*. Psychology Press.
- O'Cass, A. (2000). A psychometric evaluation of the revised version of the Lennox and Wolfe revised self-monitoring scale. *Psychology and Marketing*, 17, 397-419.
- Olson, M. A., & Fazio, R. H. (2003). Relations between implicit measure of prejudice: What are we measuring? *Psychological Science*, 14, 636-639.
- Ottoway, S. A., Hayden, D. C., & Oakes, M. A. (2001). Implicit attitudes and racism: Effects of word familiarity and frequency on the implicit association test. *Social Cognition*, 19, 97-144.
- Paradis, F. E. (1981). Themes in the training of culturally effective psychotherapists. *Counselor Education and Supervision*, 21, 136-151.
- Parker, W. M., Moore, M. A., & Neimeyer, G. J. (1998). Altering White racial identity and interracial comfort through multicultural training. *Journal of Counseling and Development*, 76, 302-310.
- 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.
- Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology*, 75, 811-832.

- 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. R. Sadowsky, & J. C. 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 and Ethnic Minority Psychology, 7*, 121-138.
- Rudman, L. A., Ashmore, R. D., & Gary, M. L. (2001). "Unlearning" automatic biases: The malleability of implicit prejudice and stereotypes. *Journal of Personality and Social Psychology, 81*, 856-868.
- Rudolph, J. (1989). Effects of a workshop on mental health practitioners' attitudes towards homosexuality and counseling effectiveness. *Journal of Counseling and Development, 68*, 81-85.
- Saucier, D. A., & Miller, C. T. (2003) The persuasiveness of racial arguments as a subtle measure of racism. *Personality and Social Psychology Bulletin, 29*, 1303-1315.
- Sigall, H., & Page, R. A. (1971). Current Stereotypes A little fading a little faking. *Journal of Personality and Social Psychology, 16*, 252-258.
- Shelton, J. N., Richeson, J. A., Salvatore, J., & Trawalter, S. (2005). Ironic effects of racial bias during interracial interactions. *Psychological Science, 16*, 397-402.
- Snowden, L. R. (2003). Bias in mental health assessment and intervention: Theory and evidence. *American Journal of Public Health, 93*, 239-243.
- Snowden, L. R. & Cheung, F. K. (1990). Use of inpatient mental health services by members of ethnic minority groups. *American Psychologist, 45*, 347-355.

- 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.
- Strahan, R., & Gerbasi, K. C. (1972). Short homogeneous versions of the Marlowe–Crowne social desirability scale. *Journal of Clinical Psychology, 28*, 191–193.
- Sue, D. W. & Sue, D. (2003) *Counseling the Culturally Diverse: Theory and Practice*. New York: John Wiley & Sons.
- Tomlinson-Clarke, S. (2000). Assessing outcomes in a multicultural training course: A qualitative study. *Counseling Psychology Quarterly, 13*, 221-231.
- Vescio, T. K., & Biernat, M. (1999). When stereotype-based expectancies impair performance: The effect of prejudice, race, and target quality on judgments and perceiver performance. *European Journal of Social Psychology, 29*, 961-969.
- Wang, V. O. (1998). Curriculum evaluation and assessment of multicultural genetic counselor education. *Journal of Genetic Counseling, 7*, 87-111.
- Yang, A. S. (1997). The polls—trends: Attitudes toward homosexuality. *Public Opinion Quarterly, 61*, 477-507.
- Young, M., Henderson, M. M., & Marx, D. (1990). Attitudes of nursing students toward patients with AIDS. *Psychological Reports, 67*, 491-497.

APPENDIX A
SAMPLE PEN AND PAPER IAT

Instructions

We are interested in how people organize concepts mentally.

For this task, your job is to go down each column as quickly as possible and categorize each item by marking the appropriate circle to the left or right. Some of the items will be words and some will be pictures. For example, for any **insect** (fly, mosquito, cricket, gnat, roach), you should mark the circle under the category "insect". For **flowers** (rose, daffodil, sunflower, daisy, tulip), you will mark the circle under the category "flower". You will do the same with words representing **good** (terrific, love, happy, joy, good), and **bad** (vomit, poison, hatred, evil, bad).

