IMPROVING INTERGROUP RELATIONS THROUGH PROGRAMMED INTERGROUP CONTACT: A MODEL PROCESS AND ITS EFFECTS

A Thesis in

Counseling Psychology

by

M. D’Andre Wilson

© 2006 M. D’Andre Wilson

Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

May 2006
The thesis of Marvin D’Andre Wilson was reviewed and approved* by the following:

Edwin L. Herr  
Distinguished Professor Emeritus of Education  
and Associate Dean Emeritus  
Thesis Advisor  
Co-Chair of Committee

Kathleen J. Bieschke  
Associate Professor of Counseling Psychology  
Co-Chair of Committee

Dennis Heitzmann  
Affiliate Professor of Counselor Education

James B. Stewart  
Professor of Labor Studies and Industrial Relations  
and African and African American Studies

Spencer G. Niles  
Professor of Counselor Education  
Head of the Department of Counselor Education,  
Counseling Psychology, and Rehabilitation Services

*Signatures are on file in the Graduate School.
ABSTRACT

Demographic trends and advances in technology have increased proximity and interdependence of diverse and historically separate races and cultures worldwide, leading to more frequent intergroup contact. Such contact can lead to mutually beneficial relationships, or to misunderstandings, discrimination, and conflict. The available current research on intergroup contact indicates that simply bringing groups into contact with each other is not enough to ensure positive outcomes. The purpose of this research study was to examine one specific programmed intergroup contact effort by looking at four major areas identified in the intergroup contact literature as warranting further attention. Specifically, this study attempted to (1) research the effects of one model programmed intergroup contact intervention program, (2) examine the impact of intergroup contact on both majority and minority participants simultaneously, (3) explore the interactions between majority and minority participants from a phenomenological perspective, and (4) examine the role of empathy in influencing the outcome of the contact experience.

For this study, both quantitative and qualitative data were collected and analyzed. The quantitative portion of this study was a quasi-experimental field study employing a pretest-posttest control group design. Pretest and posttest assessment instruments were identical and included the CoBRAS, M-GUDS, and QDI. There were three experimental conditions. First was the control which included participants that had no exposure to any form or portion of the nine-week intervention program. Participants in the other two conditions were exposed to the same nine-week intervention program with one variation; the assigned readings were experimentally manipulated so that half the participants were assigned readings identified as “factual,” and the other half readings that were identified
as “personal,” and intended to evoke greater levels of empathy. In addition, qualitative data were collected following each weekly session in the form of the Session Evaluation Questionnaire (SEQ) and a critical incident report. Also qualitative data were collected posttest in the form of a course evaluation sheet from all participants, plus semi-structured individual interviews with 12 participants.

Analyses of the pretest/posttest quantitative data included an ANOVA and two MANCOVA using the pretest scores as a covariate. A phenomenological approach to analyses of the qualitative data was employed, including thematic categorization of the weekly critical incident reports. Findings were then compared and reported based on patterns over time, experimental condition, and race. Mean scores on the four scales produced by the SEQ were also compared and reported by patterns over time, condition and race. Finally, quantitative findings were compared to qualitative findings to help gain greater understanding and generate additional hypotheses.

The results from analyses of the quantitative data found no significant differences between any of the conditions (i.e. – control, intervention with factual readings, or intervention with factual readings) at the .05 significance level for any of the seven scores or factor scores produced (i.e. – one M-GUDS score, plus three factor scores each for the CoBRAS and QDI), with the exception of the CoBRAS factor III. Participants in both of the experimental conditions were found to have improved significantly in regard to awareness of blatant racial issues compared to those in the control, as measured by the CoBRAS factor III. There were not significant differences found on any of the measures between participants in the two experimental conditions (i.e. – those exposed to factual versus personal readings).
Unfortunately, data from the posttest interviews were not of sufficient quality to properly analyze (i.e. - the tapes could not be transcribed). However, analyses of the remaining qualitative data revealed five themes: focus on Self, Other, Topic, Group Process, or Structured Exercise. Further, analyses of this data indicated that minority participants may attend to different factors at different times than majority participants. Implications for Theory, research, and practice were all considered and discussed.
TABLE OF CONTENTS

LIST OF FIGURES................................................................. xii
LIST OF TABLES................................................................. xiii
ACKNOWLEDGEMENTS....................................................... xv

CHAPTER ONE: Introduction............................................... 1
Multicultural Education...................................................... 6
The Role of Counseling Psychology...................................... 8
Need for the Study............................................................. 12
Research Questions............................................................ 13

CHAPTER TWO: Literature Review........................................ 15
Development of Prejudice.................................................... 16
Contemporary Prejudice..................................................... 18
Intergroup Contact............................................................ 23
Review of Meta-analysis..................................................... 25
The Role of Empathy........................................................... 31
Contact and Minority Participants........................................ 35
Multicultural Counselor Training.......................................... 37
Summary............................................................................. 41
The Current Study............................................................... 42

CHAPTER THREE: Method.................................................... 44
Participants......................................................................... 44
Experimental groups......................................................... 45
Comparison group............................................................. 47
<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research Question One</td>
<td>68</td>
</tr>
<tr>
<td>Assumptions of ANCOVA and MANCOVA</td>
<td>71</td>
</tr>
<tr>
<td>Research Question Two</td>
<td>86</td>
</tr>
<tr>
<td>Summary of Quantitative Research Results</td>
<td>90</td>
</tr>
<tr>
<td>Summary of Quantitative Research Results Discussion</td>
<td>91</td>
</tr>
<tr>
<td>CHAPTER FIVE: Qualitative &amp; Exploratory Research: Results and Discussion</td>
<td>93</td>
</tr>
<tr>
<td>Research Question Three</td>
<td>94</td>
</tr>
<tr>
<td>Qualitative Techniques, underlying Principles, and Procedures</td>
<td>95</td>
</tr>
<tr>
<td>Category Definitions and Exemplars</td>
<td>98</td>
</tr>
<tr>
<td>Results of Critical Incident Analyses</td>
<td>100</td>
</tr>
<tr>
<td>Critical Incidents over Time</td>
<td>100</td>
</tr>
<tr>
<td>Discussion</td>
<td>104</td>
</tr>
<tr>
<td>Critical Incidents by Condition</td>
<td>107</td>
</tr>
<tr>
<td>Discussion</td>
<td>109</td>
</tr>
<tr>
<td>Critical Incidents by Race</td>
<td>110</td>
</tr>
<tr>
<td>Discussion</td>
<td>112</td>
</tr>
<tr>
<td>Results of the Session Evaluation Questionnaire Analyses</td>
<td>113</td>
</tr>
</tbody>
</table>
APPENDIX I: Beyond This Study: A History and Update of the Dialogues on Race Program……………………... 209
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Critical Incident Reports: by Week</td>
<td>103</td>
</tr>
<tr>
<td>2</td>
<td>Critical Incident Reports: by Condition</td>
<td>108</td>
</tr>
<tr>
<td>3</td>
<td>Critical Incident Reports: by Race</td>
<td>111</td>
</tr>
<tr>
<td>4</td>
<td>Session Evaluation Questionnaire: by Week</td>
<td>115</td>
</tr>
<tr>
<td>5</td>
<td>Session Evaluation Questionnaire: by Condition</td>
<td>118</td>
</tr>
<tr>
<td>6</td>
<td>Session Evaluation Questionnaire: by Race</td>
<td>120</td>
</tr>
<tr>
<td>7</td>
<td>Session Evaluation Questionnaire: (weekly) White Participants</td>
<td>122</td>
</tr>
<tr>
<td>8</td>
<td>Session Evaluation Questionnaire: (weekly) non-White Participants</td>
<td>122</td>
</tr>
<tr>
<td>9</td>
<td>Critical Incident Reports: by Change in CoBRAS Factor 1</td>
<td>127</td>
</tr>
<tr>
<td>10</td>
<td>Critical Incident Reports: by Change in CoBRAS Factor 2</td>
<td>128</td>
</tr>
<tr>
<td>11</td>
<td>Critical Incident Reports: by Change in CoBRAS Factor 3</td>
<td>129</td>
</tr>
<tr>
<td>12</td>
<td>Critical Incident Reports: by Change in M-GUDS</td>
<td>131</td>
</tr>
<tr>
<td>13</td>
<td>Critical Incident Reports: by Change in QDI Factor 1</td>
<td>133</td>
</tr>
<tr>
<td>14</td>
<td>Critical Incident Reports: by Change in QDI Factor 2</td>
<td>134</td>
</tr>
<tr>
<td>15</td>
<td>Critical Incident Reports: by Change in QDI Factor 3</td>
<td>135</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table 1
Age: Means, Median, Standard Deviations, and Ranges for the Comparison Group and the 2 Experimental Conditions

Table 2
Summary of Frequencies of Nominal & Ordinal Variables Totals & by Condition

Table 3
Standard Deviations and Change Scores for the Comparison Group and 2 Experimental Conditions

Table 4
Contrast Results

Table 5
Bivariate Correlations for age, gender, race, and all scales & subscales at pretest

Table 6
Bivariate Correlations for age, gender, race, and all scales & subscales at posttest

Table 7
Bivariate Correlations for change in scores from pretest to posttest by age, gender, race, and all scales & subscales

Table 8
Pretest Means and Standard Deviations by Condition and Race

Table 9
Posttest Questionnaire by Intervention Condition

Table 10
Critical Incident Reports: by Week

Table 11
Critical Incident Reports: by Condition

Table 12
Critical Incident Reports: by Race

Table 13
Session Evaluation Questionnaire: Weekly by Mean Scale Score
Table 14
Session Evaluation Questionnaire: by Condition and Mean Scale Score

Table 15
Session Evaluation Questionnaire: by Race and Mean Scale Score

Table 16
Session Evaluation Questionnaire: Weekly by Race

Table 17
Critical Incident Reports: by Change in CoBRAS Factor 1

Table 18
Critical Incident Reports: by Change in CoBRAS Factor 2

Table 19
Critical Incident Reports: by Change in CoBRAS Factor 3

Table 20
Critical Incident Reports: by Change in M-GUDS

Table 21
Critical Incident Reports: by Change in QDI Factor 1

Table 22
Critical Incident Reports: by Change in QDI Factor 2

Table 23
Critical Incident Reports: by Change in QDI Factor 3
ACKNOWLEDGEMENTS

There are a great number of individuals to whom I am indebted for the various ways in which they have helped me through my dissertation process. First, I would like to thank Dr. Edwin L. Herr, Distinguished Professor Emeritus of Education and Associate Dean Emeritus. As my dissertation advisor and chair of my committee, Dr. Herr always went above and beyond in his support of my research and writing, and he continued to provide unfailing support even after his official retirement. I sincerely thank him for seeing this project through to its conclusion, but even more so for his patience, kindness, and emotional support through years of academic, professional, and personal challenges.

I would also like to thank the members of my committee: Dr. Kathy Bieschke, Associate Professor of Counseling Psychology, Dr. Dennis Heitzman, Affiliate Professor of Counseling Psychology, and Dr. James Stewart, Professor of Labor Studies and Industrial Relations and African and African American Studies. Each reader offered comments and suggestions that helped to improve my final draft, and as a committee, their combined insights and support carried me through the process.

I would also like to remember those who served on my committee along the way, but due to retirement or changes in position were unable to see this project through to completion; Dr. Donald Keat, Dr. Edmond Marks, and Dr. Beverly Vandiver. I thank each of them for the important roles they served.

Additionally, I would like to thank Dr. Chloe House and Dr. Dave Myers for all of the time and commitment that went into serving as peer raters on my qualitative analyses. Their services went above and beyond, and sincerely thank them.
Finally, I would like to personally thank the people in my life who have loved and supported me long before this particular journey began, and who I know will continue to love and support me on each step of the journeys to come. First, I would like to thank my wife, Kelli, for always trying to be there with me and for me, even when I didn’t know how to ask or even allow her to be. I would also like to thank my sons, Jadon and Tobin, who had no choice but to sacrifice in order for me to complete this task. I thank them for reminding me what is truly important in life, and I promise to spend the rest of my life trying to repay them. In addition, I want to thank my parents, Marvin and Grace Wilson. They always emphasized the importance of education and responsibility. Their sacrifices have allowed me opportunities they only dreamed of, and they truly serve as my inspiration. I regret that my mother was unable to physically see this process completed, but I know that she will always be with me in spirit. I also want to thank my sisters, Tadia and Illia for all of their love and support, plus my niece, Kim for her help with the tedious process of proofing and reformatting at the eleventh hour.

Finally, I thank God, for I know that it is only by grace that I have made it this far.
CHAPTER ONE

Introduction

In short, the habits we form from childhood make no small difference, but rather all the difference.

Aristotle (384-322 BC)

This quote by Aristotle demonstrates centuries old recognition that what is learned early in life is crucial to all that follows. These habits of which Aristotle spoke are the essence of what has since come to be referred to as culture. Our culture influences every aspect of how we interpret, experience, and interact with our world. As noted by Frederick Leong (1994), Ruth Benedict captured the full depth and importance of culture beautifully in her (1934) book, Patterns of Culture. As Leong also noted, the language used by Benedict is sexist due to the era in which it was written, but the message is timeless and clear.

No man ever looks at the world with pristine eyes. He sees it edited by a definite set of customs and institutions and ways of thinking. Even in his philosophical probings he cannot go behind these stereotypes; his very concepts of the true and false will still reference to his particular traditional customs…. From the moment of his birth the customs into which he is born shape his experience and behaviour. By the time he can talk, he is the little creature of his culture, and by the time he is grown and able to take part in its activities, its habits are his habits, its beliefs his beliefs, its impossibilities his impossibilities. Every child that is born into his group will share them with him, and no child born into one on the
opposite side of the globe can ever achieve the thousandth part. There is no social problem it is more incumbent upon us to understand than this of the role of custom. Until we are intelligent as to its laws and varieties, the main complicating facts of human life must remain unintelligible. (Leong, 1994, pp. 114-115)

The importance of valuing and understanding cultural diversity is more important today than ever. Proximity and interdependence of diverse cultures worldwide has increased dramatically in the past few decades (Cortez, 1999; Herr, 1999; Sodowsky, Taffe, Gutkin & Wise, 1994; The U.S. Bureau of Census, 1995, Table 19). Further, every indication is that this trend will continue well into the foreseeable future. The United States is no exception. Our nation has also experienced rapid increases in racial and cultural diversification in recent decades. As a result, historically separate groups are coming into contact with each other with increasing frequency. Effects from these changes in demographics and the resulting intergroup contact, has been as varied as the people involved. It is critical for all groups that we learn to understand and successfully interact with one another. This necessitates reaching across boundaries and honestly addressing issues that have historically been taboo in the United States. Racial differences and inequalities is one of the issues that it is critical for our nation to acknowledge and address.

Arguably no single aspect of the culture of the United States is more ubiquitous, firmly embedded, or ultimately oppressive than that of racism. Even prior to the inception of the United States as a nation, its seeds were firmly rooted in ethnocentric, racist, and cultural absolutist beliefs. These beliefs allowed an entire litany of atrocities
against those viewed as different (e.g. - European explorers and colonists claimed
discovery of, and sovereignty over, an already inhabited continent; committed near
genocide of North America’s indigenous peoples; kidnapped, raped, and enslaved
African peoples, etc.). Further, the atrocities continued even after the United States
became a nation. This despite the fact that the nation was founded on a constitution that
Supreme Court Justice John Marshall Harlan (Plessy v. Ferguson, 1896) asserted is
“color-blind, and neither knows nor tolerates classes among citizens. In respect to civil
rights, all citizens are equal before the law.” Yet, despite the “color-blind” nature of the
constitution on which the United States was founded: laws prohibiting people of color
(i.e.- Blacks, Asians, Native Americans) from presenting legal testimony against any
white person was upheld (People v. Hall, 1854); legalized slavery continued until 1865;
following emancipation, “Black Codes” were passed to ensure people of color could not
exercise their freedom; individual states adopted constitutions that denied rights to
immigrants of color (e.g. – California’s 1876 constitution made it unlawful to employ any
Chinese or Mongolian); Native Americans were denied citizenship even after ratification
of the fourteenth amendment (Elk v. Wilkins, 1884); the order to relocate Japanese
Americans involuntarily to “war relocation centers” during World War II was upheld
(Korematsu v. United States, 1944), and the list goes on. Further, this list is a very small
sample of only historically discriminatory racial practices officially sanctioned by the
United States legal system. It is well documented that racism is promoted and
maintained in numerous other ways. Racism is maintained daily through the media, work
place, education system, and even within the English language itself.
Bulhan (1985) listed racial oppression first among the most salient forms of oppression in the contemporary world. Within the culture of the United States, race shades every experience of an individual’s life. It influences both how one sees the world and also how one is seen by the world. Thus, because of the sociopolitical significance placed on race in the United States; the set of customs, institutions, ways of thinking, and concepts of true and false which Benedict described as combining to define one’s culture, are in large part determined, in this society at least, by race. Silberman (1964) stated, “Solving the problem of race is not only the most urgent public business facing the United States today; it is also the most difficult” (p. 5). Contemporary events such as the Amadou Diallo shooting, James Byrd Jr. being dragged to death behind a pickup truck, the Rodney King beating, the Charles Stuart and Susan Smith cases, the O.J. Simpson trial, the appeal to racial prejudices in electoral politics, the bitter debates about affirmative action, welfare reform, and racial profiling by police are all evidence that race issues are at every bit as relevant today as ever in the United States.

Finding ways to reduce prejudice and discrimination has long been identified as the central issue in combating racism in our society (Oskamp, 2000). Dating back to the 1920’s, prejudice has been one of the most studied areas in all of psychology and sociology. However, most of this study has been directed at understanding the nature, causes, and consequences of prejudice. Relatively little research has been devoted to questions of how to reduce prejudice and increase societal equality and social justice (Oskamp, 2000). However, with minority voices ever increasing in number and volume, a shift has begun. Diversity has become a favorite discussion topic in business,
academia, and government, and pro-active approaches toward reducing prejudice and
discrimination are being sought (Karp and Sammour, 2000).

Universities and colleges have been, and continue to be, among the institutions at
the very center of the multicultural storm. University students were front and center
during the civil rights movement with the formation of the Student Nonviolent
Coordinating Committee (SNCC) and the coordination of massive and effective sit-ins
and freedom rides. Black students demanding access to historically White only
universities and colleges were a critical part of legal efforts to overturn segregation. In
Sweat v. Painter 1950, the Supreme Court ruled that the University of Texas had to
integrate its law school. In September of 1962, a federal court ordered the University of
Mississippi to accept James Meredith, a Black Air Force Veteran. Meredith eventually
graduated from “Ole Miss” despite extreme efforts, including days of violence and
rioting by Whites, to prevent his attending that institution. Many students, both Black
and White, demonstrated incredible commitment and determination and made
tremendous sacrifices, some even their lives, in order to end segregation and address
racism and discrimination. One way that their efforts and sacrifices have paid off is in
the form of increased minority enrollment at universities nationwide.

Unfortunately, over the last several decades, increased diversity on campuses has
been followed by increased racial separatism and intolerance (Dalton, 1991; Ponterotto
and Pedersen, 1993; Tatum, 1997). This has led to increased polarization of forces at
colleges and universities as demonstrated by D’Souza’s (1991) attack on affirmative
action programs at universities. One side is working to implement more and better
diversity training and multicultural education programs, while the other side is working
for the complete dismantling of all university affirmative action programs (e.g., repeal of affirmative action legislation regarding state universities in Michigan and California).

While it is not currently clear which direction this debate will ultimately take, it is clear that race relations in the United States will be drastically impacted for a very long time, if not forever. Thus research and innovation regarding multicultural education and prejudice reduction efforts are critical.

**Multicultural Education**

The primary goal of multicultural education or diversity training is to assist people in competently and humanistically interacting and working with people who are different from them yet share inherent human similarities (Pedersen, 1991). Most multicultural education programs are based on the core belief that prejudices, discrimination, and intergroup conflict are primarily a result of ignorance, misinformation, or misperception. “Without significant intercultural contact, people’s perceptions of individuals representing other racial/ethnic groups are more often than not based on faulty information” (Ponterotto and Pedersen, 1993, p.28). As a result, multicultural education efforts have focused primarily on providing accurate information.

Multicultural education efforts at universities have largely been molded to fit into already existing standard education formats. This has resulted in criticism that most multicultural training programs overemphasize cognitive approaches, rely too heavily on traditional intellectual “university” teaching approaches (e.g., formal lectures, reading, and writing assignments), and under emphasize experiential training and affective learning (Parker and McDavis, 1979). These traditional teaching approaches have not proven effective in addressing key factors involved in prejudice, racism, and
discrimination such as deeply rooted feelings and attitudes or interpersonal factors (Parker and McDavis, 1979).

Cultural self-awareness has been identified as one key factor involved in the reduction of prejudice and racism. Cultural self-awareness is recognition of the unique cultural lens through which an individual filters her or his experience and understanding of the world. Development of cultural self-awareness has been established as a critical part of multicultural education. Counseling psychology training programs have been mandated to insure that their trainees develop certain multicultural competencies to best serve the needs of all who seek services. A primary competency on which scholars have focused is cultural self-awareness (e.g., Sue, Arrendondo, & McDavis, 1992). Fukuyama and Neimeyer (1985) suggested that increased self-awareness regarding cross-cultural interactions is a critical step in developing cultural sensitivity and multicultural competency. D’Andrea, Daniels, and Heck (1991) also noted the need for heightened self-awareness and advocated the use of less structured classroom discussion to help individuals examine their biases and stereotypes about groups different from their own. Group interaction where students of different cultural backgrounds are brought together in the same class, can facilitate students’ awareness of their attitudes toward members of other groups as well as the attitudes of others toward them, and has been identified as a particularly useful technique for increasing multicultural awareness by providing the experiential training and affective learning opportunities that most multicultural education efforts are currently lacking (Parker and McDavis, 1979; Sabnani, Ponterotto, and Borodovsky, 1991).
Group interaction with participants from diverse groups, or intergroup contact, is the basis for the contact hypothesis. In brief, the contact hypothesis states that intergroup exposure will experientially provide accurate knowledge regarding similarities and differences, and thus lead to better intergroup understanding and appreciation. There is a great deal of support for the basic premise behind this belief (Pettigrew, 1971, 1998b; Pettigrew and Tropp, 2000; Riordan, 1978). However, it is becoming more and more empirically clear that the conditions under which contact occurs are critical (Pettigrew, 1971, 1998b; Pettigrew and Tropp, 2000). In fact, intergroup contact lacking certain conditions (e.g., equal status between groups in the condition, common goals, no competition, authority sanction) will very often lead to increased prejudice, discrimination, misunderstanding and conflict (Cortez, 1999). Universities have largely focused on increasing the number of minority students admitted, under the belief that greater numbers will correlate with increased intergroup contact leading to greater understanding and appreciation. Unfortunately, this has only served to bring diverse students into proximity with each other, with very little attention to or effort made to provide the other necessary conditions for positive intergroup contact. This has led to increased racial separatism, intolerance, and violence (Cortez, 1999). Obviously more needs to be done to promote positive intergroup contact. One promising possibility is through intentional and programmed efforts designed to facilitate positive and meaningful intergroup contact experiences.

The Role of Counseling Psychology

Due to the crucial need for certain key conditions to be present in order to realize positive intergroup contact (i.e. – contact leading to greater understanding and
appreciation across groups), and the potential harm associated with failure to provide those conditions, programmed intergroup contact efforts require skilled facilitators. Counseling professionals are particularly well suited for this role. Unfortunately, the counseling profession has not historically been as involved as it could be in multicultural education or prejudice, racism, and discrimination prevention (Ponterotto, 1991; Ponterotto and Pedersen, 1993).

While counseling psychology, as a field, has in some ways been at the forefront in multicultural training for those entering the counseling profession, it has been notably absent regarding multicultural education efforts not directed towards its trainees. This is greatly affected by the long recognized sociopolitical nature of counseling and psychology (Betz and Fitzgerald, 1993; Bulhan, 1985; Bernard and Goodyear, 1992; Herr, 1999; Katz, 1985). As a modern discipline, psychology was born from Euro-American culture. Thus, for better or worse, modern psychology is reflective of the majority culture in the United States (Bernard & Goodyear, 1992), which makes psychology susceptible to the biases and limitations of the sociopolitical environment within which it is embedded. Bulhan (1985, p. 37) contended that, “the discipline [of psychology] has been enmeshed in that [Euro-American] history of conquest and violence.” He challenged that to discuss the psychology of racial oppression is also to pose questions about Euro-American psychology and its relationship to racial oppression. As a discipline, psychology has failed to recognize the relevance of implicit and explicit cultural differences, which has resulted in the identification of mental health professionals as status quo agents and perpetrators of racial and cultural oppression (Sue, 1991; Bulhan, 1985).
However, as greater numbers of racial and ethnic minorities have successfully made their voices heard in U.S. society, psychology, always a reflection of the society in which it exists, has also begun to hear the cries of its minority members. In recent decades there has been an increased appeal by minority and feminist practitioners for the profession to reexamine and reevaluate the theory and practice base of counseling psychology (Katz, 1985; Pedersen 1991). The basic premises on which psychology was founded are now being challenged at every level. It is becoming more and more widely accepted that United States norms, accepted by psychology, are a reflection of middle class White male values and not necessarily representative of the increasingly diverse larger U.S. culture. As a result, multiculturalism has been identified as the fourth force in counseling psychology (Pedersen, 1991). Mental health professionals are being required to shed the role of passive status quo agents and perpetrators of racial and cultural oppression, and instead embrace the role of advocates for social change.

Social advocacy fits well within the mission of counseling psychology, and working with the underrepresented and/or oppressed sends the message that the profession cares and also values the experiences of the community as a whole. Principle F of the Ethical Principles of Psychologists and Code of Conduct (American Psychologist, 1992, p.32) directly addresses psychologists’ social responsibility. It states psychologists need to be “aware of their professional and scientific responsibilities to the community and the society in which they work and live.” It further asserts the responsibility to “make public their knowledge of psychology in order to contribute to human welfare…, work to mitigate the causes of human suffering…, and encourage the development of law and social policy that serve the interests of their patients and clients
and the public.” While there continues to be much debate in some arenas regarding the value of multiculturalism and multicultural education, as a field, counseling psychology has made its position quite clear. Multiculturalism has been proclaimed the “fourth force” in counseling psychology. Counseling psychologists have an ethical responsibility to promote multiculturalism and multicultural education. However, resistance associated with multiculturalism and uncertainty about how to properly provide multicultural education has largely relegated counseling psychologists to the role of spectator regarding broader multicultural education efforts. In order to fully shed the role of status quo agents and perpetrators of racial and cultural oppression, counseling psychologists must more actively engage in multicultural education efforts, otherwise the sin has simply changed from one of commission to omission. Counseling psychologists are uniquely qualified to understand and address the profound psychological effects that racism has on everyone that is part of United States society. Thus, they have a responsibility to be among the leaders in the multicultural education movement instead of simply cheerleading from the sidelines. Many colleges and universities avoid addressing multiculturalism on any real level (Carlos Cortez, 1999). Counseling psychology has a responsibility to help rectify this situation.

Regardless of the role that counseling psychologists take, or what colleges and universities officially do or don’t do, multicultural education will continue to occur at colleges and universities. As diversity and contact between diverse groups increases globally, nationally, and on campus, students learn about multiculturalism. As they are forced to compete in a more diverse pool for admissions, assistantships, internships, and jobs, students learn about multiculturalism. As they encounter differences in the dorms,
cafeterias, and classrooms, they learn about multiculturalism. The question is not whether students will learn about multiculturalism, the question is only what they will learn. Exposure to differences can lead to misunderstandings, incompatibilities, and personal clashes, even on campuses that espouse the celebration of diversity. Or, exposure to differences can promote greater understanding of and appreciation for difference.

Carlos Cortes (1999) stated, “As one who lectures, gives diversity workshops, and troubleshoots on several dozen campuses every year, I find that the problem of space-sharing amidst growing diversity is virtually a universal higher education challenge that cannot be resolved by glowing mission statements or cheerleading platitudes about celebrating diversity or treating everyone as part of the human race…Universities face the challenge of facilitating healthy, supportive, and affirming group aggregation while simultaneously trying to inhibit calcification into self-segregation. To do so, higher education institutions need to work to build bridges among communities. The challenge, then, gives rise to an opportunity – the fostering of intergroup relations through serious, civil intergroup conversation and collaboration” (pp. 14-15).

Need for the Study

The current research study reported here was designed to investigate one particular attempt to respond to the above challenge/opportunity. The intervention program being investigated is run through a university counseling center primarily by counseling psychologists and their trainees. The hope is that studying this particular program will help add to the current understanding of what conditions and processes lead
to improved intergroup understanding and relations, and also help identify an appropriate vehicle through which counseling psychologists can play a more active role in wider university multicultural education efforts.

The primary purpose of this research is to study the process and evaluate the efficacy of one “real world” effort, run through a university counseling center, to promote and facilitate intergroup contact in a manner that helps challenge and ultimately reduce prejudice and discrimination. Specifically, the study investigates whether multicultural awareness or racial attitude changed over the course of a specific, nine-week, process-oriented, group intervention; examines whether multicultural awareness development or prejudicial attitudes are affected differently as a result of exposure to readings chosen to provide facts versus more personal readings chosen to promote empathy; explores what incidents, interactions, or other factors influence the process, and how.

Research Questions

1. Were there significant changes in attitudes or level of awareness related to racial, ethnic, or cultural diversity for students participating in the nine week intervention (Dialogues on Race) compared to a control group of students who did not participate in the intervention? More specifically, do differences exist between students participating in the intervention and students not participating in the intervention, as measured by the Millville-Guzman University-Diversity Scale (M-GUDS), the Color-Blind Racial Attitudes Scale (CoBRAS), or the Quick Discrimination Index (QDI)?

2. Did the type of readings to which participants were exposed as part of the nine week Dialogues on Race intervention make a difference in the effectiveness of the intervention? More specifically, do differences exist between students participating in
the intervention based on exposure to factual versus personal readings as measured by change in scores from pretest to posttest on the M-GUDS, CoBRAS, or QDI?

3. For all participants in the intervention, both racial majority and racial minority group members, what aspects of the intervention did they identify as important, when, and why? More specifically, examining weekly critical incident reports, posttest structured interviews, and course evaluations, what do students have to say about their experience? What does this phenomenological data add to our understanding of the intervention experience, and in what ways does this qualitative data support, challenge, or help explain the quantitative findings?
Prejudice and racism are a part of the American cultural heritage. At the same time, however, prejudice and racism are inconsistent with the American Dream and destructive for its victims, witnesses, and even perpetrators. Daily events document the pervasive and destructive nature of prejudice and racism in the United States and abroad. A New York Times series (Summer 2000), entitled, “How Race is Lived in America,” was the result of a year long examination by a team of reporters looking at current race relations in the United States. They found that race relations continue to affect the daily experiences of Americans in schools, sports arenas, workplaces, and at worship. As the global community becomes more interdependent and the United States realizes greater heterogeneity, the need for concentrated attention to intergroup relations becomes increasingly obvious (Ponterotto and Pedersen, 1993). Diversity has become a favorite discussion topic in business, academia, and government (Karp and Sammour, 2000); as a result, renewed attention is being paid to issues of intergroup conflict, prejudice, and discrimination.

It has long been understood that prejudice is often at the core of intergroup conflict and discrimination. For this reason, prejudice has been a major topic of study in the social sciences since as far back as the 1920’s. In fact, it is one of the most-studied areas in all of psychology and sociology. However, most of the research to date has aimed at describing the nature of prejudice, understanding its causes and expressions, and to some extent, at documenting its consequences (Oskamp, 2000). Recent changes in the culture of the social sciences, due largely to social and political shifts in the larger U.S.
culture, have created a demand for research to focus on finding ways to reduce prejudice and work to create a more equitable and just society. However, Oskamp (2000) stated that, “efforts to reduce prejudice necessarily have to be built on an understanding of the causes of prejudice (p.3).” Therefore, a very brief review of the development of prejudice is offered here.

*Development of Prejudice*

Interracial and interethnic animosity can be traced back to the earliest human history and interactions. Conflict over physical appearance, beliefs, and customs has been, and will continue to be, a primary source of unrest in the world (Ponterotto, 1993). Allport (1979) convincingly argued that humans have a propensity toward prejudice. Yet, regardless of how predisposed people are to developing prejudice, prejudice does not develop in a vacuum. Prejudicial attitudes are learned through explicit or implicit messages delivered at home by parents, through the media, through our education system, at work, through politics and laws, and through our social networks.

Many authors (e.g. Allport, 1979; Oskamp, 2000; Ponterotto and Pedersen, 1993; Hamilton and Trolier, 1986; Duckitt, 1992) have noted that the causes of prejudice and discrimination are multiple and elaborately intertwined. Allport (1979) proposed that an interaction of three factors easily leads to prejudicial views: (1) a tendency towards ethnocentrism, (2) lack of significant intergroup contact, (3) an inclination to organize information into pre-developed categories. Duckitt (1992) proposed a four-level model of possible factors: (1) genetic and evolutionary pre-dispositions; (2) societal organizational, and intergroup patterns of contact and norms for intergroup relations; (3) mechanisms of social influence that operate in group and interpersonal interactions; (4)
personal differences in susceptibility to prejudiced attitudes and behaviors, and in acceptance of specific intergroup attitudes.

Despite the complexity inherent in the multiple and intertwined causes of prejudice, racial prejudice could not develop without an implicit or explicit societal message of superiority or inferiority based on race. “Given the pervasive nature of negative racial prejudice and modern racism, one must conclude that society is condoning racial inequality either directly through continuing discrimination or indirectly through its refusal to alter the status quo” (Ponterotto and Pedersen, 1993, p.31). However, efforts to alter the status quo in this area have been made, and some would argue have been very successful. Few would disagree that there has been dramatic change in race relations in the United States over the past half-century. Prior to the civil rights legislation of the 1960s, blacks and other peoples of color suffered blatant restrictions in residential, educational, and employment opportunities. Civil rights legislation made such discrimination and segregation illegal. Consequently, Americans of color currently have greater access to political, social, and economic opportunities than at any previous time in the nation’s history (Dovidio and Gaertner, 1998). Race relations in the United States appear drastically better now than ever before. Yet there are signals that despite undeniable improvement, there may be much more room for further improvement than is readily apparent from the surface. In the past decade there have been race related riots in numerous United States cities, increased reports of significant racial incidents and protests on college campuses, racially motivated arson fires, and ongoing incidents of racial profiling by law enforcement agencies. Further, gaps between racial minority and majority Americans in physiological areas (e.g., infant mortality, life expectancy) and
economic areas (e.g., employment, income, poverty) not only continue to exist, but in some cases the disparities have actually increased over the past thirty years (Hacker, 1995). All of this illustrates that efforts to fight racism and discrimination have fallen short of the goal. Racism and discrimination have not been eliminated, but have instead mutated in order to survive. What the nation faces now is a more subtle form of racism that, while less overt, is just as destructive and perhaps more insidious than old-fashioned racism (Dovidio and Gaertner, 1998).

Contemporary Prejudice

Across time, racial attitudes in the United States have become less negative and more accepting. Negative stereotypes are declining. Dovidio and Gaertner (1998) give numerous examples of various research and survey efforts that dramatically indicate that Whites are becoming more accepting of minorities, Blacks in particular. For example, the number of white respondents who described Blacks as lazy decreased from 75% in 1933 to 5% in 1993 (Dovidio, Brigham, Johnson, and Gaertner, 1996). In 1958, the majority of White respondents reported that they would not vote for a well-qualified black presidential candidate. In contrast, in 1994, over 90% reported that they would (Davis, 1994). Further, the American National Election Survey (1995) found that this increased tolerance of white Americans extends beyond Blacks to other racial and ethnic minority groups as well.

Despite these encouraging trends in the intergroup attitudes of white Americans, there are still reasons for concern. First, there remains 10%-15% of the white population that still express the old-fashioned, overt form of bigotry. Second, a substantial portion of the white population expresses racial tolerance but do not endorse full racial equality.
The third reason for concern is that there is evidence that many of the 85%-90% of the white population who state, and likely believe, that they are not prejudiced, likely continue to be guilty of a more modern and subtle form of bias. This more subtle form of bias may be a major factor in the persistence of racism and discrimination in American society (Dovidio and Gaertner, 1998).

While research efforts all seem to focus on White’s racial attitudes toward minorities, especially Blacks, there is reason to believe that the same forces that exhort white Americans to express more egalitarian attitudes and beliefs, also influence minority Americans. Thus, racial minority members are probably also less likely to express blatantly biased attitudes and beliefs, either openly or to themselves. However, they may also unconsciously be practicing this more modern and subtle form of bias against members of other races or ethnicities, or even against members of their own group.

Over the past 20 years, a number of researchers (i.e., Dovidio and Gaertner, 1991; Gaertner and Dovidio, 1986; Gaertner et al., 1997) have investigated a prevalent type of modern racial bias, called aversive racism. Aversive racism is defined in contrast to “old-fashioned” racism. Old-fashioned racism is expressed directly and openly, aversive racism is a subtle, often unintentional, form of bias exercised by Americans who possess strong egalitarian values and who believe that they are nonprejudiced. Aversive racists possess negative racial feelings and beliefs of which they are unaware or that they try to dissociate from their nonprejudiced self-images. Aversive racists’ negative feelings do not reflect open hostility or hate. Instead, their reactions involve discomfort, uneasiness, disgust, and sometimes fear. They may find members of other races or ethnicities
“aversive,” but at the same time they also find any suggestion that they might be prejudiced aversive (Dovidio and Gaertner, 1998).

According to Dovidio and Gaertner (1998), because aversive racists consciously recognize and endorse egalitarian values, they will not discriminate in situations in which they recognize that discrimination would be obvious to others or themselves. Thus, if the appropriate egalitarian response in a given situation is clear, aversive racists will not discriminate. To discriminate in such a situation would directly threaten their nonprejudiced self-image. However, in situations where the appropriate response is not as obvious or when an aversive racist can justify or rationalize a negative response on the basis of some factor other than race, then their bias will be expressed.

Part of the discomfort that aversive racists experience is due to concern that they may act inappropriately and appear prejudiced when interacting with members of another race or ethnicity. As a result, aversive racists strictly adhere to established rules and codes of behavior in interracial situations. They also frequently assert that they are color-blind, because if they do not see race, then they cannot be racists. Because of the ambivalence that characterizes their racial attitudes; aversive racists’ interracial behavior is more variable than that of old-fashioned racists (Dovidio and Gaertner, 1998).

Support for the aversive racism framework has been found across a broad range of situations. Gaertner and Dovidio conducted one early test of the framework in 1977. Their experiment was modeled after the Kitty Genovese incident, which occurred in New York City in 1964. The situation created by Gaertner and Dovidio was a serious accident. They led some subjects to believe that they would be the only witness to this emergency, while others were led to believe that there would be two other people present
in the situation who were also aware of the emergency. Race of the victim was also varied. In half of the cases the victim was white. In the other half of the cases the victim was African American. The participants in the study as well as the other two people sometimes presumed to be present were all white. When the white bystanders were the only witnesses to the emergency, they helped very frequently and equivalently regardless of the race of the victim. In fact, they helped black victims somewhat more often than white victims (95% versus 83%, respectively). There was no evidence of old-fashioned racism. In contrast, when given an opportunity to rationalize not helping on the basis that one of the other witnesses could intervene, they were less likely to help, particularly when the victim was black. When participants believed that there were two other bystanders, they helped the black victim half as often as they helped the white victim (38% versus 75%).

Another area in which racial bias has historically been demonstrated is in the United States judicial system. Blacks, compared to whites, have consistently across time and location been more likely to be convicted of crimes, and, if convicted, sentenced to longer terms for similar crimes. In addition, blacks are more likely to receive the death penalty (Dovidio and Gaertner, 1998). However, racial disparities in judicial outcomes are declining (Nikerson, Mayo, and Smith, 1986). More subtle forms are replacing blatant expressions of racism in judicial decisions. Thus, racial bias continues to influence the judicial process, but mainly when it can be justified on some basis other than race. In a laboratory simulation study, Johnson, Whitestone, Jackson, and Gatto (1995) examined the effect of the introduction of inadmissible evidence, which was damaging to the defendant’s case, on white’s judgments of black or white defendant’s
guilt. No differences in judgments of guilt occurred as a function of defendant race when all the evidence presented was admissible. However, the presentation of inadmissible evidence increased judgment of guilt when the defendant was black, but not when the defendant was white. Furthermore, suggesting the unconscious or unintentional nature of the bias, participant’s self-reports indicated that they believed that the inadmissible evidence had less effect on their decisions when the defendant was black than when the defendant was white. Dovidio and Gaertner (1998), review numerous other studies supporting that modern or aversive racism plays a major role not only in judicial decisions, but also on survey responses or other self-report measures, job evaluations or hiring decisions, etc.

Despite efforts by aversive racists to be fair and just, or at least appear to be so, their biases come through and undermine their ability to be truly egalitarian. A study by Dovidio, et. al. (1996) found that even the effort to appear nonracist might be derailed by underlying attitudes of the aversive racists. The study examined the possibility that whites and blacks attend to different aspects in social interactions. A black person and white person first interacted and then completed questionnaires that asked how friendly they felt they behaved during the conversation and how friendly their partner acted. In general, whites’ perceptions of their own friendliness correlated with their self-reported prejudice scores (i.e., those who reported that they were less prejudiced also reported that they behaved in a more friendly manner with the black partner in the subsequent interaction). These perceptions were apparently guided by the conscious attitudes of the whites. Thus on a conscious level they seemed to be behaving consistently. In contrast, the perceptions of the black partners about friendliness of these same white participants
were more strongly associated with the whites’ response latency measures of bias. These results indicate that in assessing how friendly the white person was, blacks may have been considering not only the overt, consciously controlled behavior of the partner, but also concentrated on the unconscious behaviors (i.e., eye contact or other nonverbal expressions of discomfort) that whites were unable to monitor or control. These findings suggested that aversive racists might be sending very different messages than they intend to in interracial interactions.

The difference between intended and received messages can lead to potential conflict, mistrust, and frustration between well-intended individuals or groups attempting to bridge the racial divide. The evolution from old-fashioned or traditional forms to modern or aversive forms of racism brings with it a whole new set of concerns. Because modern or aversive racists view themselves as egalitarian and nonracists, their biases are not readily accessible even to themselves. This nullifies the ability for logical arguments or facts to change their biases, for they often know and support these arguments against their subconscious biases. Thus, their biases will likely only ever become apparent to them, or potentially be eliminated, through a less cognitive and more affective approach. Their cognitions are already in line with their desired egalitarian beliefs. However, their emotional responses when confronted by interracial interactions often belie their underlying racial biases. Thus the most obvious catalyst for potentially gaining access to these unconscious biases is through intergroup contact.

*Intergroup Contact*

Social psychology has long held that a major means of reducing intergroup prejudice is through contact between groups. The Human Relations Movement was the
name given to the organized effort to correct negative stereotypes following World War II. The movement was characterized by a naïve faith that getting to know each other across group lines would remedy racial tensions by helping people recognize a common humanity. The movement failed to understand the many cognitive, affective, situational and institutional barriers to positive contact effects. Yet the basic premise of this movement led to the ongoing and systematic study of intergroup contact effects (Pettigrew & Tropp, 2000).

Robin Williams (1947), followed by Gordon Allport (1954), formalized the contact hypothesis, and it has been the focal point for the study of intergroup relations ever since. Allport’s version of the hypothesis held that intergroup contact will lead to reduced intergroup prejudice if, and only if, the contact situation meets these four conditions: (1) equal status between groups in the condition, (2) common goals, (3) no competition between groups, and (4) the contact must receive some form of authority sanction (Pettigrew & Tropp, 2000).

The Contact Hypothesis has inspired extensive research over the past half century. It has found support across a variety of societies, situations, and groups. Different research methods such as field studies (e.g., Deutsch & Collins, 1951), archival research (e.g., Fine, 1979), survey research (e.g., Herek & Capitanio, 1996), and laboratory studies (e.g., Cook, 1978; Desforges et al., 1991) have yielded findings which support the theory. However, the theory’s popularity has not remained consistent. Criticism of Allport’s contentions arose both in psychology and sociology (e.g., Ford, 1986; Hewstone & Brown, 1986; Rose, 1981; Rothbart & John, 1985). Yet research on contact has expanded over recent decades to new groups and across disciplines. Going beyond racial
and ethnic groups, researchers tested the theory with a diverse range of subjects and
target groups. Participants in contact studies have included college students, children,
adolescents, and adults, with target groups ranging from the elderly to homosexuals to the
physically disabled to the mentally ill. Not only have results from various studies often
resulted in inconclusive or contradictory results, but reviews of the extensive literature in
this area has also served to further cloud the issue by also reaching conflicting
conclusions (e.g., Reviews by Pettigrew, 1971, 1998b, and Riordan, 1978, concluded by
stating that literary evidence supports the hypothesis. In contrast, Ford’s 1986 review
concluded the opposite). Because of the vastness of the literature and enormous variety
of studies, a review of Pettigrew and Tropp’s 2000 meta-analysis is an effective way to
start to make sense of and summarize research in this area.

Review of Meta-analysis

Pettigrew and Tropp (2000) performed what is believed to be the first meta-
analysis to evaluate the massive body of research that has accumulated over numerous
decades and disciplines pertaining to intergroup contact. This meta-analysis utilized 203
individual studies with 313 independent samples and 746 separate tests. A combined
90,000 subjects from 25 different nations participated in this body of research. While the
majority (73%) of the studies are from the United States, extensive research from other
nations, including work from nine developing countries, is included.

Because of the enormous variety of studies, the meta-analytic test of this research
required a clear definition of intergroup contact. Pettigrew and Tropp (2000) defined it
as actual face-to-face interaction between members of clearly distinguishable and defined
groups. From this definition they established four inclusion criteria: (1) Only studies in
which intergroup contact acted as a causal, independent variable for intergroup prejudice were included. This requirement excluded research that treated intergroup contact as a dependent variable in explaining how and why contact occurred. (2) Only research that involved contact between members of discrete, clearly distinguishable groups was included. This rule allowed groups with cross-cutting categories, but only if the categories were clearly defined. (3) The research had to involve some degree of direct intergroup interaction. The interaction could be observed or reported, or it could occur in such focused situations as classrooms where direct contact is unavoidable. This rule eliminated a variety of studies often cited in summaries of contact research. For example, it excluded research that utilized rough proximity or group proportions to infer intergroup interaction, unless the research carefully demonstrated that intergroup proximity correlated highly with actual contact. This rule also eliminated investigations that attempted to gauge contact with indirect measures such as knowledge about an outgroup. Studies asking about attitudes toward contact were also excluded, unless such indicators were directly linked to prior intergroup experience. Research that categorized subjects into groups that did not directly interact, as in many minimal group studies, was also excluded. (4) The prejudice dependent variables had to be collected on individuals, rather than simply as an aggregate outcome, and comparative data had to be available to evaluate any changes in prejudice. These comparisons made use of precontact measures in within-group designs or control-group data in between-group designs.

Within these boundaries, Pettigrew and Tropp accepted a wide variety of prejudice indicators. This enabled them to test whether intergroup contact had differential effects on various components and measures of prejudice. They examined
studies of intergroup beliefs, social distance, affective, stereotype, and sociometric indicators.

Initial results indicate that there is an inverse association between intergroup contact and prejudice for all studies, samples, and tests. It was noted that the mean estimates that were weighted for sample size were smaller than the unweighted mean estimates. This difference indicates that the less reliable studies with smaller numbers of subjects yielded higher mean effect sizes. As a conservative procedure, only effect size estimates that were weighted for sample size were employed in the analyses. Pettigrew and Tropp employed an overall conservative approach to the meta-analysis. Still, the findings clearly indicate that the effect of intergroup contact to reduce prejudice is “highly significant.” They also went to great lengths to minimize potential causal sequence problems that have historically opened the door to criticism of research findings that make the claim that intergroup contact reduces prejudice.

Pettigrew and Tropp (2000) noted that potential selection bias limits the interpretation of many studies of intergroup contact and consequently the effects found in their meta-analysis. Instead of optimal contact reducing prejudice, the opposite casual sequence could be operating. That is, prejudiced people may avoid contact with outgroups. They noted three basic methods that can overcome this limitation, and cited studies that employed these methods. The first method is longitudinal studies. This is extremely rare in intergroup research. However, Sherif’s (1966) Robbers’ Cave study is a rare exception. This famous study revealed optimal contact as the cause of reduced prejudice while eliminating the possibility of selection bias. The second method noted by Pettigrew and Tropp to overcome the casual sequence problem is statistical methods
borrowed from econometrics that allow researchers to compare the reciprocal paths (i.e., contact reduces prejudice versus prejudice reduces contact) using cross-sectional data. Research (e.g., Herek and Capitanio, 1996) using these methods indicate that prejudiced people do indeed avoid intergroup contact. However, the path from contact to reduced prejudice is even stronger (Pettigrew, 1997; Powers and Ellison, 1995). Thus it appears that while both sequences operate, the more important effect is intergroup contact reducing prejudice. The final method noted by Pettigrew and Tropp was to severely limit choice. By eliminating the possibility that initial attitudes caused differential contact, research where choice is eliminated provides a clearer indication of how intergroup contact alters prejudice. Surprisingly, the no-choice studies (N = 30) and samples (N = 44) included in Pettigrew and Tropp’s meta-analysis provided by far the largest effect sizes between intergroup contact and prejudice. Pettigrew and Tropp offered three possible factors that may work alone or in combination to explain this unexpected outcome. First, they noted that the overall quality of the no-choice studies tended to be higher than the other studies (i.e., containing more experiments and better controls). Additionally, they offered two interrelated possible explanations: fewer ceiling effects and more cognitive dissonance. No-choice contact situations typically inevitably involve a wider range of initial prejudice scores, thus they are less constrained by ceiling effects. Further, initially prejudiced individuals in no-choice situations should experience greater cognitive dissonance from intergroup contact than their less prejudiced counterparts. Regardless of which of the three methods identified to eliminate or reduce the causal sequence problem was utilized in various studies, the resulting data indicated that contact
as the cause of reduced prejudice is more important than the selection-bias possibility that prejudiced people avoid contact.

While Pettigrew and Tropp (2000) noted that their meta-analysis is not yet complete, they stated that the preliminary indications demonstrate strong trends that hold significant theoretical and empirical implications. They offered seven points deserving of emphasis. 1) Overall, face-to-face interaction between members of distinguishable groups is importantly related to reduced prejudice. Of the 203 studies used in their meta-analysis, 94% found an inverse relationship between contact and prejudice. Further, while bigoted people do avoid contact with outgroup members, the causal sequence from contact to reduced prejudice is stronger. 2) Most relevant studies demonstrated generalization of effects from immediate participants to the entire outgroups regarding the association between contact and reduced prejudice. In fact, a few studies show significant generalization of effects to outgroups not even involved in the contact. 3) One major mediator of the size of contact-prejudice effects involves whether the participants are from majority or a stigmatized minority group. Majority participants reveal much larger mean effects than do minority participants. Pettigrew and Tropp suggest that this difference may correlate with a difference in perceptions related to the contact conditions. They assert that special attention needs to be paid to make sure that contact conditions are structured in a manner where both majority and minority participants will perceive them as optimal. While this is certainly an important consideration, there are other potential explanations for the differences in effect size that come readily to mind. Differences in effects for majority versus minority participants may also represent less dissonance in the intergroup contact situation (minority participants are likely more accustomed to
intergroup encounters), or perhaps minority participants experience more of a ceiling effect (minority participants are likely to not only have more intergroup experience, but also have more developed racial attitudes as a result of the innate salience of race for them). 4) The more rigorous the study, the larger the negative contact-prejudice effects. Research employing control groups with no outgroup contact, experimental designs, and more reliable measures of both independent and dependent variables all yielded larger mean effects. This indicates that the more reliable a research study, the more it supports contact theory. 5) Intergroup contact relates to a wide range of prejudice measures. Early versions of contact theory concentrated on cognitive effects such as learning about the outgroup and altering biased beliefs and stereotypes. However, results from this meta-analysis suggest that changes are much broader. In particular, affective indicators revealed especially large effects. Affect has gained greater attention by contact researchers in recent years, just as it has in many areas of psychology. According to Pettigrew and Tropp, results from their meta-analysis suggest that this is a fruitful direction for further empirical and theoretical work. 6) Contact appears to have significantly different effects with different types of outgroups. These differences suggest that different dynamics are involved, perhaps both in the nature of the prejudices and in the typical contact between certain groups. This would seem to reduce the value of Pettigrew and Tropp’s overall meta-analysis findings that intergroup contact in general reduces prejudice, and call for more research on contact between specific groups in question. 7) When Allport’s optimal conditions were indicated, relatively large effect sizes between contact and prejudice where found. Samples that experienced carefully structured intergroup contact situations that met most or all of the optimal conditions
described by Allport achieved markedly higher mean effect size than other samples. Pettigrew and Tropp concluded that further use of, and research on, such carefully structured contact situations is strongly indicated. Further, they noted that optimal intergroup contact has the power to influence affect, a realm they noted often goes untouched by less intrusive remedies for prejudice.

Despite the broad nature of intergroup contact research, as demonstrated by the number and diversity of studies included in Pettigrew and Tropp’s (2000) meta-analysis, there are clearly a number of areas in need of greater attention. Two particular areas noted by Pettigrew and Tropp as warranting more attention are the impact of intergroup contact experiences on minority participants, and the role of empathy.

The Role of Empathy

Social scientists have explored a variety of techniques for improving intergroup relations (e.g., Aronson, Blaney, Stephan, Sikes, and Snapp, 1978; Banks and McGee-Banks, 1995; Dovidio, Gaertner, Isen, and Lowrance, 1995; Ponterotto and Pedersen, 1993; Stephan and Stephan, 1996, etc.). The majority of these programs focus primarily on providing information about out-groups (Banks and McGee, 1995; Cushner and Landis, 1996; Ferdman and Brody, 1996). The literature on the success of such informational programs indicates only moderate success (McGregor, 1993; Stephan and Stephan, 1984, 1996). Further, there has been limited systematic exploration of which components of these programs are effective (Finlay and Stephan, 2000).

Emotion is critical in intergroup contact (Pettigrew, 1998). Anxiety is common in initial encounters between groups and may interfere with potentially positive outcomes (Islam and Hewstone, 1993; Stephan, 1992; Stephan and Stephan, 1985, 1989; Wilder,
Anxiety may lead to negative encounters even when intergroup prejudice is not present (Devine, Evertt, and Vasquez-Suson, 1996). Continued contact generally reduces anxiety, but negative experiences may keep this reduction of anxiety from occurring, or even increase it. In contrast, positive emotions aroused by optimal contact can improve contact effects. Empathy has been identified as playing a major role here (Pettigrew, 1998). In fact, empathy is one component of intervention programs that has received some attention as a possible key variable in improving intergroup attitudes (Finlay and Stephan, 2000). Several investigators have suggested that certain types of intergroup relations programs are successful as a result of their ability to generate empathy for minority groups (Bridgeman, 1981; Byrnes and Kiger, 1990; McGregor, 1993; Smith, 1990). Bridgeman (1981) argues that the success of cooperative-learning groups in educational settings can be attributed to empathy. She found that cooperative-learning experiences heightened students’ ability to understand other children’s perspectives. McGregor (1993) and Smith (1990) have both proposed that empathy may be a critical element of intergroup relations programs employing active role-playing techniques. The belief is that the process of actively playing the role of out-group members involves efforts to take the perspective of others, and elicits affective reactions that leads to experiencing dissonance (McGregor, 1993; Smith, 1990). This dissonance acts as motivation to modify previously held attitudes (Festinger, 1957).

Several studies have explored the techniques to increase empathy that require no active participation (Batson et al., 1997; Finlay, 1996, Finlay and Stephan, 2000, Finaly and Trafimow, 1998). Batson et al. found that instructions to empathize with an out-
group member led to more positive attitudes toward that individual, and those positive attitudes also generalized to other members of that individual’s social group.

The studies by Finlay (1996) and Finlay and Trafimow (1998) found that empathy plays a role in changing behaviors as well as attitudes toward outgroup members. The first study found that empathy mediated the relationship between cognitions associated with the concept of common humanity and attitudes and behavioral intentions toward people with AIDS (Finlay, 1996). The second study found that empathy mediated the relationship between cognitions associated with the private self and volunteering for a community AIDS organization (Finlay and Trafimow, 1998). Both of these studies measured reactive emotions (i.e., compassion and sympathy) that arise in response to an adversity experienced by another. While Finlay and Trafimow consider these reactive emotions a type of empathy, some researchers (e.g., Duan and Hill, 1996) believe that these reactive emotions should not be considered empathic emotions at all. They suggest that empathic emotions should be restricted to emotions that match those experienced by the target individuals (e.g., sadness, hopelessness, distress, or anger). Davis (1994) labels these types of emotions parallel empathy.

Finlay and Stephan (2000) noted that these previous studies of the effects of empathy on attitudes towards out-groups did not examine racial or ethnic minority groups. They postulated that the empathic processes involved when considering racial and ethnic groups might differ from those involved for other stigmatized groups because the studies of nonracial stigmatized groups asked the participants to empathize with suffering that they played no part in causing. In contrast, when members of the majority group are asked to empathize with victims of racial injustice, they are likely aware that
their own in-group is responsible for much of the suffering. Finlay and Stephan hypothesized that the role of empathy toward out-group members mistreated by an in-group with which the participant identifies might result in a much more complex response. They suggested that when the targets are members of nonracial groups, people can easily respond with both reactive and parallel empathy, and both types of emotional reactions should then work in concert to create positive changes in attitude. However, when the target is a member of a racial minority group, the emotions involved in parallel empathy may include negative reactions toward one’s own in-group (e.g., resentment or anger). The study by Finlay and Stephan (2000) included measures of both reactive and parallel empathy to examine their effects on intergroup affect, cognition, and evaluations in response to reading about acts of discrimination toward racial minority groups. They examined the effects of giving Anglo American students information about everyday incidents of discrimination against African Americans either with or without empathy-inducing instructions. The results indicate that reading about discrimination against African Americans or inducing empathy reduces in-group-out-group bias in attitudes toward African Americans versus Anglo Americans.