You should be going down the list as fast as possible. It is okay to make some mistakes, but try not to make too many. However, if you do make a mistake, ***do not go back and change your answer***, just keep going down the list. Remember, you are being timed.

Also, **don't fill in every circle.** Simply strike a line through the appropriate circle to save time. You will have twenty seconds for each page. You might find some pages more difficult than others and might not get very far. This is common though, so don't be discouraged.

Do you have any questions?

Insect		Flower
Good		Bad
<input type="radio"/>	mosquito	<input type="radio"/>
<input type="radio"/>	hatred	<input type="radio"/>
<input type="radio"/>	cricket	<input type="radio"/>
<input type="radio"/>	love	<input type="radio"/>
<input type="radio"/>	rose	<input type="radio"/>
<input type="radio"/>	love	<input type="radio"/>
<input type="radio"/>	sunflower	<input type="radio"/>
<input type="radio"/>	joy	<input type="radio"/>
<input type="radio"/>	daisy	<input type="radio"/>
<input type="radio"/>	poison	<input type="radio"/>
<input type="radio"/>	sunflower	<input type="radio"/>
<input type="radio"/>	happy	<input type="radio"/>
<input type="radio"/>	rose	<input type="radio"/>
<input type="radio"/>	poison	<input type="radio"/>
<input type="radio"/>	fly	<input type="radio"/>
<input type="radio"/>	evil	<input type="radio"/>
<input type="radio"/>	tulip	<input type="radio"/>
<input type="radio"/>	love	<input type="radio"/>
<input type="radio"/>	daffodil	<input type="radio"/>
<input type="radio"/>	vomit	<input type="radio"/>
<input type="radio"/>	tulip	<input type="radio"/>
<input type="radio"/>	love	<input type="radio"/>
<input type="radio"/>	sunflower	<input type="radio"/>
<input type="radio"/>	happy	<input type="radio"/>
<input type="radio"/>	fly	<input type="radio"/>
<input type="radio"/>	good	<input type="radio"/>
<input type="radio"/>	fly	<input type="radio"/>
<input type="radio"/>	terrific	<input type="radio"/>
<input type="radio"/>	mosquito	<input type="radio"/>
<input type="radio"/>	hatred	<input type="radio"/>
<input type="radio"/>	tulip	<input type="radio"/>
<input type="radio"/>	terrific	<input type="radio"/>

Insect		Flower
Good		Bad
<input type="radio"/>	mosquito	<input type="radio"/>
<input type="radio"/>	good	<input type="radio"/>
<input type="radio"/>	fly	<input type="radio"/>
<input type="radio"/>	happy	<input type="radio"/>
<input type="radio"/>	rose	<input type="radio"/>
<input type="radio"/>	vomit	<input type="radio"/>
<input type="radio"/>	rose	<input type="radio"/>
<input type="radio"/>	love	<input type="radio"/>
<input type="radio"/>	sunflower	<input type="radio"/>
<input type="radio"/>	love	<input type="radio"/>
<input type="radio"/>	gnat	<input type="radio"/>
<input type="radio"/>	hatred	<input type="radio"/>
<input type="radio"/>	daffodil	<input type="radio"/>
<input type="radio"/>	bad	<input type="radio"/>
<input type="radio"/>	daisy	<input type="radio"/>
<input type="radio"/>	vomit	<input type="radio"/>
<input type="radio"/>	fly	<input type="radio"/>
<input type="radio"/>	good	<input type="radio"/>
<input type="radio"/>	fly	<input type="radio"/>
<input type="radio"/>	bad	<input type="radio"/>
<input type="radio"/>	mosquito	<input type="radio"/>
<input type="radio"/>	good	<input type="radio"/>
<input type="radio"/>	rose	<input type="radio"/>
<input type="radio"/>	terrific	<input type="radio"/>
<input type="radio"/>	gnat	<input type="radio"/>
<input type="radio"/>	hatred	<input type="radio"/>
<input type="radio"/>	rose	<input type="radio"/>
<input type="radio"/>	vomit	<input type="radio"/>
<input type="radio"/>	fly	<input type="radio"/>
<input type="radio"/>	good	<input type="radio"/>
<input type="radio"/>	gnat	<input type="radio"/>
<input type="radio"/>	love	<input type="radio"/>