While the effects of empathy on racial attitudes in these studies are quite complex and did not always occur as predicted, the sum result indicates that empathy may be a promising tool to improve the impact of information-based intergroup-relations programs. However, as Finlay and Stephan (2000) noted, additional work is needed to understand how these beneficial effects might be best achieved and why they occur.
Contact and Minority Participants

The literature on intergroup relations is peculiar in the sense that by definition it attempts to examine the interactions between majority and minority group members. However, most of the research in the area has studied only the effects on majority participants (Devine and Vasquez, 1998; Pettigrew and Tropp 2000), or the effects of majority and minority group members in isolation (Devine and Vasquez, 1998). Further, the focus of the literature on majority group members has been on the nature of the attitudes and stereotypes majority members hold with regard to minority members (e.g., Allport, 1954; Cook, 1985; Devine, Monteith, Zuwerink, and Elliot; 1991; Dovidio and Gaertner, 1986; McConahay, 1986; Stephan, 1987). The focus has been on the origins of prejudice, and ways to eliminate it. The literature on minority group members, in contrast, has focused more on self-esteem and adjustment consequences of being a member of a stigmatized group (e.g., Crocker and Major, 1989; Jones et al., 1984; Seligman and Welch, 1991).

This approach of studying the effects of intergroup interaction on majority and minority participants is limited in several ways. First, this literature makes major assumptions about the effects on both majority and minority members in regard to their level of prejudice and the effects of discrimination on their self-esteem. Pettigrew (1998) reported several studies (i.e., Smith, 1994; Works, 1961) which indicated that minority members may also hold prejudicial views of members of other groups, be it majority group members or members of other minority groups. There is also evidence that being a member of the majority group can cause dissonance that results in adjustment consequences for majority group members in intergroup contact situations where they are
confronted with the role of their group in oppressing others (Dovidio and Gaertner, 1998). Thus, it seems appropriate to start with fewer assumptions, and explore a wider range of potential effects for both majority and minority participants in intergroup contact situations.

Finally, the most limiting factor in the traditional approach of studying minority and majority group members in isolation, is that it offers little help in understanding the nature of the interpersonal dynamics of intergroup contact (Devine and Vasquez, 1998). Devine and Vasquez (1998) go on to challenge that “to effectively address the problem of intergroup tension, the study of intergroup relations must be broadened to include analyses of how individuals think about the interactions, about themselves and their partners during these interactions, and about the outcomes of the interactions (p. 241).”

Obviously, the task of improving race relations is quite complex. Many factors need to be taken into consideration both from the minority and majority perspective. Further, contemporary prejudice or aversive racism is much more subtle and therefore difficult to see and address than traditional or old-fashioned racism. Intergroup contact appears to be a very promising approach to ultimately improving race relations. However, intergroup contact can actually damage race relations if the conditions are not correct. This calls for programmed and perhaps facilitated intergroup contact efforts. Counseling psychology, as a field, has demonstrated an understanding of the importance of multiculturalism and the ways that prejudice and racism undermine it. Further, the field has made a commitment to produce multiculturally competent counselors and to proactively addressing these issues. As a result, counseling psychologists’ training in multiculturalism and group dynamics make them uniquely qualified to take on the
challenge of designing and facilitating effective intergroup contact programs. Further, much of what has been learned by counseling psychology in its efforts to develop multiculturally competent counselors can contribute to these broader efforts to essentially develop multiculturally competent students in general.

Multicultural Counselor Training

The field of counseling psychology has taken the lead in attempting to incorporate multicultural diversity issues into academic curricula and applied training experiences for its students. Researchers, educators, practitioners, and other counseling psychology professionals have emphasized the importance of counselors gaining multicultural counseling competence. Multicultural counseling competence has been defined as the aggregate of counselor’s attitudes/beliefs, knowledge, and skills in working with individuals from a variety of cultural (e.g., racial, ethnic, gender, sexual orientation) groups (Sue, Arrendondo & McDavis, 1992). Typically the overall concept of multicultural counseling competency is seen as comprised of the three areas of awareness, knowledge and skills.

The majority of literature on multicultural counseling training suggests that counseling programs need to be modified in order for trainees to obtain the necessary awareness, knowledge, and skills to realize multicultural competence. However, a consensus has not been reached regarding the most appropriate or effective means for promoting counselors’ multicultural development (D’Andrea, Daniels, and Heck, 1991). This lack of consensus reflects the newness of this area, and the lack of empirical support for any particular approach. However, training efforts focused on skill acquisition are typically structured in such a way that students are directly involved in concrete, action-
oriented learning activities, such as role-playing, behavioral modeling, and critiquing multicultural counseling interactions. In contrast, training focused on increasing multicultural awareness is often less structured and uses classroom discussion to explore students’ own stereotypes, prejudices, and misconceptions regarding members of cultural groups other than their own. Training efforts focused on multicultural knowledge acquisition typically emphasize the accumulation of large amounts of information through reading, lectures, and presentation of factual data regarding various cultural, ethnic, or racial groups (D’Andre et al., 1991). Opportunities to experience direct contact, within a counseling context or otherwise, with individuals from cultures different from the trainees, have also been encouraged to help enhance all areas of multicultural competence. Most notably such direct contact has been used to enhance skill when in a cross-cultural counseling context, and awareness when outside of the counseling context. Direct intergroup contact experiences are often more difficult to provide and control than the other experiences mentioned, which can readily be provided in any classroom. Therefore, intergroup experience has been less widely utilized as part of multicultural counselor training. However, the broader literature suggests that intergroup contact is critical in multicultural understanding. Thus, while multicultural training and general multicultural education efforts have primarily focused on more formal education approaches, efforts to structure intergroup contact approaches are also promising and deserve attention.

To formally provide multicultural education and training: Two basic approaches are used, the single course and infusion formats (Carey, Reinat, & Fontes, 1990). For many programs, the single-course format is the easiest to implement and most resource
efficient. It allows for concentrated focus on multicultural issues and thus avoids the risk of watering down these issues in the midst of trying to teach other materials. Also, the single course method typically encounters less resistance. It does not require changes in long standing courses and curriculum, it requires little coordination among trainers, and those most committed to multicultural training can teach it. Further, students and instructors not truly interested in or supportive of multiculturalism, or otherwise reluctant regarding multiculturalism, can avoid it. Or, if made mandatory, reluctant students need only survive a single course, and reluctant instructors can still typically avoid the entire topic. Because it is the easiest to implement and receives the least resistance, the single course has been the most widely implemented format both in multicultural counseling psychology training and general multicultural education programs. A survey showed that in 1977 less than 1% of counseling psychology training programs required any courses focused on multicultural issues. By 1992, 89% of the programs required at least one multiculturally focused course in their training format (handbook of MC counseling p20).

The infusion format calls for instructors to integrate multicultural information and perspectives into numerous, if not all, courses in the total curriculum. This approach is believed to present students greater opportunity to develop a wider range of multicultural understanding in multiple contexts (D’Andrea, Daniels, & Heck, 1991). However, for the reasons cited above, this is a much more difficult format to implement and is therefore less frequently utilized than the single course method.

Little research has been conducted which examines the relative merits of either of these approaches. D’Andrea et al. (1991) found preliminary support for the single course method in improving multicultural counseling awareness, knowledge, and skills of
graduate students. No studies have been located which evaluate either the effectiveness of the infusion format in isolation or the effectiveness of the two approaches in combination.

An often less formal approach to providing multicultural education and training involves programmed intergroup contact formats. Programmed intergroup contact formats require a designed intergroup contact experience. This often takes the form of an outside class assignment or a multicultural training group. Either way, this is a much more difficult format to design or control than either the single course or infusion formats. First, it requires access to groups of diverse participants (i.e. – of different race, ethnicity, sexual orientation etc), often not readily available in sufficient numbers in either counseling psychology training programs or other university settings. Second, it is harder to control since the experience depends on the participation of numerous individuals beyond just the instructor.

The other form that utilizes intergroup contact is in a multicultural training group. Multicultural training groups are typically a little more formal and controllable than experiences outside of the classroom, but still much more difficult to design and implement than more formal education approaches. However, because group interactions are the best means for providing broader intercultural contact, group interventions of this kind are of particular interest in efforts to increase multicultural awareness. Parker and McDavis (1979) discussed the utility of group interactions as a means of heightening multicultural awareness. D’Andrea et al. (1991) advocated the use of less structured classroom discussion to assist trainees in examining their biases and stereotypes about groups different from their own. Sabnani, Ponterotto and Borodovsky (1991) proposed
that in order to help in the development of multicultural awareness, encounter groups could be used to challenge trainees’ biases and stereotypes.

The lack of structure inherent in the group method provides a vehicle for spontaneous in-vivo interactions. Multicultural training groups provide students with experiences with members of other cultures without the inherent responsibility or power differential that may characterize other relationships (e.g. counseling or more formal classroom relationships). Students are afforded an opportunity for direct feedback about the impact they have on others, and can explore and express the impact that others have on them. The presence of group leaders provides a context of safety and support in which relevant issues can be addressed. Sabnani et al. (1991) allude to training groups as a place where anger and difficult feelings regarding cross-cultural interactions can be worked through successfully.

Overall, group interventions seem like a promising way to facilitate multicultural self-awareness and education at the college and university level. But, despite the promise associated with using groups as a place for positive and effective multicultural education, there has been limited empirical research to support its efficacy. However, while the narrow or structured use of intergroup contact as a part of formal multicultural counseling training or wider multicultural education at the university level may be new and lack empirical support, the value of intergroup contact more broadly defined has a long history of theory, research, and support.

**Summary**

There is a major need for research that focuses on ways to reduce prejudice and discrimination between members of different racial and ethnic groups, particularly
between majority and minority group members. There are many complex factors that contribute to the development of prejudice. In order to counteract these factors, racism and prejudice must be addressed on multiple levels. Multicultural education programs are both pervasive and fraught with challenges (Betz & Fitzgerald, 1993). There has been little empirical research to date which examines either the efficacy of these efforts, or the processes involved. One area that has received a great deal of attention and empirical support is intergroup contact. While there have been mixed results from both individual studies and overall reviews of the literature on this topic, it appears that intergroup contact under optimal conditions definitely has an inverse relationship with prejudicial attitudes and actions. There is also support for the value of promoting empathy as a way to improve intergroup relations. Experiential opportunities, such as group interactions, seem a more promising and natural way to promote empathy than traditional didactic methods. Surprisingly, the literature on intergroup contact largely focuses on either majority or minority groups at the exclusion of the other. This does not allow for any examination of the dynamic process involved between the two, on which contact theory is largely based. Finally, there is a clear call for programmed intergroup contact efforts as a primary intervention against prejudice and discrimination.

Counseling psychology training programs can provide important experience and information regarding how to best implement such an effort. Further, counseling psychologists are uniquely committed and trained to design and facilitate such efforts.

The Current Study

The following study attempts to address several areas identified as limitations in the current literature by investigating a “real world” programmed intervention effort that
is primarily based on intergroup contact theory. The overall goal of the intervention is to 
provide significant positive racial intergroup contact. The study investigates whether 
multicultural awareness and/or attitudes regarding race changed over time for those 
participating in this particular intervention. Further, exposure to factual (i.e. – articles 
chose to provide facts regarding racial issues) or personal (i.e. – autobiographical articles 
chosen to promote empathy) information through assigned readings was manipulated to 
examine the effects. The process involved in the intervention was also explored, and 
outcome and process data examined both quantitatively and qualitatively for both 
majority and minority participants together in context versus in isolation. Finally, the 
intervention was critically assessed to determine its appropriateness as a vehicle for 
university counseling centers and counseling psychologists to proactively combat racism 
and discrimination, thus, shedding the role of status quo agents, and instead embracing 
the role of social advocates.
CHAPTER THREE

Method

The purposes of this study were (a) to investigate whether multicultural awareness or racial attitude changed over time for students participating in a specific nine-week process-oriented group intervention, Dialogues on Race (DOR); (b) to examine whether multicultural awareness development or prejudicial attitudes are affected differently as a result of exposure to readings providing personal experience (e.g. – one person’s account of growing up Chicano), or factual data (e.g. – an article providing factual data about the differential treatment experienced by Latinos in the judicial system); and (c) to explore what incidents, interactions, or other factors affected the processes involved in this particular intervention, and how (i.e. – gather exploratory data through observation, critical incident reports, and interviews to gain greater understanding of the process). This chapter provides information regarding the participants, intervention, research design, instruments, and procedures used in the study.

Participants

Data were initially collected from 98 undergraduate and graduate students at The Pennsylvania State University during the spring 2002 semester. However, due to participant dropout (i.e. – failure to complete and return all required pre and posttest data), only 73 students are included in the final analyses. Of those 73, 25 served as a control group (Condition 0) and did not participate in any form of the intervention. The other 48 participated in the nine-week DOR intervention program. The students participating in the intervention program were assigned to 1 of 6 groups. The groups were split 3 each into 2 experimental conditions. The experimental conditions differed
based on exposure to different types of assigned readings. In the first condition, participants were required to read articles regarding factual information about race or differences in experience based on race, while participants in the second condition were required to read articles written by individuals sharing their very personal experiences related to race. Students in both conditions were also required to write a reaction/response paper for each assigned article. Primarily through group supervision, efforts were made to keep all other intervention activities as similar as possible (see supervision later in this chapter for details). Of the 48 students in the final data set, 25 were in groups exposed to factual readings (Condition I), and 23 were in groups exposed to personal readings (Condition II).

*Experimental groups.* All 60 of the students participating in the intervention agreed to participate in the study and turned in informed consent, pre-intervention and weekly data. However, five turned in pretest materials with required data missing or with out of range responses. Additionally, nine failed to turn in post-intervention materials. Thus, complete data sets were provided by 48 of the students participating in the intervention, only these 48 are included in the final data analyses. Of those 48 students, 25 were in Condition I, and 23 in Condition II. One of the students from Condition I failed to provide any of the requested demographic data, but was kept in the final data set since all other required data were provided. Therefore, of the 48 students participating in the intervention, demographic information were provided by 24 from Condition I and 23 from Condition II. Reported ages ranged from 18 to 24 in Condition I and 18 to 29 in Condition II, mean age was 20.17 for both Conditions I and II, median age was 20 for
both Conditions I and II, and standard deviation was 1.239 for Condition I and 2.289 for Condition II (see Table 1).

Table 1 Age: Means, Median, Standard Deviations, and Ranges for the Comparison Group and the 2 Experimental Conditions

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Median</th>
<th>SD</th>
<th>Min - Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comparison group ( (n = 25-1) )</td>
<td>20.88</td>
<td>21</td>
<td>2.193</td>
<td>18 - 27</td>
</tr>
<tr>
<td>Experimental Condition I: factual readings ( (n = 25-1) )</td>
<td>20.17</td>
<td>20</td>
<td>1.239</td>
<td>18 - 24</td>
</tr>
<tr>
<td>Experimental Condition II: personal readings ( (n = 23) )</td>
<td>20.17</td>
<td>20</td>
<td>2.289</td>
<td>18 - 29</td>
</tr>
<tr>
<td>Total: All participants: ( (n = 73-2) )</td>
<td>20.41</td>
<td>20</td>
<td>1.961</td>
<td>18 – 29</td>
</tr>
</tbody>
</table>

Of the students in Condition I, 7 identified as male and 17 female. Of the students in Condition II, 5 identified as male and 18 female. Reported class standings were 4 first year in each Condition I and II, 2 second year in Condition I and 8 in Condition II, 13 third year in Condition I and 5 in Condition II, 4 fourth year in both Conditions I and II, no fifth year in Condition I and 1 in Condition II, no sixth or more year in Condition I and 1 in Condition II, 1 graduate student in Condition I and none in Condition II. When asked to choose which of 7 categories best described their race, ethnicity, or cultural background; 5 in Condition I and 7 in Condition II identified themselves as “Black;” 2 in Condition I and 0 in Condition II “Latino/Hispanic;” none in either Condition I or II “Native American/Alaskan;” 7 in each Condition I and II “White;” 2 in Condition I and 3 in Condition II “Asian/Pacific Islander;” 2 in Condition I and 4 in Condition II “Bi-racial;” 6 in Condition I and 2 in Condition II “Other.” Upon reviewing the additional descriptions provided by students choosing “Other,” it was noted that 5 in Condition I
and 1 in Condition II identified themselves as “Middle Eastern,” or as from a specific Middle Eastern country. Thus, a decision was made to add “Middle Eastern” as a separate category. Five students in Condition I and 1 in Condition II fit this new category. Of the remaining two students choosing “Other,” one identified as being from India, and a decision was made to include her with the other students who identified as East Indian, in the Asian/Pacific Islander category. The final student choosing the “Other” category listed multiple racial and ethnic groups with which she identified. The decision was made to expand the “Bi-racial” category to include her. As a result, Condition I added a student to the “Bi-racial” category increasing the total to 3, and Condition II added a student to the “Asian/Pacific Islander” category increasing that total to 4 (see Table 2).

Comparison group. Data were also initially collected from 38 undergraduate and graduate students who were to serve as a control group and would not participate in the intervention. Of the 38 students who initially returned forms to be included in the study as part of the control group, 13 failed to return the necessary follow-up materials or returned incomplete data, as a result there are 25 students in the final control group.

The reported ages for the control group (one student failed to provide her age) ranged from 18 to 27, with a mean age of 20.88, median 21, standard deviation of 2.193 (Table 1). Reported class standings were: 3 first year, 4 second, 8 third, 6 fourth, 1 fifth, 2 sixth or more, and 1 graduate student. When asked to choose which of 7 categories best described their race, ethnicity, or cultural background, 3 identified as “Black,” 2 “Latino/Hispanic,” none “Native American/Alaskan,” 15 “White,” 2 “Asian/Pacific Islander,” 2 “Bi-racial,” and 1 “Other.” The student who chose “Other” listed multiple
racial and ethnic groups with which he identified. Consistent with the decision for the
students participating in the intervention, the “Bi-racial” category was expanded to
include him. As a result, the “Other” category was dropped, and the “Bi-racial” category
increased from 2 to 3 (see Table 2).

Table 2

| Summary of Frequencies of Nominal and Ordinal Variables by Condition and Totals |
|---------------------------------|-----------------|-----------------|-----------------|-----------------|
| Control (0) | Factual (I) | Personal (II) | Total |
| n=25       | n=25-1      | n=23           | N=43-1 |
| frequency | frequency  | frequency      | frequency      | frequency      |
| gender   | %          | %              | %              | %              |
| Male     | 7 28%      | 7 28%          | 5 21.7%        | 19 26%         |
| Female   | 18 72%     | 17 68%         | 18 78.3%       | 53 72.6%       |
| race/ethnicity | White | 15 60% | 7 28% | 7 30% | 29 39.7% |
|           | 10 40%     | 17 68%         | 16 69.6%       | 43 58.9%       |
|           | 3 12%      | 5 20%          | 7 30.4%        | 15 20.5%       |
|           | 2 8%       | 2 8%           | 0 0            | 4 5.5%         |
|           | 2 8%       | 2 8%           | 4 17.4%        | 8 11.0%        |
|           | 3 12%      | 3 12%          | 4 17.4%        | 10 13.7%       |
| biracial | 0 0        | 5 20%          | 1 4.3%         | 6 8.2%         |
| class standing | 1st year | 3 12% | 4 16% | 4 17.4% | 11 15.1% |
|            | 4 16%      | 2 8%           | 8 34.8%        | 14 19.2%       |
|            | 8 32%      | 13 52%         | 5 21.7%        | 26 35.6%       |
|            | 6 24%      | 4 16%          | 4 17.4%        | 14 19.2%       |
|            | 1 4%       | 0 0            | 1 4.3%         | 2 2.7%         |
|            | 2 8%       | 0 0            | 1 4.3%         | 3 4.1%         |
|            | 1 4%       | 1 4%           | 0 0            | 2 2.7%         |

The Intervention

The ‘Dialogues’ program consists of two major components. The first component
was the actual intervention in this study, and consisted of the process-oriented DOR
groups. These groups were experimentally manipulated for the purpose of this study.
The second component is the training and supervision of the group facilitators, which for the purposes of this study helped to ensure the quality of the intervention and minimize differences between groups.

The DOR groups represent a programmed attempt at reducing racial prejudices and discrimination through facilitated intergroup contact. The basic concept and design for the DOR groups were based on general intergroup contact theory (Pettigrew & Tropp, 2000). The primary objective of the DOR groups was to bring racially and ethnically diverse college students into contact with one another, for a sustained period of time, in a manner designed to facilitate open and honest conversation, inquiry, and self-reflection. The goal was to promote intergroup contact meeting all of the conditions noted by Allport (1954) as necessary for positive outcomes, plus the additional condition of potential friendship hypothesized by Pettigrew as also necessary.

The DOR groups were voluntary discussion groups offered and supervised through Counseling and Psychological Services (CAPS) at The Pennsylvania State University. Independent study course credit was available for participants through the Counselor Education program. During the semester in which this study was conducted, six groups were co-facilitated by students enrolled in either counseling or clinical psychology graduate programs, and supervised by CAPS staff.

Students were invited to participate in the DOR groups through an invitation letter sent out via email, posted on campus, and read by some course instructors (Appendix A). In an effort to ensure racial and ethnic diversity, advertisement for the groups was targeted to reach as diverse a student population as possible. The DOR groups were advertised at a number of diversity related events and directly to diverse student
organizations via email (e.g., Black Caucus, Latino Caucus, Asian American Student Coalition, Black Graduate Student Association, University Housing Resident Assistants, Blend Of Traditional Heritages – a bi-racial student discussion group, etc.). Interested students returned an application (Appendix A) requesting their age, gender, race/ethnicity, and major area of study. Additionally, all six DOR group section times were listed and students were asked to rank order their top three preferred times.

Participants were chosen and grouped based on three primary considerations. The first consideration was the order in which applications were received, with preference given on a first come first included basis. The second consideration was the number of participants in each group, as well as balancing the makeup of each group so that no single race or ethnicity was disproportionately represented. Further, efforts were made to ensure that no participant was the sole representative of his or her reported race or ethnicity in a given group. A final consideration was preference of meeting time, with efforts made to give applicants as highly ranked a time as possible given the other considerations.

*Group process.* The semester of this study, each DOR group met for a total of nine sessions, two hours per session, one session per week. In an effort to provide structure, a safe vehicle to start the dialogues, and consistency across groups, all groups started the session with identical opening exercises for the first five weeks of the nine week program. Beyond the opening exercises, groups were largely unstructured and facilitators were allowed the freedom to facilitate their individual groups in a manner that allowed them to evolve naturally. However, all facilitators were instructed, trained, and
supervised to keep the focus on honesty, race, and the interpersonal processes in the room.

*Group facilitators.* The semester of this study, all six sections of the DOR groups were co-facilitated by graduate students in either the APA-accredited Counseling or Clinical Psychology program. For the five weeks prior to the start of the DOR groups, all facilitators participated in two hours weekly training sessions specifically focused on leading the DOR groups. Additionally, all facilitators met weekly for two hours of supervision the entire nine-week duration of the DOR groups.

Serving as a DOR group facilitator was voluntary. Facilitators were recruited from student trainees at CAPS (i.e. – pre-doctoral interns, graduate assistants, and practicum students), as well as graduate students enrolled at the university in counseling or clinical psychology. It was a pre-requisite that all facilitators had completed a minimum amount of course work related to group facilitation and the sociology of race, class, and gender issues. An equal number of the 12 facilitators represented three broad racial/ethnic groups: Four Latinos/Hispanics, four Blacks, and four Whites. Gender representation was also equivalent with six women and six men. However, gender representation was not equivalent across racial groups. The Hispanic and White facilitators had a 3:1 male to female ratio, whereas all four of the Black facilitators were women. Facilitators were matched to allow for maximum diversity regarding race, ethnicity, and gender. One male and one female facilitator were assigned to each group, and no co-facilitator pair identified with the same racial or ethnic group. The two DOR program administrators, a black male and a white female, were employed by CAPS and supervised the group facilitators throughout the study.
Procedure Experimental Groups

Participants were recruited based on the status of their participation in the intervention under investigation, the DOR groups. All students participating in DOR groups were encouraged, but not required, to participate in this research study. In addition, students not participating in the intervention were recruited to serve as a control.

Recruitment (experimental groups). Participation in this study did not require anything additional to what was expected from all participants in the dialogue groups. Therefore, since providing the data used in the study was a requirement of participating in the groups, independent of this study, every group was provided time during group sessions to complete all evaluations/assessments used in this study. All students participating in the groups engaged in the same activities and provided the same information as part of participation in the groups, however the data they provided were not analyzed or otherwise included in the study without that student’s explicit written consent. It was explained to students that participation in the study was not a requirement for the course, and choosing to allow their responses to be included in the data analysis or not would have no effect on their grade for the course. Further, it was explained that forms indicating informed consent were to be returned in a sealed envelope, and the primary investigator, who was ultimately responsible for the assignment of grades, would not be aware of which students chose to participate until after final grades were assigned. This procedure was followed.

Students participating in the groups were told that the purpose of the research was to help evaluate the efficacy and explore the process of the DOR groups, and that a more detailed explanation of the study would be made available following completion of the
study. Further, it was explained that study results would be available once analyses were completed. Participants were asked not to discuss any aspects of the research until after completion of their involvement in the study.

**Pretest data collection (experimental groups).** At the start of the first meeting for each dialogue group, all students were given a research pre-intervention packet inside a manila return envelope. Each packet contained two explanations of the study and informed consent forms (Appendix B). They were instructed to keep one copy of this form for their personal records, and to return the other sealed in the envelope provided. A personal data sheet (Appendix C), and three paper and pencil self-report assessment instruments (Appendixes D, E, F) were also included in the pre-intervention packet. The sequence of the assessment instruments in the packets was randomly assigned to control for order effects.

The explanation and consent forms were read and explained to each group by the primary investigator, and participants were allowed time at the start of the group session to complete the study materials. It took participants approximately 15-25 minutes to complete all of the pre-intervention materials. All students, regardless of whether or not they chose to participate in the study, were instructed to return research materials to the primary investigator in the envelopes provided prior to leaving the group location.

To maintain confidentiality, the primary investigator did not open returned study materials. Instead, a pre-selected contact person (i.e. – an undergraduate student employed at CAPS in a work study position) not otherwise involved in the research, student evaluations, course grades, or facilitation of the groups, opened all materials. That contact person kept all returned consent forms, and she was the only person with the
information directly matching students to their research materials. None of the group facilitators, the course instructor, or research investigators accessed the returned research materials until after the final grades had been determined and submitted for the semester.

Students participating in the groups who failed for any reason to attend the first group session were contacted via email and instructed to pick up and complete the pre-intervention packet prior to attending their first session. Packets were made available for them at the reception desk at CAPS, and they were provided space to complete the forms. They were instructed to seal all forms in the manila envelope provided and returned them to the reception staff. This sealed envelope was then given to the person previously identified to keep the consent forms and other research materials.

*Weekly data collection (experimental groups).* In addition to collecting the pretest-posttest data described above, weekly data were also collected. This weekly data was collected at the very end of each session and included a co-leader-rating questionnaire (later dropped due to low reliability), and a session evaluation questionnaire (Appendix G). Qualitative data were also gathered weekly through the addition of a single open-ended question requesting that the student identify and write down the most meaningful single event or interaction for them in that particular session, at the bottom of the session evaluation questionnaires. Ten minutes were devoted at the end of each session for participants to complete these weekly questionnaires. Participants were instructed to place the forms in a manila envelope provided by the group facilitators prior to leaving each session. Again, none of these envelopes were opened or reviewed by group facilitators or research investigators prior to completion of the intervention.
Posttest data collection (experimental groups). During the last 25 minutes of the final group session for the semester, all students were given a research post-intervention packet inside a manila return envelope. Each packet contained a course evaluation sheet (Appendix H), and the same three assessment instruments given in the pre-intervention packets (Appendixes D, E, F). The sequence of the assessment instruments was again randomly assigned to control for order effects.

Participants were allowed the final 25 minutes of the final group session to complete the posttest research materials. They were instructed to return all research materials in the provided return envelope to the group facilitators immediately prior to leaving the group location. Students unable to attend the final group session for any reason were contacted via email by the group facilitators and informed that research packets would be available at the CAPS front desk for five days. They were further instructed to pick up, complete, and return their research packet during that time frame.

Further qualitative data were collected in the form of semi-structured open-ended individual interviews. These interviews were conducted after completion of the intervention. The primary investigator conducted semi-structured open-ended individual interviews with two participants from each of the six groups. Participants for the interviews were identified by the two facilitators from each group as having in their perceptions, gained the most and the least from participation in the intervention. These interviews were conducted the week after the final group meetings, and were not required or expected as part of participation in the intervention. Final grades had already been submitted, and information regarding final grades was given to each student interviewed prior to their interview. Students were informed that interviews would last between 30
and 45 minutes. As compensation for their time, participants agreeing to be interviewed received $5 each. All students recruited for interviews agreed to be interviewed, and all interviews were audio-taped. Informed consent separate from the consent given for the other portion of the study was obtained (Appendix B).

Procedure Comparison Group

In addition to the students participating in the DOR intervention groups, students were also recruited who would not participate the intervention to serve as a control. Participants were recruited from three primary locations, two large sections of a sociology course and one African-American and Latino leadership course. These courses were chosen in an attempt to recruit participants representing racial, ethnic, and other relevant characteristics similar to those participating in the Dialogue Groups. As incentive for participation, inclusion in a random drawing for three fifty-dollar prizes was offered.

Pretest data collection (control group). The same week that the intervention started, a total of eighty-nine pre-test packets were distributed to students not enrolled in the intervention, but who volunteered to participate in the study as a comparison group. The packets included two copies of a study explanation and informed consent form (Appendix B), and the same three instruments given to the study participants in the DOR group interventions (Appendix D, E, F). Thirty-eight (42.7%) of these packets were returned completed. Participants were also asked to provide addresses where they desired posttest packets to be sent at the end of the intervention.

Posttest data collection (control group). The same week that the experimental groups concluded, follow-up test packets were mailed to the addresses provided by the 38
control group participants. Eighteen completed posttest packets were initially returned in
the allotted time. Those 18 students were included in the random drawing for the three
$50 prizes as agreed. An additional posttest packet was then sent out to the twenty
students who had returned the completed pretest materials, but had not returned the
posttest packets. They were encouraged to complete and return the posttest packets, and
were offered $5 as incentive. Nine additional participants completed and returned the
posttest materials, for a total of 27 control group participants.

The Experimental Manipulation

Readings were initially identified by the primary investigator as related to race
and/or race relations, and appropriate as reading materials to be assigned to DOR group
participants. An additional criterion for the “personal” readings was that they be clearly
written from a first person perspective. Some articles were eliminated based on content
and some on length. Ultimately, six articles of roughly the same length were identified as
appropriate and primarily “personal” and affect oriented, and six were identified as
appropriate and primarily “factual.” These articles were then distributed to a group of
raters including: Two White and one Biracial (Black/White) psychologists, and one
White Sociology professor. They each rated the articles based on the following criteria:
(1) Overall appropriateness as an assigned reading for the DOR groups, (2) the articles
ability to either (a) relay a first person experience clearly focused on and influenced by
the racial identity of the author and provoking empathy, or (b) provide quality “factual”
information about race, race relations, or a particular racial group. One article was
eliminated because several raters felt there was some ambiguity as to whether it was
primarily “personal” or “factual” since it used several first person examples to support
the factual data provided. Another article was eliminated because three of the raters felt that it was too “academic,” “difficult to understand,” or “boring,” compared to the other readings. Ultimately, five “personal” and four “factual” articles were identified that were rated as appropriate by all raters. The three which seemed to best parallel an article from the other category were chosen to be assigned to the group participants, and thus used in the study (i.e. – The “personal” articles included one from an Asian-American woman, one from a Mexican-American man, and one from a Bi-racial (Black/White) man. The parallel “factual” articles were on the historical treatment of Asian Americans in America, treatment of Latinos in the U.S. justice system, and the scientific vagaries of racial categorization). Three of the discussion groups were then randomly assigned to the experimental conditions (i.e. – groups were numbered 1-6 and a coin was tossed to determine to which experimental condition odd and even numbered groups would be assigned). Odd numbered groups were assigned to read and write reaction papers on “factual” articles (Condition I) and the even numbered groups assigned the “personal” articles (Condition II). All participants were instructed to read the articles carefully and to write a 2-4 page double-spaced reaction paper. However, the students assigned the personal readings were given instruction to help generate empathy (e.g. – “While you are reading, try to image how you would feel if you were the writer…). The students assigned the factual readings were not given these additional instructions to help evoke empathy, but instead were instructed that their papers need to, “demonstrate that you have read the article, and also share your thoughts and reactions about what you have read.” Group members were asked not to discuss the articles or other content related
aspects of their group experiences. Facilitators also agreed not to discuss the content of the articles assigned to their groups.

**Supervision**

While it is understood that different groups will evolve differently depending on many factors including the make-up of the group and the style of the facilitators, efforts where taken to ensure as similar a group experience for all groups as possible, without drastically restricting the natural dynamic functioning of the groups. One way that consistency across groups was maintained was through two hours of weekly group facilitator training and supervision. During this time each week, group facilitators all shared questions and concerns about group facilitation and group dynamics. This ensured that all of the facilitators were at least conscious of the various concerns coming up across groups, and they collectively discussed appropriate ways to potentially intervene. The facilitators also discussed the overall progress and focus of the groups, and determined appropriate interventions to maximize the experience for the group members. For example, the facilitators collectively agreed on how much or how little structure was appropriate at various stages of the groups. The first three weeks, all of the groups started each week with the same predetermined and practiced opening exercise. This helped to maintain some consistency regarding at least where the groups started each week. The groups were then free to proceed in the directions most natural and appropriate for each particular group and its co-facilitators. Facilitators were instructed that these were to be primarily process-oriented groups, and as such, it was their primary goal to facilitate here-and-now conversation among the group members as much as possible (i.e. - versus a more didactic focus, or focus on outside events).
While group facilitators typically all met together weekly for training and supervision, there were three occasions when facilitators were broken into two smaller groups. One group was comprised of facilitators whose groups were assigned the “factual” readings (Condition I), and the other of facilitators whose groups were assigned the “personal” readings (Condition II). Since knowing the experimental conditions might affect the group leaders’ facilitation, they were not informed about the differences in the reading assignments. Further, they were all asked, and agreed, not to discuss the readings or anything related to the readings outside of their groups or in the regular supervision sessions with all of the facilitators present. However, in order to allow for appropriate discussion and supervision around issues that potentially related to the readings or reactions to them, facilitators were provided split group supervision (i.e. – the six facilitators in Condition I and the six facilitators from Condition II met in separate supervision groups for 45 minutes) following each of the three sessions when the readings and the reaction papers to the reading were discussed.

Design

The quantitative portion of this study was a quasi-experimental (due to a lack of randomization) field study employing a pretest-posttest control group design. Pretest-Posttest assessment instruments were identical (QDI, M-GUDS, and CoBRAS). In this design, there were three experimental conditions. First was the control condition which consisted of participants who where not exposed to any portion or form of the intervention. Participants in the other two conditions were exposed to the same nine-week intervention with one variation; the assigned readings were experimentally manipulated. The members of three of the groups were assigned readings identified as
“personal”, and members of the other three groups were assigned readings identified as “factual.” The design can be depicted as follows with O=observation, X=intervention, p=personal experience readings, and f=factual data readings:

<table>
<thead>
<tr>
<th>CONDITION</th>
<th>TIME 1</th>
<th>INTERVENTION</th>
<th>TIME 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental I</td>
<td>O1</td>
<td>Xf</td>
<td>O2</td>
</tr>
<tr>
<td>Experimental II</td>
<td>O3</td>
<td>Xp</td>
<td>O4</td>
</tr>
<tr>
<td>Control</td>
<td>O5</td>
<td></td>
<td>O6</td>
</tr>
</tbody>
</table>

Because different individuals co-facilitated each group, and different groups naturally evolve in different ways, weekly data were also collected to examine and potentially control for these between group differences. A leader rating assessment was collected from each group member following each group session. Plus, group participants and facilitators completed an assessment measuring their personal reactions to each particular session. The leader rating assessment used was a 24 question measure producing seven scale scores. The Cronbach’s reliability alphas for four of the seven scale scores fell below the .60 cut-off established for this research study. Thus, the entire measure was dropped from further analysis. However, the assessment measuring personal reactions to each session was utilized, and is described in detail later in this chapter.

Because quantitative and qualitative methods involve differing strengths and weakness, qualitative data were also collected as part of the study. Qualitative data were collected following each session from each group participant in the form of a brief open-
ended response regarding the most significant aspect of that session for him/her (i.e. – critical incident report). Further, posttest semi-structured individual interviews were conducted with two participants from each of the six groups: one participant from each group that the co-facilitators together assessed as having gained the most, and one participant from each group assessed to have gained the least from their participation in the intervention. This qualitative data were utilized to help gain greater understanding of the unique experiences of individual group members, and also to provide depth and description to those unique experiences.

Instruments

*Miville-Guzman Universality-Diversity Scale* (M-GUDS; Miville, 1992; Appendix D). This scale was used to measure attitudes of awareness and acceptance of both similarities and differences. This 45-item self-report instrument uses a 6-point Likert-type scale, ranging from 1 (*strongly disagree*) to 6 (*strongly agree*), to measure the respondents’ universal-diverse orientation (UDO). Fifteen items are reverse worded and scored. UDO is a construct that reflects an attitude of awareness and acceptance of both similarities and differences among people. The M-GUDS produces a single score ranging from 45-270, with higher scores reflecting greater awareness and acceptance of similarities and differences among people.

In four separate studies (*Ns* = 93, 111, 153, and 135), internal consistency and retest reliability for the M-GUDS scores ranged from .89 to .95. As theoretically predicted, the M-GUDS significantly correlated with measures of racial identity, empathy, healthy narcissism, feminism, androgyny, homophobia, and dogmatism (ref).
The M-GUDS was found not to correlate with Scholastic Achievement Test Verbal scores, but mixed results were obtained between M-GUDS and social desirability (ref).

Internal consistency was explored specifically for this study by generating Cronbach’s Alphas at both pretest and posttest for all instruments used. The alphas for the M-GUDS were .908 and .927 at pretest and posttest respectively. These findings are in line with the internal consistencies reported from previous studies.

*The Color-Blind Racial Attitudes Scale* (CoBRAS; Neville, Lilly, Lee, Duran, and Brown, 2000; Appendix E). This scale was used to measure color-blind racial attitudes and beliefs. The 20-item CoBRAS is a self-report instrument which uses a 6-point Likert-type scale, ranging from 1 (*strongly disagree*) to 6 (*strongly agree*), to assess cognitive aspects of the respondents’ color-blind racial attitudes. Ten items are reverse worded and scored. The CoBRAS produces scores on three factors: Factor I, Unawareness of racial privilege (seven items); Factor II, Unawareness of institutional discrimination (seven items); Factor III, Unawareness of blatant racial issues (six items).

Five studies (Neville et al., 2000) on the CoBRAS with over 1,100 observations provide initial reliability and validity data and support for the 3-factor solution. Alpha coefficients for each of the three factor scores and the total score were found to be acceptable: Factor I, .83; Factor II, .81; Factor III, .76, and total score .91. Intercorrelation among the three CoBRAS factors ranged from .42 to .59. The CoBRAS was found to be positively related to other indexes of racial attitudes (e.g., Quick Discrimination Index and The Modern Racism Scale), as well as two measures of belief in a just world (e.g., Global Belief in a Just World Scale and Multidimensional Belief in a Just World Scale). This indicates that greater endorsement of color-blind racial attitudes
is related to greater levels of racial prejudice and a belief that society is just and fair. A statistically significant decrease in CoBRAS scores was realized after completion of multicultural training instruction in a previous study, thus sensitivity to diversity training has been demonstrated.

Alpha coefficients for each of the three factor scores were generated for data used in this particular study at both pretest and posttest, and found to be acceptable: Factor I – pretest .834, posttest .701, Factor II – pretest .727, posttest .686, Factor III – pretest .680, posttest .741.

The Quick Discrimination Index (QDI; Ponterotto et al., 1995; Appendix F). This index was used to measure attitudes regarding racial equality. The QDI is a 30-item self-report instrument that uses a 5-point Likert-type scale ranging, from 1 (strongly disagree) to 5 (strongly agree), to assess participants’ racial and gender attitudes. Of the items, 15 are reverse worded and scored to control for order effects. The QDI produces three subscales assessing (a) general attitudes regarding racial diversity (9 items), (b) specific attitudes regarding contact and personal contact with racial diversity (7 items), and (c) general attitudes regarding women’s equality (7 items). A total instrument score can be tabulated using all 30 items; however, Ponterotto et al. recommended using a three-factor scoring procedure given that both exploratory and confirmatory factor analyses have supported tripartite conceptualization of the QDI items.

Five separate samples (Ponterotto et al., 1995; Utsey and Ponterotto, 1999), reported the following coefficient alphas: Factor I scores = .80 to .91; Factor II scores = .70 to .83; and Factor III scores = .65 to .77. Further, QDI scores were found to be relatively stable during a 15-week test-retest period. Stability coefficients for the QDI
scores across three samples were .82, .92, and .96 for Factor I; .65, .95, and .87 for Factor II; and .82, .78, and .84 for Factor II. The QDI factors were found to have negligible correlations with the Social Desirability Scale (SDS), indicating no social desirability contamination affecting scores.

Alpha coefficients for each of the three factor scores were generated for data used in this particular study at both pretest and posttest, and found to be acceptable: Factor I – pretest .773, posttest .711, Factor II – pretest .656 posttest .655, Factor III – pretest .712, posttest .735.

Session Evaluation Questionnaire (SEQ; Stiles and Snow, 1984; Appendix G). To measure participant and facilitator reactions to, and perceptions of, each session, the SEQ was used. This 24-item self-report instrument uses a 7-point Likert-type scale, ranging from 1 to 7, different anchor points are given for each question (e.g. – good/bad, rough/smooth, etc.), to help gain general insight into the session-by-session process. The SEQ consists of four scales. The Depth (5 items) and Smoothness (4 items) scales measure the participant’s reactions to characteristics of the session. The Positivity (5 items) and Arousal (4 items) scales measure post-session mood. Positivity is a measure of how positive or negative a participant feels upon completing the session. The Arousal scale is a measure of how much emotional arousal the participant feels after completing the session. The SEQ is one of the most widely used measures of participant reactions. It has also been utilized to measure both participant and counselor or facilitator reactions and perceptions of group sessions, and it has proven psychometric properties in these various conditions (Heppner, Kivlighan, and Wampold, 1992).
Course evaluation sheet (Appendix H). For exploratory purposes as well as general course evaluation and improvement, the participants completed a one-page course evaluation sheet (Appendix H) requesting participants’ opinions of the course, reading materials, and ways to improve the course. There was also a question to determine what, if any, other diversity training or structured multicultural experiences the participants were involved in during the same time period as this study. Information regarding diversity training or structured multicultural experiences during the time period encompassed by the study was also gathered from control group participants at the time posttest assessments were completed.

Interviews. In addition to the various quantitative instruments used to measure outcome and help facilitate understanding, post-intervention semi-structured open-ended interviews were conducted with individuals identified by the group facilitators as having gained the most or least from participation in the groups. The interviews were semi-structured in the sense that all interviews included the same set of predetermined general questions to insure that the same basic areas were covered in each interview, but the interviewer was free to ask any relevant follow-up questions at any time, and interviewees were encouraged to share any additional information they felt was a relevant part of the experience for them. A particular strength of this method is a practical flexibility not available with the quantitative instruments. It allowed for appropriate interview topics and questions to develop over the course of the intervention as observations and other data made them apparent, and also for the possibility of probing and extending responses. The purpose of gathering responses to open-ended questions was to enable the primary investigator to understand and capture the points of view of
other people without predetermining those points of view through prior selection of questions (Patton, 1990).
CHAPTER FOUR

Quantitative Research: Results and Discussion

The quantitative research component of this study was designed primarily to determine whether a particular intervention program (Dialogues on Race) was effective in positively influencing student participants’ level of awareness or attitudes related to racial, ethnic, or cultural diversity. Additionally, this study was designed to assess whether the type of readings assigned during the nine week intervention, either factual or personal (chosen to evoke empathy), would differentially influence any changes in awareness or attitude. The findings for these two main questions of the quantitative research portion of this study are presented and discussed below. Descriptive statistics are listed in Tables 1 and 2 in Chapter Three.

Research Question One

Were there significant changes in attitudes or levels of awareness related to racial, ethnic, or cultural diversity for students participating in the nine-week intervention (Dialogues on Race) compared to a control group of students who did not participate in the intervention? More specifically, do differences exist between students participating in the intervention and students not participating in the intervention, as measured by the differences between pretest and posttest scores on the three instruments (listed below) used in this study?

A. Attitudes regarding similarities and differences – total score on the Millville-Guzman Universality Diversity Scale (M-GUDS).
B. Color-blind racial attitudes – three factor scores on the Color Blind Racial Attitudes Scale (CoBRAS).

- Factor I: lack of awareness regarding racial privilege
- Factor II: lack of awareness regarding institutional discrimination
- Factor III: lack of awareness regarding blatant racial issues

C. Attitudes regarding racial equality – three factor scores on the Quick Discrimination Inventory (QDI).

- Factor I: general attitudes regarding racial diversity
- Factor II: specific attitudes regarding general contact and personal contact with racial diversity
- Factor III: general attitudes regarding women’s equality

Because all three instruments used in the analyses are somewhat conceptually related, results for all scores will be reported and discussed together (*See Table 3 for pretest and posttest means and standard deviations, plus change scores).
### Table 3  
Means, Standard Deviations and Change Scores for the Comparison Group and the 2 Experimental Conditions

<table>
<thead>
<tr>
<th></th>
<th>Comparison group (n = 25)</th>
<th>Pretest</th>
<th>Posttest</th>
<th>Change M</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(M-GUDS) - Attitudes regarding similarities and differences</td>
<td>202.32, 20.24</td>
<td>203.60, 24.54</td>
<td>1.28</td>
</tr>
<tr>
<td></td>
<td>(CoBRAS factor 1) - Lack of awareness regarding racial privilege</td>
<td>23.20, 6.27</td>
<td>21.52, 5.48</td>
<td>-1.68</td>
</tr>
<tr>
<td></td>
<td>(CoBRAS factor 2) - Lack of awareness regarding institutional discrimination</td>
<td>21.76, 6.04</td>
<td>21.76, 5.98</td>
<td>.00</td>
</tr>
<tr>
<td></td>
<td>(CoBRAS factor 3) - Lack of awareness regarding blatant racial issues</td>
<td>11.64, 3.50</td>
<td>12.52, 3.61</td>
<td>.88</td>
</tr>
<tr>
<td></td>
<td>(QDI factor 1) - General attitudes regarding racial diversity</td>
<td>33.56, 5.07</td>
<td>33.32, 5.13</td>
<td>-.24</td>
</tr>
<tr>
<td></td>
<td>(QDI factor 2) - Specific attitudes regarding general contact and personal contact with racial diversity</td>
<td>26.52, 3.91</td>
<td>26.40, 4.08</td>
<td>-.12</td>
</tr>
<tr>
<td></td>
<td>(QDI factor 3) - General attitudes regarding women’s equality</td>
<td>26.88, 4.10</td>
<td>26.20, 4.60</td>
<td>-.68</td>
</tr>
<tr>
<td></td>
<td>Experimental Condition 1: factual readings (n = 25)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(M-GUDS) - Attitudes regarding similarities and differences</td>
<td>216.24, 18.25</td>
<td>217.72, 19.95</td>
<td>1.48</td>
</tr>
<tr>
<td></td>
<td>(CoBRAS factor 1) - Lack of awareness regarding racial privilege</td>
<td>18.48, 5.73</td>
<td>17.56, 5.24</td>
<td>-.92</td>
</tr>
<tr>
<td></td>
<td>(CoBRAS factor 2) - Lack of awareness regarding institutional discrimination</td>
<td>18.44, 4.77</td>
<td>18.36, 4.99</td>
<td>-.09</td>
</tr>
<tr>
<td></td>
<td>(CoBRAS factor 3) - Lack of awareness regarding blatant racial issues</td>
<td>10.44, 3.42</td>
<td>9.32, 4.15</td>
<td>-1.12</td>
</tr>
<tr>
<td></td>
<td>(QDI factor 1) - General attitudes regarding racial diversity</td>
<td>37.00, 3.44</td>
<td>37.40, 3.58</td>
<td>.40</td>
</tr>
<tr>
<td></td>
<td>(QDI factor 2) - Specific attitudes regarding general contact and personal contact with racial diversity</td>
<td>29.80, 2.87</td>
<td>29.28, 3.17</td>
<td>-.52</td>
</tr>
<tr>
<td></td>
<td>(QDI factor 3) - General attitudes regarding women’s equality</td>
<td>26.24, 4.59</td>
<td>26.00, 4.70</td>
<td>-.24</td>
</tr>
<tr>
<td></td>
<td>Experimental Condition 2: personal readings (n = 23)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(M-GUDS) - Attitudes regarding similarities and differences</td>
<td>222.00, 19.30</td>
<td>226.09, 21.57</td>
<td>4.09</td>
</tr>
<tr>
<td></td>
<td>(CoBRAS factor 1) - Lack of awareness regarding racial privilege</td>
<td>19.57, 7.32</td>
<td>17.26, 5.05</td>
<td>-2.31</td>
</tr>
<tr>
<td></td>
<td>(CoBRAS factor 2) - Lack of awareness regarding institutional discrimination</td>
<td>16.43, 5.32</td>
<td>17.17, 6.03</td>
<td>.74</td>
</tr>
<tr>
<td></td>
<td>(CoBRAS factor 3) - Lack of awareness regarding blatant racial issues</td>
<td>10.65, 4.22</td>
<td>9.83, 3.66</td>
<td>-.83</td>
</tr>
<tr>
<td></td>
<td>(QDI factor 1) - General attitudes regarding racial diversity</td>
<td>37.65, 4.82</td>
<td>37.26, 4.31</td>
<td>-.39</td>
</tr>
<tr>
<td></td>
<td>(QDI factor 2) - Specific attitudes regarding general contact and personal contact with racial diversity</td>
<td>28.78, 4.08</td>
<td>29.22, 3.91</td>
<td>.43</td>
</tr>
<tr>
<td></td>
<td>(QDI factor 3) - General attitudes regarding women’s equality</td>
<td>27.30, 4.05</td>
<td>28.04, 4.32</td>
<td>.74</td>
</tr>
</tbody>
</table>
To answer question one, three separate analyses were performed, one for each instrument used in the study. Since the M-GUDS produces a single score, an analysis of covariance (ANCOVA) was performed. However, the CoBRAS and QDI each produce multiple factor scores (3-each), thus multivariate analysis of variance (MANCOVA) were performed for each of these instruments. There are multiple reasons why ANCOVA and MANCOVA, with pretest scores as a covariate, is the preferred method for analysis of pretest-posttest data (Dimitrov and Rumrill, 2003). In this particular study, the main purpose of using the pretest scores as a covariate is to adjust the posttest means for differences among groups on the pretest, which are likely to occur since the study uses a nonrandomized design.

Assumptions of Analysis of Covariance and Multivariate Analysis of Covariance

One Analysis of Covariance (ANCOVA) and two Multivariate Analysis of Covariance (MANCOVA) were performed to answer question one. Prior to performing these analyses, the assumptions underlying ANCOVA and MANCOVA were tested.

In performing ANCOVA, the variance is analyzed to determine whether a difference exists between the group variances. The underlying assumptions of ANCOVA are that a normal distribution of the data exists and that the variance is homogenous (Field, 2000). The assumption of univariate normality for the dependent variable was checked using the Kolmogorov-Smirnov test. Results indicated that the distribution was not significantly different from a normal distribution. Further, the results from Levene’s test of equality of error variances was non-significant ($F = .628$, $p = .537$), indicating that the group variances are roughly equal. An additional assumption underlying ANCOVA is the assumption of homogeneity of regression slopes (Field, 2000). This assumption
was tested by running the ANCOVA as a customized model including only the interaction between the covariate and dependent variable. The significance value of the covariate by dependent variable interaction was highly significant ($p < 0.001$), indicating that the assumption of homogeneity of regression slopes has been broken. This raises concern about the outcome for the main analysis, particularly if the findings suggest only marginal significance of the dependent variable.

MANCOVA has similar assumptions to ANCOVA but extended to the multivariate case. However, since the assumption of multivariate normality cannot be checked on SPSS, the assumption of univariate normality for each dependent variable was checked using the Kolmogorov-Smirnov test. Results indicated that the distribution was significantly different from a normal distribution for at least one group on five of the six dependent variables (3-CoBRAS and 3-QDI). Because univariate normality is a necessary condition for multivariate normality, this indicates that the assumption of multivariate normality has been violated (Field, 2000). Levene’s test of equality of error variances was not significant for any of the dependent variables. However, since Levene’s test does not take account of the covariances, the variance-covariance matrices were compared between groups using Box’s test of equality of covariance matrices, and also found to be not significant. Therefore we can be reasonably sure that the underlying assumptions of MANCOVA (other than multivariate normality) have been met.

Results

The results from the ANCOVA performed to assess for changes in participants’ attitudes regarding similarities and differences, as demonstrated by pretest and posttest scores on the M-GUDS, supported the null hypothesis that no significant differences
would be found between any of the different groups of students at the .05 significance level (F=933, p=.398).

The results from the MANCOVA performed to assess for changes in participants’ attitudes regarding racial equality, as demonstrated by pretest and posttest scores on the 3-factors of the QDI, also supported the null hypothesis that no significant differences would be found between the different groups of students on any of the factors at the .05 significance level using Roy’s Largest Root statistic (F=1.511, p=.219). Because of the small sample sizes and expected subtlety between possible differences, Roy’s largest root statistic was used to determine significance. Roy’s root represents the maximum possible between-group difference given the data collected, and is the most powerful when group differences are concentrated on the first variate (Field, 2000).

The MANCOVA performed to assess for changes in the participants’ color-blind racial attitudes, as demonstrated by pretest and posttest scores on the CoBRAS, rejected the null hypothesis that no significant differences would be found between the groups of students at the .05 significance level using the Roy’s Largest Root statistic (F=2.987, p=.037).

Since MANCOVA protects only the dependent variable for which group differences genuinely exist, a Bonferroni correction was applied to the follow up ANCOVAs (Field, 2000 & Harris, 1975). The results of the follow up univariate tests indicate that there was a non-significant difference between groups for CoBRAS factors I and II, with (F = 1.730, p = .185) and (F = .148, p = .863) respectively. However, the univariate test indicated a significant difference between groups in terms of CoBRAS factor III (F = 4.485, p = .015), lack of awareness regarding blatant racial issues. Further
follow up contrasts also found no significant difference between either of the intervention groups (factual or personal) when compared to the no-treatment control on either CoBRAS factors I or II. However, a statistically significant difference was found between both intervention groups when compared to the control group on factor III (See Table 4).

Table 4
Contrast Results

<table>
<thead>
<tr>
<th>CoBRAS (95% confidence interval)</th>
<th>Factor 1</th>
<th>Factor 2</th>
<th>Factor 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factual versus Control</td>
<td>p = .280</td>
<td>p = .677</td>
<td>p = .005*</td>
</tr>
<tr>
<td>Personal versus Control</td>
<td>p = .068</td>
<td>p = .607</td>
<td>p = .028*</td>
</tr>
</tbody>
</table>

Discussion

No significant differences were found regarding attitudes or level of awareness related to racial, ethnic, cultural diversity, or gender equity for students participating in the intervention (Dialogues on Race) compared to a comparison group of students who did not participate in the intervention as measured by either the M-GUDS or QDI. However, a MANCOVA did indicate a significant difference between students participating in the intervention and the comparison/control group as measured by the CoBRAS.
Field (2000) strongly recommends that follow up to a significant MANOVA (or in this case MANCOVA) include both univariate tests and discriminant analysis to more fully understand the data. Both the univariate tests and the discriminant analysis suggest that the group separation can be best explained in terms of one underlying dimension. In this context the dimension is likely Awareness of Blatant Racial Attitudes (CoBRAS factor III). A comparison of pretest and posttest mean scores on this measure indicate that students participating in either condition of the intervention decreased their level of unawareness of blatant racial issues, while the level of unawareness regarding blatant racial issues actually increased slightly over the same time period for students in the comparison group who did not participate in any form of the intervention.

Overall, the results from the quantitative portion of this study provide only limited support for the efficacy of this particular intervention, or the Contact Hypothesis as a whole. However, there are several factors that need to be more closely considered and further studied before final conclusions can be drawn.