Flower		Insect
Good		Bad
<input type="radio"/>	fly	<input type="radio"/>
<input type="radio"/>	evil	<input type="radio"/>
<input type="radio"/>	daffodil	<input type="radio"/>
<input type="radio"/>	terrific	<input type="radio"/>
<input type="radio"/>	daisy	<input type="radio"/>
<input type="radio"/>	evil	<input type="radio"/>
<input type="radio"/>	mosquito	<input type="radio"/>
<input type="radio"/>	hatred	<input type="radio"/>
<input type="radio"/>	tulip	<input type="radio"/>
<input type="radio"/>	bad	<input type="radio"/>
<input type="radio"/>	sunflower	<input type="radio"/>
<input type="radio"/>	evil	<input type="radio"/>
<input type="radio"/>	mosquito	<input type="radio"/>
<input type="radio"/>	terrific	<input type="radio"/>
<input type="radio"/>	daffodil	<input type="radio"/>
<input type="radio"/>	vomit	<input type="radio"/>
<input type="radio"/>	mosquito	<input type="radio"/>
<input type="radio"/>	love	<input type="radio"/>
<input type="radio"/>	roach	<input type="radio"/>
<input type="radio"/>	joy	<input type="radio"/>
<input type="radio"/>	gnat	<input type="radio"/>
<input type="radio"/>	poison	<input type="radio"/>
<input type="radio"/>	roach	<input type="radio"/>
<input type="radio"/>	joy	<input type="radio"/>
<input type="radio"/>	daffodil	<input type="radio"/>
<input type="radio"/>	bad	<input type="radio"/>
<input type="radio"/>	sunflower	<input type="radio"/>
<input type="radio"/>	love	<input type="radio"/>
<input type="radio"/>	tulip	<input type="radio"/>
<input type="radio"/>	happy	<input type="radio"/>
<input type="radio"/>	fly	<input type="radio"/>
<input type="radio"/>	love	<input type="radio"/>

Flower		Insect
Good		Bad
<input type="radio"/>	gnat	<input type="radio"/>
<input type="radio"/>	love	<input type="radio"/>
<input type="radio"/>	gnat	<input type="radio"/>
<input type="radio"/>	hatred	<input type="radio"/>
<input type="radio"/>	mosquito	<input type="radio"/>
<input type="radio"/>	terrific	<input type="radio"/>
<input type="radio"/>	daisy	<input type="radio"/>
<input type="radio"/>	good	<input type="radio"/>
<input type="radio"/>	cricket	<input type="radio"/>
<input type="radio"/>	good	<input type="radio"/>
<input type="radio"/>	daffodil	<input type="radio"/>
<input type="radio"/>	love	<input type="radio"/>
<input type="radio"/>	daisy	<input type="radio"/>
<input type="radio"/>	evil	<input type="radio"/>
<input type="radio"/>	sunflower	<input type="radio"/>
<input type="radio"/>	hatred	<input type="radio"/>
<input type="radio"/>	sunflower	<input type="radio"/>
<input type="radio"/>	love	<input type="radio"/>
<input type="radio"/>	gnat	<input type="radio"/>
<input type="radio"/>	bad	<input type="radio"/>
<input type="radio"/>	daffodil	<input type="radio"/>
<input type="radio"/>	love	<input type="radio"/>
<input type="radio"/>	tulip	<input type="radio"/>
<input type="radio"/>	poison	<input type="radio"/>
<input type="radio"/>	daffodil	<input type="radio"/>
<input type="radio"/>	vomit	<input type="radio"/>
<input type="radio"/>	fly	<input type="radio"/>
<input type="radio"/>	bad	<input type="radio"/>
<input type="radio"/>	gnat	<input type="radio"/>
<input type="radio"/>	evil	<input type="radio"/>
<input type="radio"/>	fly	<input type="radio"/>
<input type="radio"/>	love	<input type="radio"/>

White Bad		Black Good
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	hatred	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	good	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	love	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	poison	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	terrific	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	hatred	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	joy	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	happy	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	evil	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	vomit	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	happy	<input type="radio"/>

White Bad		Black Good
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	vomit	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	terrific	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	hatred	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	love	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	bad	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	good	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	joy	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	evil	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	happy	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	poison	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	hatred	<input type="radio"/>

White Good		Black Bad
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	love	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	evil	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	poison	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	terrific	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	joy	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	vomit	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	hatred	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	happy	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	good	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	bad	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	poison	<input type="radio"/>

White Good		Black Bad
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	poison	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	good	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	love	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	evil	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	hatred	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	joy	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	vomit	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	bad	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	happy	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	hatred	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	terrific	<input type="radio"/>

Straight Bad	Gay Good	
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	hatred	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	good	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	love	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	poison	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	terrific	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	hatred	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	joy	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	happy	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	evil	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	vomit	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	happy	<input type="radio"/>

Straight Bad	Gay Good	
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	vomit	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	terrific	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	hatred	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	love	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	bad	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	good	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	joy	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	evil	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	happy	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	poison	<input type="radio"/>
<input type="radio"/>		<input type="radio"/>
<input type="radio"/>	hatred	<input type="radio"/>

Straight Good	Gay Bad
<input type="radio"/>	<input type="radio"/>
	<input type="radio"/>
love	<input type="radio"/>
	<input type="radio"/>
evil	<input type="radio"/>
	<input type="radio"/>
poison	<input type="radio"/>
	<input type="radio"/>
terrific	<input type="radio"/>
	<input type="radio"/>
joy	<input type="radio"/>
	<input type="radio"/>
vomit	<input type="radio"/>
	<input type="radio"/>
hatred	<input type="radio"/>
	<input type="radio"/>
happy	<input type="radio"/>
	<input type="radio"/>
good	<input type="radio"/>
	<input type="radio"/>
bad	<input type="radio"/>
	<input type="radio"/>
poison	<input type="radio"/>

Straight Good	Gay Bad
<input type="radio"/>	<input type="radio"/>
	<input type="radio"/>
poison	<input type="radio"/>
	<input type="radio"/>
good	<input type="radio"/>
	<input type="radio"/>
love	<input type="radio"/>
	<input type="radio"/>
evil	<input type="radio"/>
	<input type="radio"/>
hatred	<input type="radio"/>
	<input type="radio"/>
joy	<input type="radio"/>
	<input type="radio"/>
vomit	<input type="radio"/>
	<input type="radio"/>
bad	<input type="radio"/>
	<input type="radio"/>
happy	<input type="radio"/>
	<input type="radio"/>
hatred	<input type="radio"/>
	<input type="radio"/>
terrific	<input type="radio"/>

APPENDIX B

INDEX OF HOMOPHOBIA

This questionnaire is designed to measure the way you feel about working or associating with homosexuals. It is not a test, so there are no right or wrong answers. Answer each item as carefully and accurately as you can by placing a number by each question as follows.