First, were the quantitative measures effective in measuring the types of changes potentially produced as a result of the intervention? While efforts were made to use quantitative measures that captured various different aspects of issues of awareness and/or acceptance of similarities and differences, the measures were all self-report and somewhat transparent or face valid. Considering that the participants all self-selected, and therefore likely considered themselves somewhat aware and egalitarian, they may have actually found through participation in the intervention that they had more biases or less awareness than they had previously believed. Ironically, this increased awareness of ignorance may have lead to a more negative, but also more accurate assessment of
awareness or egalitarian beliefs at posttest than at pretest. Thus, actual positive changes may have been muted. This is a potential concern that needs to be addressed in future research, perhaps with additional measures that are not self-report. However, one way to examine the measures that were used in this study is to look at how scores and scale or factor scores related to each other. This can be accomplished by looking at correlation matrices. Three separate correlation matrix were utilized, one comparing pretest scores, one posttest scores, and one change scores (Tables 6, 7, and 8 respectively), correlations based on age, gender, and race were also examined. It should be noted that higher scores on the CoBRAS indicate greater lack of awareness, thus lower scores are preferable or better. As a result, scores on the CoBRAS factor scores are likely to be negatively correlated with measures on the M-GUDS or QDI were larger scores indicate greater acceptance or more egalitarian attitudes. Age did not correlate highly with any measures at either pretest or posttest. However, age did show a significant negative correlation at the .05 level for change on the CoBRAS factor II. It would appear that as age increased, change scores decreased. Since decreases in scores on the CoBRAS are good, this would indicate that increase in participant age correlates with more positive change on this measure. This is somewhat of a surprise since there was concern about potential ceiling effects. However, older participants may have been better able to understand the more subtle components addressed by this measure (i.e. – awareness of institutional discrimination). Gender significantly correlated at both pretest and posttest with scores on the CoBRAS factor II and the QDI factor III, with women scoring better than men on both. This is not a surprise on the QDI factor III since it explicitly measures attitudes related to gender equity, but the difference on the CoBRAS factor II warrants further
investigation. At pretest, race significantly correlated with scores on the CoBRAS factor I and QDI factor III, with nonwhites scoring better on the CoBRAS factor II and whites scoring better on the QDI factor III. However, the correlation with race and the CoBRAS factor II was not present at posttest, but the correlation with the QDI factor III remained. A number measures and factor scores correlated significantly at both the .05 and the .01 levels with other measures. This was consistent across both pretest and posttest measurements. This may indicate that the instruments are not differentially measuring attitudes or beliefs for the population in this study. Again, how to accurately measure the attitudes and beliefs that intergroup contact hopes to change is a challenge that needs to be addressed further in future research.
Table 5

Bivariate Correlations for age, gender, race, and all scales & subscales at pretest

<table>
<thead>
<tr>
<th></th>
<th>Age</th>
<th>Gender</th>
<th>Race</th>
<th>CoBRAS factor I</th>
<th>CoBRAS factor II</th>
<th>CoBRAS factor III</th>
<th>M-GUDS</th>
<th>QDI factor I</th>
<th>QDI factor II</th>
<th>QDI factor III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>-.167</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race</td>
<td>-.223</td>
<td>-.042</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CoBRAS factor I</td>
<td>.012</td>
<td>-.004</td>
<td>.036</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CoBRAS factor II</td>
<td>.195</td>
<td>-.316**</td>
<td>-.237*</td>
<td>.463**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CoBRAS factor III</td>
<td>-.062</td>
<td>-.065</td>
<td>.127</td>
<td>.595**</td>
<td>.401**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M-GUDS</td>
<td>.103</td>
<td>.089</td>
<td>-.009</td>
<td>-.273*</td>
<td>-.278*</td>
<td>-.417**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>QDI factor I</td>
<td>-.079</td>
<td>.208</td>
<td>.007</td>
<td>-.655**</td>
<td>-.670**</td>
<td>-.685**</td>
<td>.442**</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>QDI factor II</td>
<td>.141</td>
<td>-.027</td>
<td>.057</td>
<td>.038</td>
<td>-.219</td>
<td>-.311**</td>
<td>.540**</td>
<td>.231*</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>QDI factor III</td>
<td>-.044</td>
<td>.432**</td>
<td>-.291*</td>
<td>-.227</td>
<td>-.330**</td>
<td>-.451**</td>
<td>.331**</td>
<td>.430**</td>
<td>.213</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 6

Bivariate Correlations for age, gender, race, and all scales & subscales at posttest

<table>
<thead>
<tr>
<th></th>
<th>Age</th>
<th>Gender</th>
<th>Race</th>
<th>CoBRAS factor I</th>
<th>CoBRAS factor II</th>
<th>CoBRAS factor III</th>
<th>M-GUDS</th>
<th>QDI factor I</th>
<th>QDI factor II</th>
<th>QDI factor III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>-.167</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race</td>
<td>-.223</td>
<td>-.042</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CoBRAS factor I</td>
<td>.018</td>
<td>-.088</td>
<td>-.011</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CoBRAS factor II</td>
<td>.007</td>
<td>-.257*</td>
<td>-.014</td>
<td>.555**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CoBRAS factor III</td>
<td>.035</td>
<td>.211</td>
<td>.070</td>
<td>.588**</td>
<td>.478**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M-GUDS</td>
<td>.035</td>
<td>.184</td>
<td>.027</td>
<td>-.281*</td>
<td>-.345*</td>
<td>-.565**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>QDI factor I</td>
<td>.001</td>
<td>.158</td>
<td>.000</td>
<td>-.483**</td>
<td>-.674**</td>
<td>-.632**</td>
<td>.562**</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>QDI factor II</td>
<td>.173</td>
<td>.081</td>
<td>.024</td>
<td>-.234*</td>
<td>-.211</td>
<td>-.449**</td>
<td>.729**</td>
<td>.455**</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>QDI factor III</td>
<td>-.079</td>
<td>.421**</td>
<td>-.234*</td>
<td>-.240*</td>
<td>-.403**</td>
<td>-.502**</td>
<td>.454*</td>
<td>.557**</td>
<td>.263</td>
<td>1</td>
</tr>
</tbody>
</table>

Note. For both of the above: Gender (1=male, 2=female), Race (1=white, 2=nonwhite) / **=p<.01 and *=p<.05.
### Table 7

Bivariate Correlations for change in scores from pretest to posttest by age, gender, race, and all scales & subscales

<table>
<thead>
<tr>
<th>Age</th>
<th>Gender</th>
<th>Race</th>
<th>CoBRAS factor I</th>
<th>CoBRAS factor II</th>
<th>CoBRAS factor III</th>
<th>M-GUDS</th>
<th>QDI factor I</th>
<th>QDI factor II</th>
<th>QDI factor III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>-.167</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race</td>
<td>-.223</td>
<td>-.042</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CoBRAS factor I</td>
<td>.004</td>
<td>-.108</td>
<td>-.071</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CoBRAS factor II</td>
<td>-.267*</td>
<td>.070</td>
<td>.309**</td>
<td>.077</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CoBRAS factor III</td>
<td>.120</td>
<td>-.197</td>
<td>-.061</td>
<td>.374**</td>
<td>.041</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M-GUDS</td>
<td>-.102</td>
<td>.197</td>
<td>-.066</td>
<td>-.140</td>
<td>-.036</td>
<td>-.189</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>QDI factor I</td>
<td>.113</td>
<td>-.072</td>
<td>-.010</td>
<td>-.105</td>
<td>-.119</td>
<td>-.317**</td>
<td>.243*</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>QDI factor II</td>
<td>.060</td>
<td>.181</td>
<td>-.054</td>
<td>.175</td>
<td>.010</td>
<td>-.069</td>
<td>.040</td>
<td>-.120</td>
<td>1</td>
</tr>
<tr>
<td>QDI factor III</td>
<td>-.061</td>
<td>.039</td>
<td>.052</td>
<td>.093</td>
<td>-.129</td>
<td>-.083</td>
<td>.360**</td>
<td>.279</td>
<td>.210</td>
</tr>
</tbody>
</table>

*Note. Gender (1=male, 2=female), Race (1=white, 2=nonwhite) / **=p<.01 and *=p<.05.*
Second, did this particular intervention provide all of the necessary components for optimal intergroup contact? Pettigrew and Tropp (2000) defined the following as necessary prerequisites for optimal contact conditions: 1) equal status 2) common goals 3) no competition 4) authority sanction. They further asserted that special attention should be paid to make sure that contact conditions are structured in a manner where both majority and minority participants perceive them as optimal. Every effort was made to provide conditions that would be perceived as optimal by all participants (see Chapter Three: method section – the intervention), and participants as a whole rated the intervention experience very highly on post-intervention evaluations (see appendix I). However, it is not clear that the overall intervention experience was long enough to provide the necessary time to overcome the initial discomfort and anxiety that are typical at the beginning of any intergroup contact (Islam & Hewstone; Pettigrew, 1998; Stephan, 1992). The findings may also be an indication that awareness of blatant racial issues is more easily obtained through intergroup contact than awareness or positive attitude change on the other dimensions assessed.

Third, was there anything unique about this particular study that may have influenced the results? There are two aspects of this particular study that definitely warrant further consideration. One is the sample, and the other is the unique historical time and place in which this study was conducted.

The sample in this study was completely voluntary and relatively small. It may not be representative of the larger population of University students. In fact, the combined students in the intervention had “better” scores (i.e. – lower on the M-GUDS and QDI, and higher on the CoBRAS) at pretest than those in the comparison group on
every measure. These differences between participants in the control condition and the intervention conditions were found to be significant at the 95% confidence interval for all pretest measures except for the CoBRAS factor III (unawareness of blatant racial issues) and QDI factor III (general attitudes regarding women’s equality), with the mean score for the control group being worse (i.e. – lower on the M-GUDS and QDI, and higher on the CoBRAS) on every significantly different measure. While pretest scores were used in all analyses as a covariate in an attempt to control for potential pretest differences, it seems that this may not have provided sufficient protection. Upon closer examination, it becomes apparent that the primary difference between students in the control and students in the intervention conditions is primarily among the White students. (See Table 8)
Table 8
Pretest Means and Standard Deviations by Condition & Race

<table>
<thead>
<tr>
<th>Comparison group (n = 25; 15 White, 10 of color)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(M-GUDS) - Attitudes regarding similarities and differences</td>
</tr>
<tr>
<td>(CoBRAS factor I) - Lack of awareness regarding racial privilege</td>
</tr>
<tr>
<td>(CoBRAS factor II) - Lack of awareness regarding institutional discrimination</td>
</tr>
<tr>
<td>(CoBRAS factor III) - Lack of awareness regarding blatant racial issues</td>
</tr>
<tr>
<td>(QDI factor I) - General attitudes regarding racial diversity</td>
</tr>
<tr>
<td>(QDI factor II) - Specific attitudes regarding general contact and personal contact with racial diversity</td>
</tr>
<tr>
<td>(QDI factor III) - General attitudes regarding women’s equality</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Experimental Condition 1: factual readings (n = 24; 7 White, 17 of color)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(M-GUDS) - Attitudes regarding similarities and differences</td>
</tr>
<tr>
<td>(CoBRAS factor I) - Lack of awareness regarding racial privilege</td>
</tr>
<tr>
<td>(CoBRAS factor II) - Lack of awareness regarding institutional discrimination</td>
</tr>
<tr>
<td>(CoBRAS factor III) - Lack of awareness regarding blatant racial issues</td>
</tr>
<tr>
<td>(QDI factor I) - General attitudes regarding racial diversity</td>
</tr>
<tr>
<td>(QDI factor II) - Specific attitudes regarding general contact and personal contact with racial diversity</td>
</tr>
<tr>
<td>(QDI factor III) - General attitudes regarding women’s equality</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Experimental Condition 2: personal readings (n = 23; 7 White, 16 of color)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(M-GUDS) - Attitudes regarding similarities and differences</td>
</tr>
<tr>
<td>(CoBRAS factor I) - Lack of awareness regarding racial privilege</td>
</tr>
<tr>
<td>(CoBRAS factor II) - Lack of awareness regarding institutional discrimination</td>
</tr>
<tr>
<td>(CoBRAS factor III) - Lack of awareness regarding blatant racial issues</td>
</tr>
<tr>
<td>(QDI factor I) - General attitudes regarding racial diversity</td>
</tr>
<tr>
<td>(QDI factor II) - Specific attitudes regarding general contact and personal contact with racial diversity</td>
</tr>
<tr>
<td>(QDI factor III) - General attitudes regarding women’s equality</td>
</tr>
</tbody>
</table>
The White students who chose to participate in the intervention were significantly more racially aware and egalitarian regarding racial differences, than the White students in the control condition. No such differences were found for the students of color in the control compared to the students of color in the intervention. The mean score for students of color in the control was better than the mean score for their White counterparts in the control on every measure. However, the difference was only statistically significant at the .05 level for the M-GUDS and the CoBRAS factor II. Interestingly, the White students in the intervention conditions scored significantly better than the students of color, either in the control or intervention conditions. This is of particular interest, since Whites overall do not generally demonstrate as great a level of racial awareness or egalitarian attitudes regarding race as people of color (Pettigrew, 1998). This raises concerns related to potential ceiling effects for the White students in the intervention conditions. Because White students in the intervention conditions appear much more racially aware and egalitarian related to racial issues at pretest then their control group counterparts, it is likely that they have already had greater experience with intergroup contact, whether as a cause or effect of their greater awareness and egalitarian attitudes. Thus, the additional intergroup contact provided by the intervention, as with any subsequent intergroup contact, will likely produce diminishing effect. Further, previous authors have suggested that a type of ceiling effect might be somewhat built in for people of color. The idea is that people of color in U.S. society are much more exposed to intergroup contact and issues of race by simple virtue of being a minority. Thus, the effect of any new or additional experience is likely diminished (Pettigrew, 1998).
A somewhat unique feature of this study is the fact that both majority and minority group members were included in the same study, and examined on the same measures. Historically, most of the research in this area has studied only the effects on the majority participants (Devine and Vasquez, 1998; Pettigrew and Tropp, 2000), or the effects of majority and minority group members in isolation (Devine and Vasquez, 1998). Further, while it has been noted that minority group members may also hold prejudicial views of members of other groups, be it majority group members or members of other minority groups (Smith, 1994; Works, 1961), few if any studies have focused on the effects of intergroup contact on measures of prejudice or negative attitudes held by minority group members. For minority group members, the focus has typically been on self-esteem and adjustment consequences of being a member of a stigmatized group (Crocker and Major, 1989; Jones et al., 1984; Seligman and Welch, 1991). Thus, it is not clear how minority group members typically respond to intergroup contact in regard to their own levels of prejudice or awareness. Another complicating factor is the potential effect of being confronted about one’s own potential prejudices and biases in the company of members of other groups. Finlay and Stephen (2000) suggested that the role of empathy is very important in facilitating change related to attitudes about outgroups. They further postulated that dissonance may play a major role in the ability to gain empathy, and that one’s own ingroup being implicated as contributing to an outgroup’s suffering may significantly complicate the manner in which empathy develops. This is likely confounded even more if that implication of responsibility occurs in the presence of members of other groups (e.g. – I have found that any dialogue on race is drastically influenced/altered by the inclusion of members from another racial group: i.e. –
dialogues among all African American participants is qualitatively different from
dialogues among only people of color, but including not just African Americans but also
Hispanics/Latinos or Asians, which is qualitatively different from dialogues which
include various peoples of color and also Whites). All of these factors make any
conclusions from this study very complex and somewhat tentative and raises further
research issues discussed in Chapter Six.

The second issue that needs to be considered regarding this particular study is the unique time and place in which it was conducted. This study was conducted in the spring of 2002. The previous spring there had been major racial unrest and protests on campus. Residuals from these events on campus were clearly still being felt, and seemed to significantly change the racial discourse on campus. On an even larger scale, the 9/11 terrorist attacks on the U.S. occurred only a few months before the start of this study. This drastically changed the racial and cultural discourse nationwide, and in response Muslim and/or Middle Eastern students were specifically invited to participate in the Dialogues on Race groups. There is little doubt that both of these events significantly affected the DOR discourse in very complex ways that make any results hard to generalize to other times or populations.

A final important consideration is to assess the results of these quantitative findings in the context of all available data before drawing final conclusions. As noted, there are numerous factors that may limit the conclusions that can be drawn from the quantitative portion of this study. Therefore it is important to explore all data to determine what aspects of this intervention may have had some effect, either positive or negative, and thus show direction for the development of future interventions and
research in this area. Additional data in the form of qualitative responses to the course evaluation questionnaire, weekly session evaluations, and critical incident reports were all examined in an attempt to better understand the experiences of the participants, and will be discussed in the next chapter. Efforts to identify and discuss potential patterns, understanding of outliers, and content analysis of qualitative data will be discussed in a manner that attempts to integrate that data with the quantitative results in Chapter Six, in an attempt to provide a deeper and more complete understanding of the experience and effects of this intervention.

Research Question Two

Did the type of readings to which participants were exposed as part of the nine week Dialogues on Race intervention make a difference in the effectiveness of the intervention? More specifically, do differences exist between students participating in the intervention based on exposure to factual versus personal readings as measured by change in scores from pretest to posttest on the M-GUDS, CoBRAS, or QDI?

To answer question two, three separate analyses were performed, one for each instrument used in the study. Since the M-GUDS produces a single score, an ANCOVA was performed with the posttest score as the dependent variable and the pretest score as a covariate. Because the CoBRAS and QDI both produce 3-factor scores, MANCOVAs were performed for these two instruments, again with the posttest scores as dependent variables and pretest scores as covariates.
Results

The results from the ANCOVA (M-GUDS) and MANCOVAs (CoBRAS and QDI) all supported the null hypotheses that no significant differences would be found at the .05 confidence interval, with (F = .783, p = .381), (F= .523, p = .669), (F = 1.269, p = .298) respectively.

Discussion

While these findings fail to support the suggestion that empathy is important in promoting change regarding deeply held attitudes or biases, there are two important limitations to consider that potentially nullify any conclusions that can be drawn regarding the role of empathy on intervention outcomes as measured here. First, there was a failure to empirically validate whether or not the personal readings were successful in evoking greater levels of empathy, compared to the factual readings, as intended. If the two types of readings did not produce different levels of empathy in participants, then the portion of this research designed to study the role of empathy clearly could not do so as intended. Second, even if we assume the readings effectively produced different levels of empathy as intended, the assigned readings represent a relatively small portion of the overall intervention and group facilitators reported that they produced very little direct effect on the group dialogues. Therefore, any differences generated by the difference in readings may have been lost in the overall intervention. Thus it may be helpful to examine the effects of the readings separate from the overall outcome measures. One way to do this is to examine how the participants assessed the readings as part of their evaluation of the course. Looking at participant responses on the course evaluations (appendix H), it would appear that the students did value or experience the different types
of readings differently. In fact, while there seems to be no notable difference in participants’ responses on the course evaluation to general questions (i.e. – questions 1-4) about their overall assessment of the intervention experience regardless of condition (i.e. – factual or personal readings), there does appear to be a difference on responses to questions focused more specifically on the assigned readings (i.e. – questions 5-7). Participants seem to more positively assess the effects of the personal readings compared to the factual readings. Results from ANOVA on each individual question finds no significant difference between participants assigned the factual readings versus those assigned the personal readings at the .05 confidence level on any of the more general questions (1-4). However, ANOVA performed on the questions specifically focused on the readings (5-7) did find significant differences for responses on questions 5 and 6 with (F = 8.524, p = .005 and F = 4.312, p = .043) respectively. The difference between the two groups was not found to be statistically significant for question 7 at the .05 level (F = 1.450, p = .235) (see table 9). Qualitative data from the two conditions will be examined in the next chapter to see if it can provide additional insight into potential effects of the readings that were not picked up by the pretest posttest assessments.
Table 9

Post-test questionnaire by Intervention Condition

<table>
<thead>
<tr>
<th>Question</th>
<th>Condition 1 (factual)</th>
<th>Condition 2 (personal)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>1. Did your dialogue group meet your expectations?</td>
<td>2.00</td>
<td>.913</td>
</tr>
<tr>
<td>2. Was your group facilitated in a manner that encouraged open and honest dialogue?</td>
<td>1.24</td>
<td>.597</td>
</tr>
<tr>
<td>3. Would you recommend participation in a similar group to a friend?</td>
<td>1.12</td>
<td>.440</td>
</tr>
<tr>
<td>4. Did your participation in this group help you to explore/challenge your views and/or biases regarding race and race relations?</td>
<td>1.88</td>
<td>.781</td>
</tr>
<tr>
<td>5. Did the assigned readings help you gain greater understanding of race and race relations?</td>
<td>*2.60</td>
<td>1.258</td>
</tr>
<tr>
<td>6. Did the assigned readings help you gain greater empathy for the experience of others?</td>
<td>*2.16</td>
<td>1.068</td>
</tr>
<tr>
<td>7. How do you feel the assigned readings facilitated your group’s ability to engage in open and honest dialogue?</td>
<td>2.80</td>
<td>1.225</td>
</tr>
</tbody>
</table>

Note - All responses are based on a 5-point Likert type scale with more positive responses anchored at 1, and less positive responses anchored at 5. *- indicates a statistically significant difference between groups at the .05 level. ** - 2 respondents left question 1 blank. So n=21 for condition 2 on question 1.
Summary of Quantitative Research Results

This chapter attempted to address the first two primary research questions.

1. Were there significant changes in attitudes or level of awareness related to racial diversity or gender equity for students participating in a nine-week intervention program (Dialogues on Race) compared to a control group of students not participating in the intervention? 2. Did the type of readings (factual or personal) to which participants were exposed as part of the nine-week intervention make a difference in the effectiveness of the intervention?

To answer the above questions, three quantitative measures (M-GUDS, CoBRAS, and QDI) were administered to all students participating in the intervention, as well as to a group of students serving as a comparison group and not participating in the intervention, immediately prior to the start of the nine-week intervention and again immediately following completion of the intervention. The M-GUDS produces one total score while the CoBRAS and QDI each produce three factor scores.

Question One Results. An ANCOVA using pretest scores as a covariate was performed for the M-GUDS, and MANCOVAs also using pretest scores as covariates were performed for both the CoBRAS and QDI. No statistically significant differences were found between those participating in the intervention compared to the control group at the .05 level for: the M-GUDS, which assesses attitudes regarding similarities and differences; the CoBRAS Factor I, which assesses lack of awareness regarding racial privilege; the CoBRAS Factor II, which assesses lack of awareness regarding institutional discrimination; the QDI Factor I, which assesses general attitudes regarding racial diversity; the QDI Factor II, which assesses specific attitudes regarding general contact
and personal contact with racial diversity; or the QDI Factor III, which assesses general attitudes regarding women’s equity.

However, a statistically significant difference at the .05 level between those participating in the intervention and the comparison group was found for the CoBRAS Factor III, which assesses lack of awareness regarding blatant racial issues. Further, a comparison of pretest and posttest mean scores on this measure indicated that students participating in either condition of the intervention (i.e. – exposed to either factual or personal readings) significantly decreased their level of unawareness regarding blatant racial issues over the course of the intervention compared to students in the control group who did not participate in the intervention.

Question Two Results. Again an ANCOVA using pretest scores as a covariate was performed for the M-GUDS, and MANCOVAs also using pretest scores as covariates were performed for both the CoBRAS and QDI. No statistically significant differences at the .05 level were found between those participating in the intervention assigned “factual” readings compared to those participating in the intervention assigned “personal” readings for any of the scores or factor scores on the three quantitative measures (M-GUDS, CoBRAS, and QDI) used in this study. However, it was noted that when course evaluations were examined, participants did seem to rate the “personal” readings more favorably than the “factual” readings.

Summary of Quantitative Research Results Discussion. Overall, the findings for question one gave very limited support for the efficacy of this particular intervention, and for the Contact Hypothesis as a whole, as measured by the quantitative instruments used in this study. Further, the findings provided no support on pretest and posttest
measurements for the idea that exposure to “factual” compared to “personal” readings would differentially affect the experience.

However, numerous factors that may have affected the results were also noted and discussed. These included issues of possible sampling bias, potential ceiling effects, and lack of empirical validation to confirm that the different types of assigned readings evoked different levels of empathy. It was noted that White participants in the intervention condition started with much higher levels of awareness and more positive views on similarities, differences, and racial equity than White participants in the comparison condition, and this might have led to a ceiling effect that artificially limited the amount of change demonstrated by the scores on the quantitative instruments from pretest to posttest for the participants exposed to the intervention compared to those in the control group. Potential ceiling effects for non-White participants were also discussed. Additionally, unique historical factors were also raised as an issue to consider: Specifically, the 9/11 terrorist attacks and racially motivated protests and sit-ins on camps all occurred within the 10 months prior to the start of this study. These major campus and national/international events definitely influenced the context in which the intervention occurred, and they were often topics within the dialogues (e.g. – the subject of racial profiling took on a very different meaning than it had the previous semester). A final note was to consider the quantitative findings in this chapter as only part of the available data from this study. The findings in this chapter should be viewed along with the qualitative data presented in the next chapter before any final conclusions are drawn.
CHAPTER FIVE

Qualitative & Exploratory Research: Results and Discussion

Chapter five addresses the third research question of this study, what aspects of the intervention did participants identify as important, by presenting the findings generated from weekly session evaluation questionnaires (SEQ) and content analyses of the weekly critical incident reports provided by intervention participants at the end of each session. The qualitative research component of this study was broadly designed to explore the phenomenological experiences of the intervention participants. The design included collecting weekly SEQ and critical incident reports from all intervention participants, plus 12 post-intervention, semi-structured, follow-up interviews with two participants from each of the six different groups (i.e. – the participant deemed by the facilitators of that group to have gained the most, and also the participant deemed to have gained the least, from their participation in the intervention). Unfortunately, data from the 12 interviews was determined to be of insufficient quality to allow proper content analyses. Specifically, the audio-tapes used to record the interviews were not of sufficient quality to be properly transcribed. The tapes were taken to several different professionals to be transcribed, and contact was made with several audio-visual laboratories to see if the tape quality could be enhanced. Ultimately, all efforts to salvage the data from the audio taped interviews failed, and the data was determined to be unusable. As a result, only findings generated by the SEQ and from the content analyses of the weekly critical incident data will be presented in this chapter. Themes and patterns will be explored and also compared to themes and patterns from the pretest/posttest
quantitative findings. Specifically, this chapter reports on six components in the following sequence:

1. A description of the qualitative analyses techniques and underlying principles used by the investigator in this study.

2. A presentation and description of the themes or categories that emerged from the weekly critical incident data and illustrative events for each of the categories.

3. A presentation of Critical Incident theme patterns based on when they occurred over the course of the intervention, as well as by condition and race.

4. A presentation of Session Evaluation Questionnaire response patterns based on when they occurred over the course of the intervention, as well as by condition and race.

5. A presentation of patterns which were identified when comparing participants whose pretest and posttest scores on the quantitative measures used in this study (i.e. – CoBRAS, M-GUDS, and QDI) indicated the most versus least positive change over the course of the intervention.

6. An overall summary and discussion of the qualitative results and patterns identified in this chapter.

Research Question Three

For all participants in the intervention, both racial majority and racial minority group members, what aspects of the intervention did they identify as important, when, and why? More specifically, examining weekly critical incident reports, posttest structured interviews, and course evaluations, what do students have to say about their experience? What does this phenomenological data add to our understanding of the
intervention experience, and in what ways does this qualitative data support, challenge, or help explain the quantitative findings?

Based on data collected weekly in the form of the SEQ and critical incident reports obtained through an open-ended request (“Please briefly describe the exchange or interaction that you feel was the most significant for you personally during this session.”) included at the end of the SEQ, the participants’ phenomenological experiences of the intervention were examined. Specifically, what topics, events, inter or intrapersonal processes did the participants identify as particularly significant? When during the course of the intervention did various themes emerge? How did participants rate the experience over time on the SEQ? Did participants respond to the weekly critical incident request in ways that help explain their measured scores on the quantitative research? Do participants’ responses differ when compared across intervention condition (i.e. – exposure to either factual or personal readings) or race (i.e. – White or non-White)?

Qualitative Techniques, Underlying Principles, and Procedures

Merriam (1988) noted that the qualitative research paradigm, particularly case studies, focuses on discovery, insight, and understanding from the perspectives of those being studied, and thus in her opinion offers the greatest promise of making significant contributions to the knowledge base and practice of education. Further, qualitative case studies are typically hypothesis-generating rather than hypothesis-testing.

Thus, the qualitative portion of this study was designed to generate further hypotheses about the workings of one particular intervention program, Dialogues on Race. Hypotheses are included in the discussions following each result section in this
chapter, and selected hypotheses are summarized at the end of this chapter. The fundamental qualitative analytic technique used in this study is a phenomenology. A phenomenological approach works well for hypotheses-generation because it is inductive in nature, meaning that it generates a theory after the data is collected rather than before it is collected. The primary goal of this approach is to study and understand the problem or experience by attempting to enter the field of perception of the participants; seeing how they experience, live, and display the phenomenon; and looking for the meaning of the participants’ experiences (Creswell, 1998). As described in Chapter Three, various qualitative techniques were used to help control for internal and external validity and reliability. Triangulation, a reliance on multiple methods of measurement and multiple investigators, is one technique used to assist in ensuring internal and external validity and reliability (Merriam, 1988). The multiple methods of measurement used included: Three quantitative pretest/posttest instruments (i.e. – the CoBRAS, M-GUDS, and QDI); weekly data collected in the form of the SEQ and also critical incident reports; and post-intervention open-ended follow-up interviews with six participants from each condition. Another strategy used to help ensure internal validity was peer examination.

The process for thematic categorization of events was guided by the coding system described by Merriam (1988), and used by Bieschke et al. (2003). The primary investigator solicited the other members for the peer examination team by asking for volunteers from among the qualified members of the CAPS staff and trainees to help sort critical incident reports into categories. Qualifications for being part of the peer examination team of raters included familiarity with the intervention (Dialogues on Race), and a variety of demographic characteristics. The team included the primary
investigator who is an African American, heterosexual, male counseling psychology
doctoral student who has worked full time at a university counseling center for more than
five years; one Caucasian, heterosexual, female counseling psychologist who at the time
worked primarily in a university counseling center; and one Caucasian, heterosexual,
male counseling psychologist who also at the time worked primarily in a university
counseling center. All of the raters had taken doctoral-level group courses and had co-
facilitated numerous therapy, support, or discussion groups. Additionally, all had
previously served as D.O.R. co-facilitators, but not during the semester this data was
collected, nor with any of the students involved in this study.