1	2	3	4	5
Strongly agree	Agree	Neither agree or disagree	Disagree	Strongly disagree
1. ___				
I would be comfortable working closely with a male homosexual.				
2. ___				
I would enjoy attending social functions at which homosexuals were present.				
3. ___				
I would feel uncomfortable if I found out that my neighbor was a homosexual.				
4. ___				
If a member of my sex made a sexual advance toward me I would feel angry.				
5. ___				
I would feel comfortable knowing that I was attractive to members of my own sex.				
6. ___				
I would feel uncomfortable being seen in a gay bar.				
7. ___				
I would feel comfortable if a member of my sex made an advance toward me.				
8. ___				
I would be comfortable if I found myself attracted to a member of my sex.				
9. ___				
I would feel disappointed if I learned that my child was a homosexual.				
10. ___				
I would feel nervous being in a group of homosexuals.				
11. ___				
I would feel comfortable knowing that my clergyman was a homosexual.				
12. ___				
I would deny to members of my peer group that I had friends who were homosexual.				
13. ___				
I would feel that I had failed as a parent if I learned that my child was gay.				
14. ___				
If I saw two men holding hands in public I would be disgusted.				
15. ___				
If a member of my own sex made an advance toward me I would be offended.				
16. ___				
I would feel comfortable if I learned that my daughter's teacher was a lesbian.				
17. ___				
I would feel uncomfortable if my spouse or partner was attracted to members of his or her same sex.				
18. ___				
I would like my parents to know that I have gay friends.				
19. ___				
I would feel uncomfortable kissing a friend of the same sex in public.				
20. ___				
I would like to have friends of my sex who were homosexual.				
21. ___				
If a member of my sex made an advance toward me I would wonder if I was a homosexual.				
22. ___				
I would feel comfortable if I learned that that my best friend of the same sex was homosexual.				
23. ___				
If a member of my sex made an advance towards me I would feel flattered.				
24. ___				
I would feel uncomfortable knowing that my son's male teacher was homosexual.				
25. ___				
I would feel comfortable working with a female homosexual.				

APPENDIX C

MODERN RACISM SCALE

Please answer the following questions according to this scale.

Strongly Agree	Mildly Agree	Agree	Neutral	Disagree	Mildly Disagree	Strongly Disagree
1	2	3	4	5	6	7

1. ___ If a black were put in charge of me, I would not mind taking advice and direction from him or her.
2. ___ If I had a chance to introduce black visitors to my friends and neighbors, I would be pleased to do so.
3. ___ I would rather not have blacks live in the same apartment building I live in.
4. ___ I would probably feel somewhat self-conscious dancing with a black in a public place.
5. ___ I would not mind it at all if a black family with about the same income and education as me moved in next door.
6. ___ I think that black people look more similar to each other than white people do.
7. ___ Interracial marriage should be discourages to avoid the 'who-am-I?' confusion which children feel.
8. ___ I get very upset when I hear a white make a prejudicial remark about blacks.
9. ___ I favor open housing laws that allow more racial integration of neighborhoods. _
10. ___ It would bother me if my new roommate was black.
11. ___ It is likely that blacks will bring violence to neighborhoods when they move in.
12. ___ I enjoy a funny racial joke, even if some people might find it offensive.
13. ___ The federal government should take decisive steps to override the injustices blacks suffer at the hands of local authorities.
14. ___ Black and white people are inherently equal.
15. ___ Black people are demanding too much too fast in their push for equal rights.
16. ___ Whites should support blacks in their struggle against discrimination and segregation.
17. ___ Generally, blacks are not as smart as whites.

APPENDIX D

SOCIAL DESIRABILITY SCALE

Listed below are a number of statements concerning personal attitudes and traits. Read each item and decide whether the statement is true (T) or false (F) as it pertains to you personally.

1. ___ I never hesitate to go out of my way to help someone in trouble.
2. ___ I have never intensely disliked anyone.
3. ___ There have been times when I was quite jealous of the good fortune of others.
4. ___ I would never think of letting someone else be punished for my wrong doings.
5. ___ I sometimes feel resentful when I don't get my way.
6. ___ There have been times when I felt like rebelling against people in authority even though I knew they were right.
7. ___ I am always courteous, even to people who are disagreeable.
8. ___ When I don't know something I don't at all mind admitting it.
9. ___ I can remember "playing sick" to get out of something.
10. ___ I am sometimes irritated by people who ask favors of me.