Following the completion of the nine-week program, the hand written weekly
responses to the request (“Please briefly describe the exchange or interaction that you feel
was the most significant for you personally during this session.”) at the end of each post-
session questionnaire was entered into the computer spread sheet along with the session
number and unique student identification number. This was done so that once the
responses were classified into one of five pre-determined categories it would be easy to
check the responses based on criteria such as session, condition, race, etc. The critical
incident responses were then transferred verbatim, one each, onto individual index cards
so that they could be easily sorted. The primary investigator then read through all of the
individual critical incident responses three times to get a sense of all of the responses. He
then read through the events again and identified themes. It took several complete
readings before five specific categories were identified. The peer examination team then
met collectively to discuss the overall task and the definition for each category. All raters
then agreed to separately read through all critical incident events one time prior to
sorting, and then to sort the individual events into the five identified categories. Any event that a rater felt did not fit appropriately into one of the predetermined categories was set aside. The team of raters then met again to discuss a sample of the events that were not unanimously categorized. Raters discussed and fine tuned the definitions for all categories and then agreed to individually re-examine those critical incident events they had previously not categorized, or events the raters had not unanimously agreed on the category the event best fit. Raters then returned their final ratings for all events to the primary investigator. Some events remained uncategorized by some raters, and unanimous agreement was not reached on how several others should be classified. Including the same 48 participants that represented the final data set used in the quantitative analyses, there was the potential for 432 individual reports of critical incident events (i.e. – 48 participants x 9 weeks). From the potential 432 responses, 102 or 23.61% were either not returned or left blank, 37 or 8.56% were dropped because they were illegible or too vague, 34 or 7.87% were determined by the peer examination team to either not fit appropriately into any of the five categories or the raters failed to unanimously agree regarding which category was most appropriate. Ultimately, the three person peer examination team unanimously agreed on the categorization of 259 reported critical incidents into 1 of the 5 categories.

**Category Definitions and Exemplars**

The five categories that emerged from the reported critical incidents were labeled: *Self*, *Other*, *Topic*, *Group Process*, and *Structured Exercise*. Definitions and exemplars for each of the categories follow.
**Self:** The major criterion for this category was that the primary focus of the identified critical incident be on the individual reporting the incident. The response reflects that the incident was identified as important because of something the author shared or felt, or some personal insight he or she gained. Examples: “I came out to the group”; “I felt misunderstood”; “I was uncomfortable and realized that I need to be more accepting of people who are more religious than I am”

**Other:** The criterion for this category was that the primary focus of the identified critical incident is on comments, thoughts, feelings, or information shared by other members of the group. The person reporting the incident is not clearly active in the exchange, nor do they note any intrapersonal insight. Examples: “When Mary discussed her bisexuality with the group”; “When Susan shared her feelings about her mother’s silence”; “Listening to people talk personally about their experiences”

**Topic:** The criterion for this category was that the primary focus of the reported critical incident was a particular identified topic. Examples: “talking about affirmative action”; “when we talked about stereotypes”; “it was good to talk about September 11th”

**Group Process:** Criteria for this category were that the identified critical incident focused on feelings about the group, general group process, or the group as a whole, with no identified individual or exchange as the focal point. Examples: “I liked the way we were able to talk about the things that were on our minds instead of having to stick to a certain topic all the time”; “The fact that we were very honest with each other and people showed their true emotions and feelings”; “this session was very emotional”; “this session was good because it was less threatening than last”
Structured Exercise: The criterion for this category was that the identified critical incident focused on a structured exercise or assignment. Examples: “I thought the ice-breaker was successful and insightful”; “I really liked the opening exercise”

The peer examination team also unanimously rated some critical incident reports as miscellaneous or vague: This included any reported critical incident raters felt could not clearly be interpreted as relevant to the experience in a way that could be categorized. Examples: “I was very tired and had a hard time focusing”; “the last 20 minutes”; “everything”; “nothing.”

Results of Reported Critical Incident Analyses

Critical Incidents Over Time. This section examines when during the course of the D.O.R. group intervention various themes seemed to emerge. Since not all participants reported a critical incident every session and not all of those reported were ultimately categorized, the raw numbers in each critical incident category are not directly comparable. In order to compensate for this, the percentages of the critical incidents that were unanimously classified in each category are compared and reported.

It should be noted that the amount of structure imposed on the groups externally diminished over the course of the intervention. That is, the co-facilitators for each group were allowed increasingly greater latitude about how and when during each session it was best to introduce a structured exercise so as not to negatively affect the natural “flow” or process of the group. Thus, for the first three weeks there was a very strict protocol followed with each group starting with the same structured exercise at the very beginning of the group to get the session started. Over the remaining six sessions the co-facilitators discussed during their weekly group supervision how their respective groups
were progressing and what type of structured or unstructured interventions would be most appropriate for each group. Prior to session 7, facilitators for all groups agreed to introduce a structured group exercise focused on group members giving each other feedback regarding their perceptions of each other and the progress of the group up to that point. This was decided because the majority of the group facilitators felt as though the groups were beginning to stray from the topic and also share less deeply. Thus, for sessions 1, 2, 3, and 7 all groups experienced very similar structured exercises. However, for the other five sessions the co-facilitators for each particular group decided on the appropriate level and type of interventions to introduce to their groups much more independently. While this allowed groups the autonomy to progress naturally, it also allowed for greater divergence between how groups were facilitated. All efforts were made to minimize this effect without unnaturally restricting the group or co-facilitators. However, this difference may have influenced how participants experienced given sessions, particularly when participants reported structured exercises as the most critical incident.

For the first session the percentage of critical incident reports categorized as “Other” started notably higher than all of the others, while “Self,” “Topic,” and “Process,” were identical in frequency. The percentage of critical incident reports categorized as “Structured Exercise” started and remained low relative to the other categories, but had small spikes on weeks one, three, and seven. While all of the other categories fluctuated in prevalence over the course of the nine week intervention, “Self” had by far the greatest fluctuation with lows on weeks one, seven, and eight and spikes on weeks two and nine. At first, no discernable patterns or trends were recognized. However, upon further
examination it was noted that there might be a slight inverse relationship between the percentage of critical incidents categorized as “Structured Exercise” and those categorized as “Self.” The percentage of the critical incidents categorized as “Other” fluctuated around the same level over the entire course of the interventions except for a notable spike in week seven and a severe drop in week nine. The percentage of the critical incidents accounted for by “Topic” had a sharp drop between weeks one and two, then peaked at week three. From there is seemed to demonstrate a slight downward trend for the remainder of the intervention. In contrast, “Process” had a notable drop between weeks two and three, and then had a relatively steady trend upward over the remainder of the intervention. (See Table 10 and Figure 1)
Table 10

Critical Incident Reports: by Week

<table>
<thead>
<tr>
<th>Category</th>
<th>Week 1</th>
<th>Week 2</th>
<th>Week 3</th>
<th>Week 4</th>
<th>Week 5</th>
<th>Week 6</th>
<th>Week 7</th>
<th>Week 8</th>
<th>Week 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self</td>
<td>19.44</td>
<td>40.74</td>
<td>25.00</td>
<td>26.92</td>
<td>32.00</td>
<td>26.47</td>
<td>14.81</td>
<td>17.24</td>
<td>37.04</td>
</tr>
<tr>
<td>Other</td>
<td>30.56</td>
<td>29.63</td>
<td>28.57</td>
<td>34.62</td>
<td>28.00</td>
<td>29.41</td>
<td>40.74</td>
<td>34.48</td>
<td>18.52</td>
</tr>
<tr>
<td>Topic</td>
<td>19.44</td>
<td>7.41</td>
<td>25.00</td>
<td>23.08</td>
<td>20.00</td>
<td>23.53</td>
<td>14.81</td>
<td>17.24</td>
<td>11.11</td>
</tr>
<tr>
<td>Process</td>
<td>19.44</td>
<td>18.52</td>
<td>10.71</td>
<td>15.58</td>
<td>20.00</td>
<td>20.59</td>
<td>18.52</td>
<td>31.03</td>
<td>33.33</td>
</tr>
<tr>
<td>Struct.</td>
<td>11.11</td>
<td>3.70</td>
<td>10.71</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>11.11</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Critical Incident Reports

Figure 1
Discussion. Several interesting fluctuations and potential patterns were identified regarding the percentage of reported critical incidents that fit into each category over the course of the nine-week intervention. Hypotheses generated to help understand or explain these fluctuations and patterns will be discussed by critical incident category.

Self - One explanation for the apparent spike in the percentage of critical incidents being reported during the final session of the intervention fitting into the “Self” category might be found in the very nature of a scheduled final session. The focus was on personal reflection of the experience and insight into any effect or change that may have occurred as a result of having participated in the intervention. Discussion was focused inward even during the process of saying goodbye to other members. Facilitators were free to self-disclose more intimately their feelings about the group and modeled using “I statements” to express how having shared this experience with group members had effected them personally. Members seemed to express more personal and potentially vulnerable statements in the final session, perhaps due to the fact that this was their last opportunity, or perhaps because they would not have to face other members again if their disclosures were not well received. The possible inverse relationship between “self” and “structured exercise” was more surprising. It should be noted that “structured exercise” was too small and infrequent to establish a strong relationship, but this might be worth further study. Perhaps participants’ focus on self is particularly susceptible to external interventions.

Other – One potential explanation for the reason that the focus seemed to start out much higher on others, as noted by the higher percentage of the critical incident reports
unanimously placed in the “Other” category, might be that new members to a group typically start out assessing others and seeking similarities or common ground (Yalom, 1995). Had these categories been established prior to the intervention, it would have been predictable that the “Other” category would represent a relatively large percentage of the critical incident responses throughout the entire intervention. After all, the idea of learning from and about others is the foundation on which intergroup contact theory and this intervention were founded. The spike in week seven is likely best explained by the specific group intervention used in all of the groups during that week. The intervention was designed to help facilitate the members sharing thoughts, feelings, and assessments of other members. The hope was that this would help focus and bring the group closer together, and also push the level of discourse “deeper.” The fact that the percentage of critical incident responses focused on “Other” spiked to its highest point during this session, and that several members noted the structured exercise intervention as the most critical event during this session, support the idea that the intervention was successful. Analyses of the SEQ, reported later in this chapter, indicated that participants reported having experienced session seven, the one in which the facilitators intervened, as “deeper.” This may also be interpreted as a positive indication of the efficacy of the intervention. This lends support to the idea that specific facilitator interventions can have a notable effect on how members experience a session, and perhaps how critical it is that efforts to programmatically engage people in intergroup contact be well designed and well facilitated.

**Topic** – One potential explanation for why the percentage of the critical incident reports accounted for by the “Topic” category peaked early and then seemed to decrease
over the remainder of the intervention may be that participants were primed to identify the topic since that is the expectation in most didactic courses. Further, the topics may have initially been somewhat taboo since these are topics U.S. societal norms teach individuals to either avoid all together or discuss in a “politically correct” manner. The percentage of critical incident reports focused on topic may have diminished over time as norms were established which encouraged participants to focus more on the interactions among members, and their own feelings and reactions to each other.

*Group Process* – In contrast to topic, the percentage of critical incident reports focusing on group process may have been low early and steadily increased because it was a relatively new concept for many participants. Once they gained an awareness and language regarding group process, they began to identify it as more and more critical.

*Structured Exercises* – The percentage of participants identifying structured exercises as the most critical event for a given session likely stayed low in part because facilitators were always instructed that any structured exercises were only meant to focus and facilitate the group process and not be the focus. However, the literature indicates (Oskamp, 2000) that it is very important to provide structure to help facilitate what is essentially an asocial process of sharing openly, honestly, and personally in mixed company about topics that are socially and taboo (e.g. – race, religion, SES, sexual orientation, gender roles, etc.). My personal experience with these groups is that a certain level of structure is critical to providing a framework in which to start having open and honest dialogues, particularly in mixed company, around topics that U.S. culture has socialized its members to either avoid completely or approach very cautiously. The idea that structure is important is also supported by the apparent effect that structured
exercises may have had during the sessions when they were utilized. This is
demonstrated by the reported perceived “depth” of the session as demonstrated by
analyses of the SEQ reported later in this chapter. Briefly, the sessions that included
specific structured exercises, tended to be rated as at least slightly “deeper” than those not
providing such structure.

Critical Incidents by Condition. This section compares critical incident
reports across experimental conditions (i.e. - exposed to either “factual” or
“personal” readings as part of the intervention). Overall, participants’ critical
incident reports were very similar over the course of the nine-week intervention
regardless of whether they were exposed to “factual” or “personal” readings as
part of the intervention. The most notable difference in critical incident reports
between the two conditions is the percentage of reports that fall into the “Topic”
and “Process” categories. Participants exposed to the “personal” readings were
more likely to identify critical incidents categorized by the peer rater team as
focused on topic, at a rate of 22.69% compared to 14.29% for those exposed to
the “factual” readings. While those exposed to the “factual” readings reported
critical incident events categorized as focused on group process 25.71% of the
time compared to 15.13% for participants exposed to “personal” readings. (See
Table 11 and Figure 2)
Table 11

Critical Incident Reports: by Condition

<table>
<thead>
<tr>
<th>Condition</th>
<th>% Self</th>
<th>% Other</th>
<th>% Topic</th>
<th>% Process</th>
<th>% Structured</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factual</td>
<td>23.57</td>
<td>32.14</td>
<td>14.29</td>
<td>25.71</td>
<td>4.29</td>
</tr>
<tr>
<td>Personal</td>
<td>29.41</td>
<td>28.57</td>
<td>22.69</td>
<td>15.13</td>
<td>4.20</td>
</tr>
</tbody>
</table>

Figure 2
Discussion. Overall it is not surprising that critical incident reports were very similar regardless of the type of readings, factual or personal, to which the participants were exposed. There were only three assigned readings for each group over the course of the nine-week intervention, and despite opportunities to discuss the readings in group, they seemed to have very little direct influence on the group process (i.e. – group facilitators reported that participants spent very little time directly discussing the assigned readings). The original hypothesis was that the personal readings would better facilitate development and expression of empathy than the factual readings. Anecdotal observational data would suggest that there was little difference in the level of empathy expressed by participants based on the readings. In fact, it was noted by group facilitators and the primary investigator that participants assigned the factual readings often moved from discussing the facts or statistics about a particular topic to discussing individual examples, in support of the facts, that they had personal knowledge of or experience with (e.g. – when they discussed the article about Latinos in the legal system, members shared little about the overall situation of Latinos in the legal system, but instead shared about specific individuals or individual cases of which they were aware). Thus, factual readings seemed to lead to personal dialogues potentially facilitating the development and expression of empathy very similar to that produced by the personal readings. I have no readily available hypothesis to explain the differences in the prevalence rates for critical incident reports ultimately categorized by raters as “topic” or “process.” In fact, this is the opposite of what would have been expected based on the initial hypothesis. These differences may simply be the natural result of the differences between groups, and the fact that there was no overlap between participants in the factual
condition and the personal condition to mute other potential group differences. However, this phenomenon is worth further attention in future research.

*Critical Incidents by Race.* This section compares how frequently critical incident themes or categories were identified as most important based on participants’ self-identified racial category. While race was initially coded into seven categories, it was recoded into two categories (White and non-White) to achieve better balance. Overall, participants’ reported critical incidents coded by the peer examination team into the five identified categories at a very similar percentage over the course of the nine-week intervention regardless of whether participants identified racially as White or non-White. The largest differences in the overall percentage rate for any of the five categories were “Process,” “Other” and “Topic.” Participants identifying as White were more likely to identify critical incidents categorized by the peer rater team as “Other” 35.71% of the time compared to 25.15% for participants identifying as non-White, and “Topic” 22.62% compared to 16.77%. However, those identifying as White reported critical incidents rated as “Process” less often, 10.71% of the time, compared to 26.95% for participants identifying as non-White. (See Table 12 and Figure 3)
### Table 12

<table>
<thead>
<tr>
<th>Race</th>
<th>% Self</th>
<th>% Other</th>
<th>% Topic</th>
<th>% Process</th>
<th>% Structured</th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td>27.38</td>
<td>35.71</td>
<td>22.62</td>
<td>10.71</td>
<td>3.57</td>
</tr>
<tr>
<td>Non-white</td>
<td>26.35</td>
<td>25.15</td>
<td>16.77</td>
<td>26.95</td>
<td>4.79</td>
</tr>
</tbody>
</table>

#### Figure 3

Critical Incident Reports
**Discussion.** The results by race are consistent with expectations based on the literature. Whites in general have often had less exposure to non-Whites than vice versa, and as a result are not forced to think about race as often (Orbe & Harris, 2001). Thus, more of what they learn from initial exposure is new and it is therefore not surprising that they would identify more critical incidents that would be categorized as “Other” or “Topic.” The literature also suggests that Whites and non-Whites often experience the same intergroup interaction or exchange very differently (Dovidio, et. al. 1996). The findings here suggest that Whites might focus more on the topic or content of the exchange and explicitly what the other person is saying, while non-Whites might focus more on the “Process” or other intangibles of the exchange. These differences seem to be present even with White participants who were much more aware than their counterparts in the control group regarding issues of race and equity, and likely also much more aware than the average White student at the university. Herr (1999) noted that culture provides structure for the world by designating what we pay attention to and what we ignore. Thus it is critical that we understand what matters and to whom. The findings here may serve as an additional indication that facilitators of intergroup/intercultural contact need to be aware of many different dynamics in every interaction, and potentially different criteria for judging those interactions used by different racial, ethnic, or cultural groups. For example, considering the historical context of interracial contact in the United States, non-Whites may have developed less trust in the specifics of what is explicitly being said, but may instead be socialized to be more sensitive to the interaction between group members, the process, than Whites. Thus, non-Whites may be more attuned to the implicit aspects of the communication such as how people speak to each other, show
respect, non-verbal cues, etc., in a group context than Whites. This is definitely seems to be an area worthy of further exploration.

Results of Session Evaluation Questionnaire Analyses

SEQ Over Time. This section examines patterns in how the participants evaluated the sessions over the course of the nine-week intervention according to their responses on the SEQ. The SEQ produces four factor scores: depth, smoothness, positivity, and arousal. The depth and smoothness scores measure participants’ reactions to those particular characteristics of the session (i.e. – how “deep” and how “smooth” did the participant find this particular session to be). The positivity and arousal scores measure participants’ post-session mood (i.e. – positivity measures how positive or negative a participant feels immediately after a session and arousal measures the participants’ level of emotional arousal immediately following the session). All four scale scores fluctuated around or near their starting points, except Smoothness which fluctuated more widely. It was noted that all scores were above their average level for the nine-weeks on the final session, and all were also up for the final session compared to the previous session. It should also be noted that the range of the scale scores seemed somewhat restricted. Scale scores on the SEQ have a potential range from 1-7. In this study, the absolute lowest average weekly score on any of the four SEQ scale scores was 4.14 and the highest 6.05. This indicates that participants evaluated each session of the intervention relatively positive with little change from week to week. Because there was not large variation in scores from week to week on any of the four SEQ scale scores, small changes in scores may need to be given greater attention. Since we are using this section to generate hypotheses rather than to test or prove them, we will consider any change in mean score
from one session to the next of .5 or more as warranting particular consideration.

Considering changes from week to week, “Depth” scores only changed .5 or greater between session 8 and 9, increasing by .66 from 5.39 to 6.05. It was also noted that the “Depth” score of 6.05 on week 9 was the only “Depth” score above 6.00. Scores of “Smoothness” fluctuated up and down more than .5 several times: down from week 1 to 2, up from week 3 to 4, down from week 4 to 5, and up again from week 5 to 6, before seeming to stabilize into a slight upward trend. Scores on “Positivity” also fluctuated up and down for the first few weeks; down from week 1 to 2, up from 3 to 4, and down from week 4 to 5, before also seemingly stabilizing into an upward trend. Scores for “Arousal” started and remained relatively low throughout, never changing .5 or greater between any two sessions. (See Table 13 and Figure 4)
Table 13

SESSION EVALUATION QUESTIONNAIRE: Weekly by Mean Scale Score

Week 1, Week 2, Week 3, Week 4, Week 5, Week 6, Week 7, Week 8, Week 9

<table>
<thead>
<tr>
<th></th>
<th>Depth</th>
<th>Smooth</th>
<th>Positiv.</th>
<th>Arousal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>5.67</td>
<td>5.79</td>
<td>5.81</td>
<td>4.57</td>
</tr>
<tr>
<td>Scale</td>
<td>5.71</td>
<td>4.68</td>
<td>5.12</td>
<td>4.31</td>
</tr>
<tr>
<td>Score</td>
<td>5.77</td>
<td>4.40</td>
<td>5.26</td>
<td>4.49</td>
</tr>
<tr>
<td></td>
<td>5.73</td>
<td>5.00</td>
<td>5.79</td>
<td>4.38</td>
</tr>
<tr>
<td></td>
<td>5.50</td>
<td>4.33</td>
<td>5.13</td>
<td>4.14</td>
</tr>
<tr>
<td></td>
<td>5.29</td>
<td>5.12</td>
<td>5.62</td>
<td>4.44</td>
</tr>
<tr>
<td></td>
<td>5.63</td>
<td>4.99</td>
<td>5.40</td>
<td>4.14</td>
</tr>
<tr>
<td></td>
<td>5.39</td>
<td>5.05</td>
<td>5.58</td>
<td>4.40</td>
</tr>
<tr>
<td></td>
<td>6.05</td>
<td>5.12</td>
<td>5.75</td>
<td>4.48</td>
</tr>
</tbody>
</table>

Figure 4
Discussion. Early fluctuations in scale scores for “Smoothness” and “Positivity” may simply represent the volatility involved in the early stages of group development, and the eventual leveling off and upward trends may be the natural outcome of group development towards better cohesion once group norms were established and understood (Yalom, 1995). Another possible explanation might be that the first session was very structured with very little controversy or real challenge of social norms or taboos, while in subsequent sessions participants were challenged to push beyond their normal social politeness and political correctness toward being more honest and “real.” Fluctuations in “Smoothness” scores may indicate the struggles inherent in the process of adjusting to expectations and norms for this type of structured intergroup contact that often go directly against previous socialization regarding intergroup contact in general. Scores on “Positivity” pretty much paralleled scores on “Smoothness” and may indicate that at least for these types of groups, perceptions of how smooth the session was directly influence how positive or negative the participants felt following that session. Scores on “Depth” may have started and remained relatively high over the course of the intervention due to the nature of these groups. In the United States we are often socialized to avoid many of the topics directly and explicitly addressed in this intervention. As a result, participants may feel that simply addressing such topics represents a certain level of session depth. Scores on “Depth” may also be starting to demonstrate an upward trend toward the end of the intervention, but it is too early in any potential pattern to say for sure. The fact that all scores were up some for the final session may be due to a “halo” affect associated with a final session, or from participants’ separation anxiety as they come to terms with this being a meaningful and important experience for them. It was a little surprising that
“Arousal” scores were relatively low and flat throughout the course of the intervention considering the levels of emotional intensity reported by both facilitators and participants over the course of the intervention. This could suggest “guardedness” by participants, but regardless, seems to be an area worth exploring in future research studies.

*SEQ by Condition.* This section compares overall SEQ scale scores by condition (i.e. – factual or personal) averaged over the nine-week course of the intervention. Of the participants included in the final data set, there were 25 assigned to condition I (assigned factual readings) and 23 assigned to condition II (assigned personal readings). The average scores over the course of the intervention for all four scales on the SEQ differ very little when compared by condition. In fact, the scores differ by less than .25 for all four scales (see Table 14 and figure 5). However, if we examine the scores weekly, there are some notable differences, but no clear patterns to indicate anything significant over the course of the intervention, or that could be projected as potentially significant in the future had the intervention continued. Further, it should be noted again that comparing across conditions means that all comparisons are made across group with no overlap. This increases the likelihood that other natural variances between groups may account for minimal differences.
### Table 14

Session Evaluation Questionnaire by Condition and Mean Scale Score

<table>
<thead>
<tr>
<th>Condition</th>
<th>Depth</th>
<th>Smoothness</th>
<th>Positivity</th>
<th>Arousal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factual</td>
<td>5.73</td>
<td>4.92</td>
<td>5.61</td>
<td>4.47</td>
</tr>
<tr>
<td>Personal</td>
<td>5.55</td>
<td>4.97</td>
<td>5.38</td>
<td>4.28</td>
</tr>
</tbody>
</table>

#### Figure 5

Session Evaluation Questionnaire

![Graph showing mean scale scores for factual and personal conditions for Depth, Smoothness, Positivity, and Arousal](image-url)
Discussion. Based on results from the analyses of the SEQ scale scores, there does not seem to be any notable data to indicate that exposure to “factual” versus “personal” readings as part of the overall group intervention had any notable affect on how participants evaluated the experience. Again, this is not an unexpected finding given the relatively small role the assigned readings had in the overall intervention, and the concern related to the readings effectiveness in differentially evoking empathy.

SEQ by Race. This section compares SEQ factor scores by race, defined as White or non-White, averaged over the nine-week course of the intervention. Of the participants included in the final data set, 14 identified as “White,” 33 as “non-White” and 1 did not report race. The average scores over the course of the intervention for three of the four scales on the SEQ differ markedly (i.e. - .5 or more) when compared by race. Overall, White participants on average reported feeling less positive after the intervention sessions and also evaluated the sessions as having gone less smoothly than non-White participants, as indicted by the SEQ scale scores for “Smoothness” and “Positivity.” On average, White participants over the course of the intervention reported experiencing greater levels of anxiety or arousal after sessions versus their non-White counterparts, as measured by the SEQ scale score “Arousal.” On average, the SEQ scale scores for “Depth,” indicating how meaningful or “deep” participants experienced the session to be, indicated virtually no difference between White and non-White participants over the entire intervention. (see Table 15 and figure 6 )
Table 15

<table>
<thead>
<tr>
<th>Race</th>
<th>Depth</th>
<th>Smoothness</th>
<th>Positivity</th>
<th>Arousal</th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td>5.57</td>
<td>4.34</td>
<td>5.01</td>
<td>4.74</td>
</tr>
<tr>
<td>Non-white</td>
<td>5.67</td>
<td>5.17</td>
<td>5.68</td>
<td>4.23</td>
</tr>
</tbody>
</table>

Figure 6
However, when the scale scores are examined by race on a weekly basis, it tells a slightly different story. White participants evaluated the sessions consistently lower every session for “Smoothness” and “Positivity,” and consistently higher every session for “Arousal” than their non-White counterparts. This was consistent with the findings when the averages for the entire intervention were compared. In contrast, while overall scores on “Depth” had been virtually the same for White and non-White participants, this was not the case when compared weekly. Average weekly scores for “Depth” indicated notable differences in both scores and patterns based on race. Compared to non-White participants, scores of session “Depth” started much lower for White participants, rose steadily through the first four sessions, then fluctuated over sessions 6-9 with an extreme spike following the last session. In comparison, scores of session “Depth” started higher, then tended to drop consistently through session 6, and fluctuate over the final 3 sessions with a spike back to the original level following session 9 for non-Whites. (See Table 16 and figures 7 & 8)

Table 16

Session Evaluation Questionnaire: Weekly by Race

| Week 1, Week 2, Week 3, Week 4, Week 5, Week 6, Week 7, Week 8, Week 9 |
|-----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| **Depth**       | White          | 5.02           | 5.11           | 5.60           | 5.89           | 5.88           | 5.26           | 5.63           | 5.22           | 6.25           |
|                 | Non-W          | 5.96           | 5.85           | 5.84           | 5.68           | 5.34           | 5.28           | 5.65           | 5.46           | 5.96           |
| **Smooth**      | White          | 5.38           | 4.20           | 3.96           | 4.35           | 3.98           | 4.31           | 4.39           | 4.56           | 3.96           |
|                 | Non-W          | 5.98           | 4.92           | 4.58           | 5.24           | 4.48           | 5.37           | 5.17           | 5.25           | 5.57           |
| **Positivity**  | White          | 5.27           | 4.65           | 4.77           | 5.36           | 5.09           | 4.60           | 4.77           | 5.29           | 5.33           |
|                 | Non-W          | 6.06           | 5.37           | 5.46           | 5.93           | 5.14           | 5.89           | 5.61           | 5.70           | 5.93           |
| **Arousal**     | White          | 4.63           | 4.63           | 4.88           | 5.21           | 4.50           | 4.73           | 4.43           | 4.50           | 5.14           |
|                 | Non-W          | 4.54           | 4.14           | 4.34           | 4.06           | 3.98           | 4.39           | 4.05           | 4.35           | 4.21           |
Session Evaluation Questionnaire: (weekly) White participants

Week 1 2 3 4 5 6 7 8 9
Mean Scale Score
Arousal Positivity Smoothness Depth

Session Evaluation Questionnaire: non-white participants

Week 1 2 3 4 5 6 7 8 9
Mean Scale Score
Depth Smoothness Positivity Arousal

Figure 7

Figure 8
Discussion. In regard to how participants experienced the intervention sessions, participant responses on the SEQ seem to indicate that race does indeed matter. The results in this section seem to indicate that intergroup contact experiences, at least this one, invoke greater arousal levels for White participants than non-White participants. This is consistent with the literature that suggests that, in general, Whites tend to have fewer experiences with interracial contact than non-Whites, and therefore such contact tends invoke greater arousal levels for them (Pettigrew, 1995). The results here seem to indicate that this was potentially true for White participants in this study, even though they had self selected to participate in this particular intervention and had scored significantly better on quantitative pretest measures (i.e. – The CoBRAS, M-GUDS, and QDI) assessing multicultural awareness and attitudes about race and equality than their White counterparts in the control group who did not self-select to participate in the intervention. Further, they even scored better at pretest on these same measures than the non-White participants in the study who had also self-selected to participate in the intervention. Therefore, it seems logical that the average White college student is likely less aware and egalitarian regarding racial differences than those participating in this intervention, and would likely experience even greater levels of arousal under the same conditions. Another possibility is that along with the novelty of the intergroup contact experience for White students, they may also experience anxiety around fear of being seen as racist, or they may experience what is known as “White guilt” associated with having racial privilege. These same factors might also help explain why White participants consistently reported lower perceived session “Smoothness” or “Positivity” following the sessions. These findings suggest that different conditions may need to be
facilitated for different racial/ethnic groups to have a positive/successful intergroup contact experience. All of this begs for more research to be done that includes both sides of the intergroup contact experience, something that the literature noted as severely lacking.