APPENDIX E
SELF-MONITORING SCALE

Please rate the following statements using the scale below.

Strongly Disagree	Disagree	Slightly Disagree	Slightly Agree	Agree	Strongly Agree
0	1	2	3	4	5

1. ___ In social situations, I have the ability to alter my behavior if I feel that something else is called for
2. ___ I have the ability to control the way I come across to people, depending on the impression I wish to give them
3. ___ When I feel that the image I am portraying isn't working, I can readily change it to something that does
4. ___ I have trouble changing my behavior to meet the requirements of any situation I find myself in
5. ___ I have found that I can adjust my behavior to meet the requirements of any situation I find myself in
6. ___ Even when it might be to my advantage, I have difficulty putting up a good front
7. ___ Once I know what the situation calls for, it's easy for me to regulate my actions accordingly.

APPENDIX F

THE CROSS CULTURAL COUNSELING INVENTORY-REVISED

The purpose of this inventory is to measure your perceptions about your own Cross Cultural Counseling Competence. We are interested in your opinion so please make a judgment on the basis of what the statements in this inventory mean to you. In recording your response, please keep the following points in mind:

- a. Please circle the appropriate rating under each statement.
- b. Please circle only one response for each statement.
- c. Be sure you check every scale even though you may feel that you have insufficient data on which to make a judgment—please do not omit any.

Rating Scale:	1 = strongly disagree	2 = disagree	3 = slightly disagree	4 = slightly agree	5 = agree	6 = strongly agree
---------------	-----------------------	--------------	-----------------------	--------------------	-----------	--------------------

1.	I am aware of my own cultural heritage.	1	2	3	4	5	6
2.	I value and respect cultural differences.	1	2	3	4	5	6
3.	I am aware of how my own values might affect clients.	1	2	3	4	5	6
4.	I am comfortable with differences between myself and clients.	1	2	3	4	5	6
5.	I am willing to suggest referral when cultural differences are extensive.	1	2	3	4	5	6
6.	I understand the current socio-political system and its impact on clients.	1	2	3	4	5	6
7.	I demonstrate knowledge about clients' culture.	1	2	3	4	5	6
8.	I have a clear understanding of counseling and therapy process.	1	2	3	4	5	6
9.	I am aware of institutional barriers which might affect clients' circumstances.	1	2	3	4	5	6

Rating Scale:	1 = strongly disagree	2 = disagree	3 = slightly disagree	4 = slightly agree	5 = agree	6 = strongly agree
---------------	-----------------------	--------------	-----------------------	--------------------	-----------	--------------------

10.	I am able to elicit a variety of verbal and non-verbal responses from clients.	1	2	3	4	5	6
11.	I accurately send and receive a variety of verbal and non-verbal messages.	1	2	3	4	5	6
12.	I am able to suggest institutional intervention skills that favor clients.	1	2	3	4	5	6
13.	I send messages that are appropriate to the communication of clients.	1	2	3	4	5	6
14.	I attempt to perceive the presenting problem within the context of clients' cultural experience, values, and/or lifestyle.	1	2	3	4	5	6
15.	I present my own values to clients.	1	2	3	4	5	6
16.	I am at ease talking with cross cultural clients.	1	2	3	4	5	6
17.	I recognize those limits determined by the cultural differences between clients and myself.	1	2	3	4	5	6
18.	I appreciate clients' social status as an ethnic minority.	1	2	3	4	5	6
19.	I am aware of the professional and ethical responsibilities of a counselor.	1	2	3	4	5	6
20.	I acknowledge and am comfortable with cultural differences.	1	2	3	4	5	6

APPENDIX G

IRB APPROVAL STATEMENT

All research conducted as part of this dissertation was approved by the Iowa State University IRB as human subjects projects 03-813 and 04-568.