**Qualitative and Quantitative Findings Together**

This section attempts to examine the qualitative findings in the context of the quantitative findings. Specifically, if we examine the qualitative data (i.e. - the critical incident reports) based on changes in scores from the pretest to posttest on the quantitative measures (i.e. – M-GUDS, CoBRAS, and QDI), do the findings from the qualitative data in any way support, challenge, or bring additional understanding to the quantitative findings? To answer this question, participants’ critical incident reports were grouped based on the amount of change in score from pretest to posttest on each of the quantitative measures. Two groups were formed representing the most extreme ends of the change spectrum for each of the seven quantitative scores generated (i.e. – CoBRAS factors I, II, and III, M-GUDS, and QDI factors I, II, and III). One group consisted of the participants whose scores on the quantitative measure demonstrated the greatest amount of desirable change (Best), while the other group was made up of the participants whose scores demonstrated the least amount of desirable change (Worst). More specifically, the group defined as “Best” consisted of the participants with the greatest change from pretest to posttest in the desired direction (i.e. – decreased scores on the CoBRAS indicating a decrease in lack of awareness regarding racial issues, increased scores on the M-GUDS indicating increased awareness and acceptance of both similarities and differences, and increased scores on the QDI indicating more positive attitudes regarding
racial diversity and gender equity), while the group defined as “Worst” consisted of the participants with the greatest change opposite the desired direction or least change in the desired direction. At least five participants were included in each group, but when there were identical or duplicate change scores such that there was not a clear break point separating the five scores defined either “Best” or “Worst” from the other scores, then all scores up to the next break point were included (e.g. – If the top seven change scores representing the “Best” changes were 8, 5, 5, 4, 3, 2, 2, then only the first five were included. But if the top seven change scores representing the “Best” changes were 8, 5, 5, 4, 3, 3, 2, then the first six were included). The critical incident category percentages were calculated based on the total number of critical incidents categorized into one of the five categories by the peer rater team for that subset of cases (i.e. – if the group of participants separated as “Best” had 25 total critical incident reports categorized by the peer rater team, then the percentage reported for each critical incident category is out of a possible 25). This again adjusted for blank or vague critical incident reports which were dropped, and allowed uneven numbers of cases to be easily compared. Once this process was completed, the percentage of critical incident responses categorized by the peer rater team into each of the five critical incident categories (i.e. – Self, Other, Topic, Group Process, and Structured Exercise) was compared “Best” change cases versus “Worst” change cases for each quantitative measure on the M-GUDS, CoBRAS, and QDI.

Results

CoBRAS. The Structured category represented a very small percentage of the reported critical incidents, making it particularly susceptible to skewing by only one or two responses. Therefore, it was not considered in the findings in this section. For the
CoBRAS factor I (unawareness of racial privilege), the only major differences were for the Topic and Process categories. Topic was identified as the most critical incident at a much greater percentage in the “Best” change cases, and Process in the “Worst” change cases. For the CoBRAS factor II (unawareness of institutional discrimination), Topic was identified as the most critical incident a much larger percentage of the time in the “Best” cases compared to the “Worst” cases, and Self and Other were both identified a slightly lower percentage of the time in the “Best” cases compared to the “Worst” cases. For the CoBRAS factor III (unawareness of blatant racial issues), the greatest difference between the “Best” and “Worst” cases was that Topic was identified as most critical a much greater percentage of the time for “Best” cases. Self, Other, and Process were all identified a markedly lower percentage of the time in “Best” cases versus the “Worst” cases. (See Tables 17-19 and Figures 9-11)
Table 17

<table>
<thead>
<tr>
<th>Change</th>
<th>% Self</th>
<th>% Other</th>
<th>% Topic</th>
<th>% Process</th>
<th>% Structured</th>
</tr>
</thead>
<tbody>
<tr>
<td>Best</td>
<td>32.14</td>
<td>28.57</td>
<td>35.71</td>
<td>3.58</td>
<td>0</td>
</tr>
<tr>
<td>Worst</td>
<td>34.48</td>
<td>20.69</td>
<td>3.45</td>
<td>31.03</td>
<td>10.34</td>
</tr>
</tbody>
</table>

Figure 9
### Table 18

Critical Incident Reports: by Change in CoBRAS Factor 2

<table>
<thead>
<tr>
<th>Change</th>
<th>% Self</th>
<th>% Other</th>
<th>% Topic</th>
<th>% Process</th>
<th>% Structured</th>
</tr>
</thead>
<tbody>
<tr>
<td>Best</td>
<td>23.81</td>
<td>21.43</td>
<td>26.19</td>
<td>23.81</td>
<td>4.76</td>
</tr>
<tr>
<td>Worst</td>
<td>33.33</td>
<td>31.11</td>
<td>11.11</td>
<td>22.22</td>
<td>2.22</td>
</tr>
</tbody>
</table>

![Critical Incident Reports: by Change in CoBRAS Factor 2](image)

**Figure 10**
Table 19

Critical Incident Reports: by Change in CoBRAS Factor 3

<table>
<thead>
<tr>
<th>Change</th>
<th>% Self</th>
<th>% Other</th>
<th>% Topic</th>
<th>% Process</th>
<th>% Structured</th>
</tr>
</thead>
<tbody>
<tr>
<td>Best</td>
<td>15.79</td>
<td>18.42</td>
<td>47.37</td>
<td>13.16</td>
<td>5.26</td>
</tr>
<tr>
<td>Worst</td>
<td>24.14</td>
<td>34.48</td>
<td>10.34</td>
<td>24.14</td>
<td>6.90</td>
</tr>
</tbody>
</table>

Figure 11

Critical Incident Reports: by Change in CoBRAS Factor 3
**M-Guds.** For the M-GUDS (attitudes of awareness and acceptance of both similarities and differences), the greatest differences between the “Best” and “Worst” cases was that Topic was identified as the most critical incident a greater percentage of the time for “Best” cases versus “Worst” cases, while Process and Other were identified as most critical a slightly lower percentage of the time for “Best” cases compared to “Worst” cases. (see Table 20 and Figure 12)
Table 20

<table>
<thead>
<tr>
<th>Change</th>
<th>% Self</th>
<th>% Other</th>
<th>% Topic</th>
<th>% Process</th>
<th>% Structured</th>
</tr>
</thead>
<tbody>
<tr>
<td>Best</td>
<td>29.41</td>
<td>29.41</td>
<td>26.47</td>
<td>14.71</td>
<td>0</td>
</tr>
<tr>
<td>Worst</td>
<td>30.77</td>
<td>34.62</td>
<td>11.54</td>
<td>23.08</td>
<td>0</td>
</tr>
</tbody>
</table>

Critical Incident Reports: by Change in M-GUDS

Figure 12
**QDI.** For the QDI factor I (general attitudes regarding racial diversity), the greatest differences between the “Best” and “Worst” cases was that Topic was identified as the most critical incident a greater percentage of the time for “Best” cases compared to “Worst” cases, while Self and Other were identified as the most critical incidents a slightly lower percentage of the time for “Best” cases versus “Worst” cases. For the QDI factor II (specific attitudes regarding general contact and personal contact with racial diversity), the largest differences were that Self was identified as the most critical incident a much greater percentage of the time in “Best” cases, while Topic and Process were identified as the most critical incident a somewhat higher percentage of the time in the “Worst” cases. For the QDI factor III (general attitudes regarding women’s equality), the greatest differences between “Best” cases and “Worst” cases was that Process was identified a much greater percentage of the time as the most critical incident for the “Best” cases, while Other was identified a larger percentage of the time as the most critical incident for the “Worst” cases. (See Tables 21-23 and Figures 13-15)
### Table 21

**Critical Incident Reports: by Change in QDI Factor 1**

<table>
<thead>
<tr>
<th>Change</th>
<th>% Self</th>
<th>% Other</th>
<th>% Topic</th>
<th>% Process</th>
<th>% Structured</th>
</tr>
</thead>
<tbody>
<tr>
<td>Best</td>
<td>16.13</td>
<td>19.35</td>
<td>41.94</td>
<td>22.58</td>
<td>0</td>
</tr>
<tr>
<td>Worst</td>
<td>27.27</td>
<td>33.33</td>
<td>6.06</td>
<td>27.27</td>
<td>6.06</td>
</tr>
</tbody>
</table>

![Bar chart showing change in QDI Factor 1](image)

*Figure 13*
Table 22

<table>
<thead>
<tr>
<th>Change</th>
<th>% Self</th>
<th>% Other</th>
<th>% Topic</th>
<th>% Process</th>
<th>% Structured</th>
</tr>
</thead>
<tbody>
<tr>
<td>Best</td>
<td>43.33</td>
<td>16.67</td>
<td>13.33</td>
<td>20.00</td>
<td>6.67</td>
</tr>
<tr>
<td>Worst</td>
<td>25.49</td>
<td>17.65</td>
<td>21.57</td>
<td>33.33</td>
<td>1.96</td>
</tr>
</tbody>
</table>

Critical Incident Reports: by Change in QDI Factor 2

Figure 14
Table 23

Critical Incident Reports: by Change in QDI Factor 3

<table>
<thead>
<tr>
<th>Change</th>
<th>% Self</th>
<th>% Other</th>
<th>% Topic</th>
<th>% Process</th>
<th>% Structured</th>
</tr>
</thead>
<tbody>
<tr>
<td>Best</td>
<td>23.08</td>
<td>23.08</td>
<td>7.69</td>
<td>34.62</td>
<td>11.54</td>
</tr>
<tr>
<td>Worst</td>
<td>14.29</td>
<td>37.14</td>
<td>5.71</td>
<td>4.00</td>
<td>2.86</td>
</tr>
</tbody>
</table>

Figure 15
**Discussion**

The most obvious and interesting pattern uncovered by the analyses in this section was that participants’ focus on Topic seemed to be the factor that distinguished between “Best” cases and “Worst” cases on the majority of the quantitative measures (i.e. – CoBRAS factors I, II, and III, M-GUDS, and QDI factor I). A much greater percentage of the critical incident reports identified topic as most important for the “Best” cases compared to the “Worst.” One potential explanation is that these quantitative measures are more sensitive to changes in awareness or knowledge related to particular topics versus changes in levels of awareness or insights about self, others, or the interactions between self and other (process).

Focus on topic did not appear to be the difference between “Best” case and “Worst” case on the QDI factor II or III. What are the potential differences in these two measures compared to the other five? One obvious difference is that the QDI factor III relates to gender equity while all of the other measures focus on race. While gender toles and equity were discussed, it was often included in the conversation related to values espoused by a particular racial or ethnic group or community, and was not the primary topic in and of itself. Another potential difference for both QDI factors II and III is that they focus on more specific attitudes compared to the other measures. While the M-GUDS, CoBRAS factors I, II, and III, and QDI factor I all assess more general awareness, lack of awareness, or attitudes, the QDI factors II and III are more specific (i.e. – specific attitudes and personal contact or attitudes specifically regarding women’s equity). Maybe topic becomes less important and self-awareness or focus on process become more critical as the attitudes become more specific. Maybe the focus needs to be
different depending on what attitudes we are attempting to change? Again, this seems like potentially an important area to research further.

All of these findings seem to indicate a need for more research regarding intergroup contact in general. For instance, continued research regarding what aspects of intergroup contact promote change is important. If knowledge of topic is what promotes the most positive change, then an effective structured intervention may look very different than if empathic relationship is what promotes positive change. Also, more research on what constitutes positive change and how to measure that change seems important. Longitudinal studies would be helpful to understand what types of changes make a difference over the long term, and also what types of differences those changes make. For example, small “vector” changes now may demonstrate very large changes later. Several of the student participants have informed me that participation in this intervention has led to changes in what topics they discuss with family and friends, plus how they interpret those discussions. What ultimately changes a person’s behavior so that they become less discriminatory and more egalitarian remains the crucial question.

**Summary of Qualitative Research Results**

This chapter addressed the third research question of this study: What aspects of the intervention did participants identify as important? When and Why? The qualitative analyses techniques and underlying principles used in this study were reported. The themes or categories that emerged from the weekly critical incident data were described along with illustrative events for each of the categories. Patterns across the course of the intervention regarding when various critical incidents themes emerged were examined, reported, and discussed. This was done for all participants collectively and then also by
condition (exposure to factual or personal readings) and race (White or non-White).
Session Evaluation Questionnaire (SEQ) response patterns were also examined, reported, and discussed. Again, this was done for all participants collectively and the also by condition and race. Finally, comparisons of critical incident reports from participants identified as having made the “Best” or “Worst” changes over the course of the intervention based on their pretest to posttest scores on the quantitative measures used in the study (i.e. – M-GUDS, CoBRAS, QDI) were examined, reported, and discussed.

**Analysis Process.** The fundamental qualitative analytic technique used was a phenomenology. This technique was chosen because it is inductive in nature. The primary investigator examined all weekly critical incident reports provided by intervention participants and by identifying themes, established five distinct critical incident categories: Self, Other, Topic, Group Process, and Structured Exercise. A peer examination team was used to classify each critical incident into one of the five categories identified by the primary investigator.

**Critical Incident Patterns.** Several interesting fluctuations and potential patterns were identified regarding the percentage of reported critical incidents classified as fitting into each category over the course of the nine-week intervention. Hypotheses were generated to potentially help understand or explain fluctuations and patterns.

The percentage of critical incidents reported that focused on “self” fluctuated more than other critical incident categories over the course of the intervention. One hypothesis is that focus on self is particularly susceptible to external interventions.

For the first session of the intervention, reports categorized as “other” were reported much more often than any of the other categories. This category represented a
relatively large percentage of the critical incidents reported over the entire course of the intervention, with a notable exception in the final week. One hypothesis is that this reflects typical group development with group members assessing others and attempting to find common ground during the first stage and learning from others being the primary strength and focus throughout.

Reports of critical incidents categorized as Topic peaked early and then decreased over the course of the intervention. Hypotheses generated to explain the early focus on topic include various forms of socialization. For example, students are often socialized in academic setting to focus on topic. Also, participants may have been socialized to avoid many of the topics discussed during this intervention in intergroup contact situations. Thus the “taboo” nature of the topics may have made them more salient. The decrease in the focus on topic may be due to changes in established norms. Facilitators encouraged participants to focus on interactions more than topics, and the more the topics were discussed the less taboo they became.

The percentage of the incident reports identifying group process as the most critical event started relatively low early in the course of the intervention and seemed to rise steadily. One hypothesis is that this was a relatively new concept for many participants, and once they gained an awareness and language regarding group process they began to identifying it as increasingly important.

The percentage of incident reports identifying structured exercises as the most critical event was relatively low over the entire course of the intervention. However, the literature and anecdotal evidence suggest that providing a certain amount of structure is important to help establish safe boundaries and help facilitate the asocial process of
sharing openly, honestly, and personally in mixed company about topics which are socially taboo.

Critical incident reports were also examined by assigned reading condition. Overall, participant’s critical incident reports were very similar regardless of which type of readings they were assigned (i.e. – personal or factual readings). This was not surprising given the relatively small part of the overall intervention the readings represented, along with the fact that it was not empirically validated that the readings achieved the desired effect of differentially evoked empathy in the participants. The largest difference in reported critical incidents between those assigned the factual reading compared to those assigned the personal readings was that those exposed to the personal readings were more likely to report critical incidents categorized as focused on topic more often while those exposed to factual readings were more likely to report incident identified as focused on group process more often. The differences were somewhat surprising, but not very large, and may simply represent the natural differences between groups.

Critical incident reports were also examined by race (i.e. – White or non-White). Overall incidents identified as critical were very similar regardless of race. The most notable differences in the overall percentage rate for any of the five categories were for “other,” “group process,” and “topic.” Participants identifying as White were more likely to report incidents focused on other or topic as most critical a greater percentage of the time, while participants who identified as non-White were more likely to identify incidents focused on the group process as critical a greater percentage of the time. It was hypothesized that Whites tend to experience less interracial contact than non-Whites and
therefore focus more on the “other” in the contact, as well as potentially on other explicit aspects of the interactions. While in comparison, non-Whites may focus on the “process” and other intangibles. It seems particularly important that more research is done examining the possible differences in ways different identity groups may experience the same interaction.

*SEQ Patterns.* Patterns related to how participants rated each intervention session on the four scales (depth, smoothness positivity, and arousal) of the Session Evaluation Questionnaire were examined, reported, and discussed. All four scale scores fluctuated around or near their starting points, except smoothness which fluctuated more widely. It was also noted that all scores were above their average level for the nine-weeks on the final session, and all were also up for the final session compared to the previous session.

It was hypothesized that early fluctuations in scale scores for “Smoothness” and “Positivity” may simply represent the volatility involved in the early stages of group development, and the eventual leveling off and upward trends may be the natural outcome of group development towards better cohesion once group norms were established and understood (Yalom, 1995). Another hypothesis was that scores on “Depth” may have started and remained relatively high over the course of the intervention due to the nature of these groups. In the United States we are often socialized to avoid many of the topics directly and explicitly addressed in this intervention. As a result, participants may have felt that simply addressing such topics represented a certain level of session depth. The fact that all scores were up some for the final session may reflect participants’ separation anxiety as they came to terms with this being a meaningful and important experience for them. It was a little surprising that
“Arousal” scores were relatively low and flat throughout the course of the intervention considering the levels of emotional intensity reported by both facilitators and participants over the course of the intervention. This could suggest “guardedness” by participants, but regardless, seems to be an area worth exploring in future research studies.

Results from analyses of SEQ data also did not seem to indicate exposure to “factual” versus “personal” readings had any differential affect on how participants evaluated the experience. Again, this was not unexpected given the relatively small role the assigned readings had in the overall intervention.

However, analyses of SEQ responses by race did find some potentially important differences. The results indicated that White participants experienced greater levels of arousal than non-White participants. It was hypothesized that this may be due to White participants having had fewer contact experiences in the past compared to non-Whites, fear associated with being seen a racist, or guilt associated with racial privilege. These same factors may also help explain why White participants reported lower perceived session “smoothness” or “positivity.” Once again, these findings beg for more research including both sides of the intergroup contact experience simultaneously.

Critical Incidents by Quantitative Findings. Participants’ critical incident reports were compared “Best” versus “Worst” defined by the results from pretest to posttest on the quantitative measures used in the study (i.e. – M-GUDS, CoBRAS, and QDI). The most obvious and interesting finding from this comparison was that participants identified as “Best” or having experienced the most change in the desired direction over the course of the intervention based on pretest to posttest scores on the M-GUDS, CoBRAS, and QDI, consistently identified incidents categorized as “topic” as the most
critical more often than those identified as “Worst” based on the same criterion (with the exception of the QDI factors II and III). This finding was somewhat surprising given the emphasis in the literature on self-awareness, understanding of other, and group process. One hypothesis was that the quantitative measures used here are more sensitive to changes in awareness or knowledge of certain topics, versus awareness or knowledge about self, others, or the interaction between self and other. This also seems to warrant further study.
CHAPTER SIX

Summary of the Study, Findings, and Implications

This chapter, summarizes what has occurred throughout each stage of this study as well as presents final conclusions and future recommendations. In particular, Chapter Six summarizes the study’s importance, statement of problem, research questions, methodology, and findings. Also included in this chapter are final conclusions based on the results and their implications for relevant theory, practice, counselor training, and further research

Need for the Study

Demographic trends and advances in technology have increased proximity and interdependence of diverse and historically separate races and cultures worldwide. This situation has some times and some places led to mutually beneficial relationships and at other times to misunderstandings, discrimination, and conflict. As it has become clear that intergroup contact will occur with increasing frequency, the need to understand what leads to positive versus negative outcomes when such contact occurs has also become clear. It has been long understood that prejudice is the core reason for intergroup misunderstanding, discrimination, and conflict. For this reason, prejudice has been a major topic of research in the social sciences to varying degrees since the 1920’s. However, the focus was historically on causes, expressions, and to a lesser degree consequences. Recently there has been a shift towards focusing more on ways to reduce or eliminate prejudice.
Counseling Psychology as a field seems particularly well suited to play a central role in modern efforts to reduce prejudice and discrimination. The field of counseling psychology has taken the lead in attempting to incorporate issues of multiculturalism and diversity into academic curricula and applied training experiences for its students, the importance of counselors gaining multicultural counseling competency has been emphasized at every level. Efforts to gain multicultural competency have taken numerous forms including single courses focused on multiculturalism, infusion of multicultural issues into multiple courses, and less structured experiential interventions including programmed intergroup contact formats. While intergroup contact formats tend to be much more difficult to design or control than either single course or infusion formats, they demonstrate particular promise in efforts to increase multicultural awareness and successfully address anger and other difficult feelings regarding cross-cultural interactions. Much can be learned from the multicultural education efforts of counseling psychology.

In addition to counseling psychologists being trained to recognize and respect cultural differences, training in this field also typically includes a focus on group theory and facilitation. While programmed intergroup contact experiences typically take place in settings outside of therapy, an understanding of the therapeutic factors associated with group interactions is a critical component of any programmed intergroup contact intervention. Facilitators need to be skilled at dealing with group process issues as well as the potentially high levels of internal dissonance and confusion that can be triggered by intergroup contact combined with contemporary forms of prejudice. Again, the values
and training associated with counseling psychology seem particularly well suited to engage in efforts to reduce or eliminate prejudice.

Prejudice has proven very difficult, if not impossible, to truly eliminate. While great strides have been made at reducing certain forms of prejudice in the United States, it seems prejudice has evolved or mutated in ways that help it survive and perhaps even thrive in contemporary U.S. society. In many ways, prejudices in the United States have become increasingly more subtle and thus harder to identify and confront. However, the detrimental consequences to victims, witnesses, and even perpetrators remain. This contemporary form of prejudice can be seen in a prevalent form of racial bias called aversive racism (Davidio and Gaertner, 1998). Aversive racism is defined in contrast to “old-fashioned” racism. Old-fashioned racism is expressed directly and openly, aversive racism is a subtle, often unintentional, form of bias exercised by Americans who possess strong egalitarian values and who believe that they are nonprejudiced. Aversive racists possess negative racial feelings and beliefs of which they are unaware or that they try to dissociate from their non-prejudiced self-images. Aversive racists’ negative feelings do not reflect open hostility or hate. Instead, their reactions involve discomfort, uneasiness, disgust, and sometimes fear. They may find members of other races or ethnicities “aversive,” but at the same time they also find any suggestion that they might be prejudiced aversive. As a result, aversive racists will not discriminate based on race when such discrimination would be obvious. However, in situations where the appropriate response is not as obvious or when an aversive racist can justify or rationalize a negative response on the basis of some factor other than race, then their bias will be expressed (Dovidio and Gaertner, 1998). Findings from a study by Dovidio et. al. (1996)
suggest that aversive racists might be sending very different messages than they intend to in interracial interactions. The difference between intended and received messages can lead to potential conflict, mistrust, and frustration between well-intended individuals or groups attempting to bridge the racial divide. This contemporary form of prejudice is what must now be understood and confronted. The most obvious catalyst for potentially gaining access to these unconscious biases is through intergroup contact.

Social psychology has long held that a major means of reducing intergroup prejudice is through contact between groups. Allport’s version of the contact hypothesis holds that intergroup contact will lead to reduced intergroup conflict if, and only if, the contact meets these four conditions: (1) equal status between groups in the condition, (2) common goals, (3) no competition between groups, and (4) the contact must receive some form of authority sanction (Pettigrew and Tropp, 2000). The Contact Hypothesis has inspired extensive research over the past half century, with mixed results. However, the findings from a meta-analysis of the research by Pettigrew and Tropp (2000) strongly indicate an inverse association between intergroup contact and prejudices. They also identified several areas that call for more research. These areas included research on programmed interventions, impact on minority participants, interactions between majority and minority participants, and the role of empathy.

*Statement of the Problem*

The available current research on intergroup contact indicates that simply bringing groups into contact with each other is not enough. It seems that something more is necessary to help combat the prejudice and misunderstanding that often interferes with positive outcomes when intergroup contact occurs. The purpose of this research study
was to examine four major areas identified in the literature as warranting further attention. Specifically, this study attempted to (1) research the effects of programmed intergroup contact interventions, (2) examine the impact on both majority and minority participants, (3) explore interactions between majority and minority participants, and (4) look at the role of empathy in influencing the outcome of the intergroup contact. This was done by examining whether or not a specific programmed intergroup contact experience, Dialogues on Race that attempted to provide all of the theoretically necessary components, resulted in positive changes in awareness or attitudes related to racial, cultural, or gender diversity. In addition to examining outcome measures, this study also attempted to explore the process of the intervention to determine what aspects of it were important, and to whom. This was done for both majority and minority participants, in addition, efforts to understand interactions between the two groups were included.

Finally, in order to examine the role that empathy may play in effective intergroup contact experiences, assigned readings were experimentally manipulated in an effort to evoke greater levels of empathy from half of the participants.

This was a relatively ambitious undertaking requiring that data be collected in numerous forms. Data were collected pretest and posttest in the form of quantitative measures of awareness, attitudes, and acceptance of similarities and differences. Additional data were also collected on a weekly basis in the form of critical incident reports and session evaluations. Finally, posttest only data were also collected in the form of overall intervention evaluations, and face-to-face interviews. Data were analyzed for outcomes, and also for emerging themes that would help generate hypotheses for future investigation. The following research questions guided the study.
Research Questions

1. Were there significant changes in attitudes or level of awareness related to racial, ethnic, or cultural diversity for students participating in the nine week intervention (Dialogues on Race) compared to a control group of students who did not participate in the intervention? More specifically, do differences exist between students participating in the intervention and students not participating in the intervention, as measured by the Millville-Guzman University-Diversity Scale (M-GUDS), the Color-Blind Racial Attitudes Scale (CoBRAS), or the Quick Discrimination Index (QDI)?

2. Did the type of readings to which participants were exposed as part of the nine week Dialogues on Race intervention make a difference in the effectiveness of the intervention? More specifically, do differences exist between students participating in the intervention based on exposure to factual versus personal readings as measured by change in scores from pretest to posttest on the M-GUDS, CoBRAS, or QDI?

3. For all participants in the intervention, both racial majority and racial minority group members, what aspects of the intervention did they identify as important, when, and why? More specifically, examining weekly critical incident reports, posttest structured interviews, and course evaluations, what do students have to say about their experience? What does this phenomenological data add to our understanding of the intervention experience, and in what ways does this qualitative data support, challenge, or help explain the quantitative findings?
Design

The quantitative portion of this study was a quasi-experimental (due to a lack of randomization) field study employing a pretest-posttest control group design. Pretest-Posttest assessment instruments were identical (QDI, M-GUDS, and CoBRAS). In this design, there were three experimental conditions. First, was the control condition that consisted of participants who were not exposed to any portion or form of the intervention. Participants in the other two conditions were exposed to the same nine-week intervention with one variation; the assigned readings were experimentally manipulated. The members of three of the groups were assigned readings identified as “personal”, and members of the other three groups were assigned readings identified as “factual.” These made up the other two experimental conditions identified as Condition I – “factual”, and Condition II – “personal,” based on which readings participants were assigned.

Because quantitative and qualitative methods involve differing strengths and weaknesses, qualitative data were also collected as part of the study. Qualitative data were collected following each session from each group participant and facilitator in the form of a brief open-ended response regarding the most significant aspect of that session for him/her (i.e. – critical incident report). Further, qualitative data were also collected posttest in the form of a course evaluation sheet completed by all students participating the intervention, and semi-structured individual interviews with two participants from each of the six groups. The qualitative data were utilized to help gain greater understanding of the unique experiences of individual group members, provide depth and description to those unique experiences, and generate hypotheses for further study.
Procedures

Participants. Data were initially collected from 98 undergraduate and graduate students at The Pennsylvania State University during the spring 2002 semester. However, due to participant dropout (i.e. – failure to complete and return all required pre and posttest data), only 73 students are included in the final analyses. Of those 73, 25 served as a control group (Condition 0) and did not participate in any form of the intervention. The other 48 participated in a nine-week intervention program. Of those 48 students, 25 were in groups exposed to factual readings (Condition I), and 23 were in groups exposed to personal readings (Condition II).

Recruitment. All students enrolled in the Dialogues on Race intervention course were encouraged, but not required, to participate in this research study. All students participating in the intervention engaged in the same activities and provided the same information regardless of whether they chose to be included in the study. They were informed that any information they provided would be included only with their explicit written consent. Students not participating in the intervention were also recruited to serve as a comparison group. These students were recruited primarily from two large sociology courses and one African American and Latino leadership program.

Data Collection. Data were collected pretest for all students included in the study, regardless of which condition of the intervention they represented (i.e. – control, intervention with factual readings, or intervention with personal readings), in the form of three quantitative instruments (M-GUDS, CoBRAS, QDI). Data were then also collected on a weekly basis for those participants involved in either condition of the intervention program in the form of (the SEQ, critical incident reports, and Facilitator Rating Sheet).
Data were then collected posttest again for all participants including those in the control condition on the same three quantitative measures used at pretest (M-GUDS, CoBRAS, QDI). Additional posttest data were collected from students exposed to either condition of the intervention in the form of a course evaluation questionnaire, and follow-up semi-structured interviews were conducted with 12 students (i.e. six from each intervention condition in the form of two from each group, the one identified by facilitators as having gained the most from their participation and the one identified as having gained the least).

**Instrumentation**

*Miville-Guzman Universality-Diversity Scale* (M-GUDS; Miville, 1992; Appendix D). This scale was used to measure attitudes of awareness and acceptance of both similarities and differences. This 45-item self-report instrument uses a 6-point Likert-type scale, ranging from 1 (*strongly disagree*) to 6 (*strongly agree*), to measure the respondents’ universal-diverse orientation (UDO). Fifteen items are reverse worded and scored. UDO is a construct that reflects an attitude of awareness and acceptance of both similarities and differences among people. The M-GUDS produces a single score ranging from 45-270, with higher scores reflecting greater awareness and acceptance of similarities and differences among people.

*The Color-Blind Racial Attitudes Scale* (CoBRAS; Neville, Lilly, Lee, Duran, and Brown, 2000; Appendix E). This scale was used to measure color-blind racial attitudes and beliefs. The 20-item CoBRAS is a self-report instrument which uses a 6-point Likert-type scale, ranging from 1 (*strongly disagree*) to 6 (*strongly agree*), to assess cognitive aspects of the respondents’ color-blind racial attitudes. Ten items are reverse worded and scored. The CoBRAS produces scores on three factors: Factor I,
Unawareness of racial privilege (seven items); Factor II, Unawareness of institutional discrimination (seven items); Factor III, Unawareness of blatant racial issues (six items).

The Quick Discrimination Index (QDI; Ponterotto et al., 1995; Appendix F). This index was used to measure attitudes regarding racial equality. The QDI is a 30-item self-report instrument that uses a 5-point Likert-type scale ranging, from 1 (strongly disagree) to 5 (strongly agree), to assess participants’ racial and gender attitudes. Of the items, 15 are reverse worded and scored to control for order effects. The QDI produces three subscales assessing (a) general attitudes regarding racial diversity (9 items), (b) specific attitudes regarding contact and personal contact with racial diversity (7 items), and (c) general attitudes regarding women’s equality (7 items). A total instrument score can be tabulated using all 30 items; however, Ponterotto et al. recommended using a three-factor scoring procedure given that both exploratory and confirmatory factor analyses have supported tripartite conceptualization of the QDI items. The latter procedure was used in this study.

Session Evaluation Questionnaire (SEQ; Stiles and Snow, 1984; Appendix G). To measure participant reactions to, and perceptions of, each session, the SEQ was used. This 24-item self-report instrument uses a 7-point Likert-type scale, ranging from 1 to 7, different anchor points are given for each question (e.g. – good/bad, rough/smooth, etc.), to help gain general insight into the session-by-session process. The SEQ consists of four scales. The Depth (5 items) and Smoothness (4 items) scales measure the participant’s reactions to characteristics of the session. The Positivity (5 items) and Arousal (4 items) scales measure post-session mood. Positivity is a measure of how positive or negative a
participant feels upon completing the session. The Arousal scale is a measure of how much emotional arousal the participant feels after completing the session.

Course evaluation sheet (Appendix H). For exploratory purposes as well as general course evaluation and improvement, the participants completed a one-page course evaluation sheet (Appendix H) requesting participants’ opinions of the course, reading materials, and ways to improve the course.

Interviews. In addition to the various quantitative instruments used to measure outcome and help facilitate understanding, post-intervention semi-structured open-ended interviews were conducted with individuals identified by the group facilitators as having gained the most or least from participation in the groups. The interviews were semi-structured in the sense that all interviews included the same set of predetermined general questions to insure that the same basic areas were covered in each interview, but the interviewer was free to ask any relevant follow-up questions at any time, and interviewees were encouraged to share any additional information they felt was a relevant part of the experience for them.

Data Analysis

To answer research questions one and two, three quantitative instruments (M-GUDS, CoBRAS, and QDI) were utilized pretest/posttest. Separate analyses were performed, one for each of the three quantitative pretest/posttest instruments used in the study. Since the M-GUDS produces a single score, an analysis of covariance (ANCOVA) was performed. However, since the CoBRAS and QDI produce multiple factor scores, three each, multivariate analysis of variance (MANCOVA) were performed for each of these instruments. Pretest scores were used as a covariate primarily to adjust
the posttest means for differences among groups on the pretest, which were likely to occur due to the nonrandomized design of the study.

To answer research question three, several different measures and means of analysis were utilized. This included data collected weekly in the form of participant identified critical incidents and the self-report SEQ, which produces four scales measuring reactions to and perceptions of each session, as well as data collected post-intervention in the form of semi-structured follow-up interviews with 12 participants (two from each of the six groups). Unfortunately, data for the 12 follow-up interviews were determined to be of insufficient quality to allow proper content analyses. As a result, only the weekly data from the SEQ and the critical incident reports were analyzed.

The fundamental qualitative technique used to analyze the critical incident reports was a phenomenological approach, and the process of thematic categorization of critical incident events was guided by the coding system described by Merriam (1988). A peer examination team including the primary investigator and two others sorted each reported critical incident event into one of five categories identified by the primary investigator as having emerged from the data. Ultimately, only critical incident events that all raters unanimously categorized the same were considered. The percentages of critical incident reports in each category were then compared and reported by week, by condition (exposed to factual readings versus personal readings), and by race (White versus non-White). Patterns, themes, and potential hypotheses were discussed.

Mean scores on the four scales of the SEQ were also compared and reported by week, condition (exposed to factual readings versus personal readings), and race (White versus non-White). Again, patterns themes, and potential hypotheses were discussed.
Findings from the qualitative content analyses of the critical incident reports were also viewed in comparison with the SEQ findings to gain potentially greater understanding. Themes and patterns between the two forms of data were examined and discussed.

Finally, as an exploratory analysis meant to generate further hypotheses regarding what types of incidents lead to positive change, critical incident reports for participants whose differences in scores on the three quantitative measures (M-GUDS, CoBRAS, and QDI) from pretest to posttest represented the “Best” and “Worst” changes over the course of the intervention were compared. Emerging themes, patterns, and potential hypotheses were considered and discussed.

Results and Conclusions

Question One. The results from the ANCOVA performed to assess for changes in participants’ attitudes regarding similarities and differences, as demonstrated by pretest and posttest scores on the M-GUDS, supported the null hypothesis that no significant differences would be found between any of the conditions (i.e. – control, intervention with personal readings, intervention with factual readings) at the .05 significance level (F=.933, p=.398). The results form the MANCOVA performed to assess for changes in participants’ attitudes regarding racial and gender equality, as demonstrated by pretest and posttest scores on the three factors of the QDI, also supported the null hypothesis at the .05 level (F=1.511, p=.219). However, the MANCOVA performed to assess for changes in the participants’ color-blind racial attitudes, as demonstrated by pretest and posttest scores on the CoBRAS, rejected the null hypothesis at the .05 level (F=2.987, p=.037). The follow-up univariate tests indicated no significant differences between
conditions for CoBRAS factors I or factor II (F=1.730, p=.185 respectively). However, follow-up tests did find a statistically significant difference between conditions for the CoBRAS factor III, lack of awareness regarding blatant racial issues. Follow-up contrasts found statistically significant differences on this measure between the comparison condition and both intervention conditions, but no significant difference between the two intervention conditions.

These results suggest limited support for the overall efficacy of the intervention. However, there are several possible reasons why only awareness of blatant racial issues was found to be significantly different for those who participated in either form of the intervention compared to those who did not experience the intervention. First, this may indicate that intergroup contact, as represented by this particular intervention, is not effective in changing attitudes or awareness related to potentially more subtle issues such as institutional discrimination, racial privilege, etc., but can facilitate change for more obvious issues such as those measured by the CoBRAS factor III, awareness of blatant racial issues. These more subtle forms of potential bias or inequity do tend to be much harder to identify, and may also prove to be harder to change. Perhaps these more complex and less obvious forms of prejudice or bias are also easier to rationalize as something other than racial prejudice. This would then generate less dissonance with a non-prejudiced self-image, and perhaps be less likely internally to require change. Another hypothesis is that greater length or intimacy of contact may be required than was provided by the intervention in this study to effectively change more subtle or perhaps more deeply held attitudes. Finally, the samples in this study were relatively small, and somewhat skewed in a manner that may have produced ceiling effects for the participants.
in the experimental groups compared to the control. Specifically, scores on most measures found the White participants in the intervention conditions to be significantly more aware and egalitarian than their White counterparts in the comparison group at pretest. This difference may have resulted in significant ceiling effects that would mute potential positive results of the intervention. All of these possibilities suggest further research.

**Question Two.** No statistically significant differences were found on any of the measures between participants in the two conditions (i.e. – exposure to factual or personal readings). One likely explanation is that the assigned readings simply represented too small a part of the overall intervention, and thus lacked sufficient overall impact to significantly change the overall experience in any measurable way. Another possibility is that the personal articles did not differentially evoke increased empathy compared to the factual articles. While the articles were chosen by a team of raters based on the raters assessment of the articles’ ability to evoke empathy, it was never confirmed either quantitatively or qualitatively that the participants felt any greater empathy as a result of reading the personal articles compared to the factual ones. It seems that more needs to be done to first measure levels of empathy, and then compare any changes in attitude or bias before any conclusions can be drawn.

**Question Three.** This final research question examined more phenomenological aspects of the participants’ experiences of the intervention. This was done primarily through analyses of qualitative data collected on a weekly basis at the end of each intervention session. The content for this process included Session Evaluation Questionnaires (SEQ) and Critical Incident Reports. The fundamental qualitative
analytic technique used was based on a phenomenological approach. Analyses of the qualitative data revealed several interesting themes and patterns.

Through initial analysis of the critical incident reports by the primary investigator, five themes emerged. Participants identified critical incident events focused on **Self**, **Other**, **Topic**, **Group Process**, and **Structured Exercise**. A team of reviewers then rated each individual critical incident event for fit into one of the five categories that had been identified by the primary investigator. Once categorization of the critical incident reports by the review team was complete, then the findings were further analyzed by comparing the percentage of critical incidents falling into each category in relation to the time they occurred during the intervention, experimental condition, and race.

The analyses based on critical incident reports by time during the intervention explored potential patterns that emerged over the course of the intervention related to the percentage of the reported critical incident events fitting into each of the five categories. The findings revealed several interesting fluctuations and potential patterns. Hypotheses were generated to potentially help understand or explain these fluctuations and patterns.

The percentage of critical incident events reported that focused on “Self” fluctuated more than any of the other critical incident categories over the course of the intervention. One hypothesis is that focus on self is more susceptible to external interventions. A possible inverse relationship was noted between the percentage of incidents categorized as self and those categorized as “Structured Exercise.” It was noted that structured exercise was too infrequent to draw any definitive conclusions, but a relationship between the two categories would potentially support this hypothesis.
The percentage of reported critical incident events categorized as “Other” was much greater than any other category after the initial intervention session, and this category represented a relatively large percentage of the reported critical incident events throughout the intervention with the notable exception of the final session. One hypothesis is that this reflects typical group development with members assessing others and attempting to find common ground early, and, then, continuing to learn from each other over the life of the group.

Reports of critical incidents categorized as “Topic” peaked early and decreased over the course of the intervention. Hypotheses generated to explain this phenomenon were based on several possible forms of socialization. One was that, since focus on topic is typically encouraged in most didactic courses, the students involved in the intervention might have been thus socialized to focus primarily on topic. A second possibility is that societal norms make many of the topics discussed in the intervention “taboo,” which may have resulted in the topics themselves becoming more salient. The decrease in focus on topic may have been due to changes in group norms that occurred over the course of the group (e.g. facilitators encouraged participants to focus on interactions over topics, and the more topics were discussed the less taboo they became).

Analyses of critical incident reports based on condition (i.e. – assigned factual or personal readings) demonstrated virtually no difference. The largest difference based on condition was that participants exposed to personal readings were more likely to report incidents categorized as topic, while those exposed to factual readings were more likely to report incidents categorized as “Group Process.” These differences were unexpected
and relatively small, and may simply represent natural differences between groups. However, it may also warrant further study.

Analyses of critical incident reports compared by race (i.e. – White or non-White) found that the percentage of critical incidents fitting into each of the five categories were very similar regardless of race. The most notable difference found based on analyses by race was that participants identifying as White were more likely to report critical incidents categorized as other more often, while participants identifying as non-White were more likely to report critical incidents categorized as group process. It was hypothesized that since Whites tend to experience less interracial contact than non-Whites, they may focus more on the explicit aspects of the experience and information provided by others, while in comparison non-Whites may focus more on the “process” or other intangibles of the experience. This seems like a particularly important area for further research.

Analyses of the SEQ were performed based on all four scale/factor scores produced by the SEQ (i.e. – depth, smoothness, positivity, and arousal). Patterns and differences were identified and discussed for all four scores by session over the course of the intervention, as well as by experimental condition and also race. Findings from these analyses of the SEQ scale scores over the course of the intervention revealed that scores for Depth, Smoothness, and Positivity all started around the same point, while Arousal started much lower. Scores for all scales fluctuated very little around or near their starting points with the exception of Smoothness, which fluctuated more widely. It was also noted that scores on all four scales were above their average level for the nine-weeks on the final session, and all were also up for the final session compared to the previous
session. Additionally, it was noted that scale scores for smoothness and positivity somewhat paralleled each other throughout, while scores for depth and arousal seemed much more independent. It was hypothesized that early fluctuations in scale scores for Smoothness and Positivity may simply represent the volatility inherent in early stages of group development, and the eventual leveling off and upward trends may be the natural outcome of group development towards better cohesion once group norms become more established and understood. It was hypothesized that scores on Depth started and remained relatively high due to the very nature of these groups. Simply addressing topics that are typically avoided in intergroup contact situations may automatically constitute a certain level of depth for many people. One final hypothesis was that scores for all four scales were relatively elevated following the last session due to separation anxiety as participants came to terms with the fact that this was the end of what had come to be a meaningful and important experience for them.

Findings from the SEQ scale scores compared by condition (i.e. – assigned factual versus personal readings) demonstrated virtually no differences regardless of which type of readings the participants were assigned. Again, this was not a surprise given the relatively small portion of the intervention represented by the readings. There is also no empirical support to indicate that the personal readings evoked greater levels of empathy as they were intended. Thus findings by condition fail to provide evidence to either support or challenge the role of empathy in changing attitudes.

Findings from analyses comparing SEQ scale scores by race (i.e. – White or non-White) suggested that average scores across the entire course of the intervention were virtually the same for Depth, and notably higher for non-Whites compared to Whites on
Smoothness and Positivity, while the opposite was true for Arousal. When the comparison by race was examined session by session, it was found that scores on Smoothness, Positivity, and Arousal were consistent with the overall averaged scores (i.e. – comparatively, non-Whites rated every session higher on Smoothness and Positivity, while Whites rated every session higher for Arousal). However, examination of Depth scores session by session revealed that non-Whites rated early sessions significantly higher in terms of Depth compared to their White counterparts, but average scores on Depth increased steadily during the first half of the intervention for White participants while steadily decreasing for non-White participants. At week four of the intervention, the average scale scores for White participants on Depth surpassed scores for non-Whites, and after that point scores began to fluctuate up and down for both Whites and non-Whites, with what appeared to be a slight upward trend for both. Ultimately scale scores for Depth were up for the final session for both groups.

Findings based on analyses and interpretation of scale scores of the SEQ by race lead to several interesting hypotheses. First, the findings that average scores for Depth often headed in opposite directions from session to session depending on race, lead to the hypothesis that there may be significant difference in how Whites and non-Whites evaluate or interpret the depth of an interaction. Second, consistently lower average scores on Smoothness, and Positivity, coupled with consistently higher average scores on Arousal for White participants might indicate that they, likely having had fewer past contact experiences than their non-White counterparts, find such contact provokes greater arousal levels, and is also experienced as less smooth or positive. These differences may be accounted for simply by the relative novelty of the situation for Whites compared to
non-Whites, or, it may be as complex as White guilt associated with racial privilege, the manifestation of dissonance associated with aversive racism, or some combination of all of these things. It seems important to understand these potential differences in how Whites and non-Whites evaluate intergroup interactions, making this another potentially valuable area for further research.

Finally, critical incident reports were grouped and compared based on change from pretest to posttest on the three quantitative measures (i.e. – the M-GUDS, CoBRAS, and QDI). The two groups represented the most and least change in the desired direction from pretest to posttest on measures produced by the three quantitative instruments (i.e. – Increases in M-GUDS scores were considered positive since they indicated improvement in attitudes regarding awareness and acceptance of both similarities and differences. Increases in the QDI factor scores were also positive since they indicated improvement in attitudes regarding racial diversity, contact with racial diversity, and gender equity. However, decreases in the CoBRAS factor scores were considered positive since they indicated a decrease in lack of awareness). The critical incident reports associated with the participants with the most improvement in scores from pretest to posttest were placed in a group labeled “Best,” while critical incident reports associated with participants with the least improvement in scores from pretest to posttest were placed in a group labeled “Worst.” The percentage of the critical incidents reported in each of the five categories was then compared between groups.

The most obvious and interesting findings from this comparison were that the group identified as “Best” was more likely to identify and report critical incidents related to Topic notably more often than those in the group identified as “Worst” for the M-
GUDS, the CoBRAS factors I, II, and II, and the QDI factor I. One hypothesis was that topic is a critical component for positive intergroup contact, much more so than just the contact itself. Maybe it is crucial what is discussed during intergroup contact. This would support the idea that programmed and facilitated interventions are important since they would help make sure relevant topics were discussed. Another hypothesis is that the difference is more a reflection of what the instruments actually measure (i.e. – knowledge related to particular topics may be what is being measured, more than actual attitudes). A consideration leading to another potential hypothesis was that the only scores not seemingly more affected by focus on topic were scores on the QDI factors II and III, factors that seem to be potentially focused more specifically (i.e. – factor I: specific attitudes and personal contact, factor II: specifically focused on attitudes related to women’s equity). Thus the hypothesis was generated that topic may decrease in importance, while self-introspection and awareness of group process might increase in importance, as specificity of the attitudes being measured increases. The exploration of what specific aspects of intergroup contact lead to positive changes for both majority and minority participants is an area that has been historically neglected, and deserves further attention.

*Implications for Theory*

Findings from the current study provide several considerations for theory related to reduction of prejudice, and specifically to contact theory. First, it seems that contact theory may need to be expanded to consider potentially different needs depending on race, or possibly majority or minority standing. The theory also tends to assume that everyone needs the same components for a positive intergroup contact experience.
Perhaps ways that minority and majority group members differ regarding what they attend to, or how they interpret the same situation, may need to be better accounted for by theory. The findings in this study suggest that minority group members may attend to different aspects of the contact experience than majority group members (i.e. — white participants tended to focus more on topic and other while nonwhites focused more on group process). Therefore there may be a difference in what components members of various groups need in order for the contact to have a positive outcome for them. Thus, theory may need to expand to be able to differentiate what is needed for majority versus minority participants, or even further for specific subgroups. Further, the magnitude, duration, or intensity of the intervention might also need to be considered based on group or subgroup. Findings from this study suggest that arousal levels may differ a great deal for minority versus majority group members. The level of arousal necessary to engage or appropriately challenge one group or individual to promote positive change might be so great that it shuts another down.

Second, contact theory may need to be updated to account for the apparent evolution of prejudice. Modern forms of prejudice such as aversive racism tend to be much more subtle, both in their external manifestations as well as in the internal dissonance that is generated. The components identified as necessary by the current contact theory may not be enough to make clear these more subtle forms of bias. Perhaps contact historically would by definition bring biases and negative attitudes about other groups into awareness, now it may be necessary for specific relevant topics to be addressed in methods that make the more subtle biases associated with modern prejudices accessible. Further research is needed to identify such methods and their outcomes.
Implications for Practice

One important, if indirect, pragmatic finding of this study was that programmed intergroup interventions can be run through a University Counseling Center. Such interventions can fit into the primary mission of a University Counseling Center by providing a service to university students in the form of the contact groups, and also serve as a valuable and unique training experience for students training at the center. The efficacy of such a program received limited empirical support from this study, but anecdotal support in the form of enrollment and positive feedback from participants has been very positive and abundant. This type of program is potentially a great opportunity for a unique education experience for participants, training for counseling psychology students who serve as facilitators, and fertile ground for further “real world” research on intergroup/intercultural contact, an area that all indications suggest will only continue to grow in importance.

Implications for Training

While the intervention serving as the focus of this study, Dialogues on Race, was initially developed as a means to provide positive intergroup contact experiences for the participants, it has become increasingly evident that it also has the potential to provide a unique role in the training of the facilitators. Given the proper support and resources, this type of program is a great opportunity for counseling psychology trainees to engage in experiential learning as part of their group supervision, gain experience in group facilitation, and get exposure to diverse populations. With proper instruction and supervision, this type of program or intervention represents a unique opportunity to serve
as a social advocate, provide a service, and also gain in all areas of cultural competency (i.e. - awareness, knowledge and skills).

Implications for Future Research

The major implications for future research stem from the tentative finding that participants in intergroup contact experiences may attend to different aspects of the experience based on race, or minority versus majority status. This finding again emphasizes the need already identified in the literature, that future research includes both majority and minority participants, along with interrelationships between the two. Since previous research, as well as findings in this study, provides mixed results regarding the efficacy of intergroup contact, it suggests that further outcome research is warranted. However, findings from this study also suggest that it is crucial to consider the content and process of the intergroup contact as well. Described below are suggestions for replicating this study as well as suggestions for additional research topics that emerged from the results of this study.

Suggestions for Replicating this Study

As noted throughout, there were numerous limitations to this study that could be potentially minimized or eliminated in any future replications. First, the loss of the data from the semi-structured follow-up interviews was significant. In future studies the collection of interview data should be done with greater care and better equipment. The interviews were by far the most flexible means in the study for collecting relevant data, and it was a major loss to not be able to analyze that data.
Another suggestion would be to include larger more representative samples. One recommendation regarding how to achieve a larger sample would be to collect data for more than one semester, or intervention cycle. Also, participants in the experimental groups were self-selected, which ultimately led to a less representative sample compared to the control group, and likely to the larger student body. One suggestion is that a comparison group might be formed by students self-selecting to participate in the intervention groups, but placed on a wait list until the next semester. Another possibility would be to make participation in the groups mandatory for some university students and use that same population as the comparison.

Another limitation of the current study was that the experimental manipulation designed to examine the role of empathy was too small a percentage of the overall intervention to know for sure if it had an effect. Further, the “personal” readings chosen and assigned to evoke greater levels of empathy as compared to the “factual” readings, were not empirically validated regarding whether or not they succeeded in effectively eliciting the greater level of empathy as designed. One suggestion is that levels of empathy be empirically examined, and that empathy be looked at separate from the overall outcome results (e.g. – perhaps levels of empathy should be examined for each session and then some form of session evaluation or productivity could be the measure on the effect of increased empathy).

Another suggestion for replicating this study would be to collect more in depth weekly data. While the critical incident reports provided a great deal of information, it was often hard to place them into context. Perhaps collecting data in the form of
participant weekly journals would provide more context and also greater insight into each participant’s experience.

One final consideration might be to include follow-up data collection at six months, one year, and possibly even later. Feedback from participants has suggested that this type of intervention might initiate a type of vector change, the ramifications of which might not be truly felt for some time. For example, I have had students well into the next semester who have emailed stating that they find themselves going to different diversity related events, or engaging in different conversations than they would have prior to their participation in the intervention. They also noted that they hear and see things from a different perspective. One student wrote that her participation in the groups had “ruined her vacation” with her family because she was now conscious of the biases in many of their statements and actions. This all indicates that positive outcomes may sometimes be gradual, and may not be fully appreciated unless examined longitudinally.

Suggestions for Additional Research Topics

This research study suggested two additional research topics. The first is a result of the findings that participants may differ significantly on what aspects of the intervention they attune to based on either race or majority/minority status. This suggests that it might be important to measure both majority and minority reactions to specific interactions. It seems critical to understand what participants mean to present, how it is interpreted, and what causes any discrepancies between the two. Not only is it important that future research include both majority and minority perspectives, but also the intersection of majority and minority interactions.
The second additional research topic was generated by several factors. First was the idea that modern forms of prejudice or racism tend to be less conscious, combined with the fact that no significant outcome differences were found on most of the self-report quantitative measures, led to the suggestion that measures of less consciously controllable reactions may be important (i.e. - measurements that would be more likely to capture modern forms of bias or prejudice that may not be reflected by reports of conscious attitudes). Perhaps signs of aversive reactions could be measured through things such as eye movement, heart rate, galvanic skin response, or certain changes in voice that can indicate stress or anxiety.

In summary, this research study has attempted to add to the body of research that has investigated prejudice reduction and intergroup contact. This study was unique in its methodology, using qualitative as well as quantitative research, including both minority and majority group members, exploring the role of empathy by attempting to experimentally manipulate empathy levels through different types of assigned readings, and by examining an intervention program run through a University Counseling Center. The qualitative analyses added insight and explanation to some of the quantitative findings, and also generated further hypotheses for future consideration.

One final note, a major part of this study was to determine if an intervention program with a design based on intergroup contact theory and utilizing programmed and facilitated groups could be effectively run through a University Counseling Center. The fact that the Dialogues on Race program has continued to grow and evolve, and is now in its forth year, would indicate that it is possible with sufficient support and commitment
particularly on the part of the Counseling Center, but also throughout the University. A brief review and update of the program are included in appendix J.
REFERENCES


cognitive analysis of the effects of intergroup contact. *Journal of social Issues,*
41(3), 81-104.

development and cross-cultural counselor training: A stage model. *The*
*Counseling Psychologist,* 19, 76-102.


S. Tisdale, L. Heath, & E. J. Posavec (Eds.), *Social influence processes and*

*Social Inq. 64,* 438-455.

multicultural inventory: A self-report measure of multicultural competencies.
*Journal of Counseling Psychology,* 41, 137-148.

(Ed.), *Review of personality and social psychology: Group processes and*

*Sch.,* 2, 145-158.


APPENDIX A

DIALOGUES ON RACE INVITATION / APPLICATION
YOU ARE INVITED TO PARTICIPATE IN:

A Dialogue on Race

Beginning the week of February 11th, these intensive discussion groups will meet for 2 hours per week and last the duration of the semester. Each group will be co-led by trained facilitators and consist of approximately 8-10 students from diverse backgrounds.

The purpose is to encourage and develop honest communication that will lead to a better understanding of others…and oneself.

Course credit is available.

If you want more information or are interested in joining us, please send the following information to <mdw166@psu.edu> by Wednesday, January 30, 2002

Name:     Phone #: 
Gender:    Age: 
Email Address: 
Major: 
Semester you plan to graduate: 
Grad or undergrad status: 
Race/ethnic background: 
Student Number:
Indicate your top three group meeting times in order of preference:

___ Mondays:  5:30  7:30pm
___ Wednesdays:  2:45  4:45pm
___ Wednesdays:  3:30  5:30pm
___ Wednesdays:  6:00  8:00pm
___ Thursdays:  2:00  4:00pm
___ Thursdays:  6:00  8:00pm

Thanks for your interest.

D'Andre Wilson and Laurie Mulvey
Supervisors
APPENDIX B

EXPLANATION OF STUDY / INFORMED CONSENT FORMS

- Intervention Participant Form
- Control Participant Form
- Interview Participant Form
INFORMED CONSENT FORM

The Pennsylvania State University

Title of Project: Dialogues on race: Process and effects (IRB#02B0041)
Principal Investigator: M. D’Andre Wilson, 242 Ritenour, University Park, PA 16802
(814) 863-0395 mdw166@psu.edu

1. Purpose of the Study: The purpose of this research is to assess the effectiveness of the “Dialogues on Race” groups at promoting open and honest conversation about race, and to explore what factors influence the process.

2. Procedures to be followed: Participation in this research will not require any activities beyond those expected as part of participation in the dialogue group. All group participants will be asked to answer approximately 100 questions on several survey forms prior to the start of the first group session, again following the final group session, and to answer approximately 50 questions at the end of each session. Completing the various survey forms will be expected regardless of whether or not you choose to participate in the research study. Agreement to participate in the research study simply gives consent to the analysis of your participation for research purposes. If you wish to participate in the group but not in the research study, you will still be expected to complete all forms, however, any information that you provide will be excluded from formal analysis or reporting as part of the research.

3. Discomforts and Risks: There are no risks in participating in this research beyond those experienced in everyday life. Some of the questions are personal and might cause discomfort.

4. Benefits:
   a. The knowledge that your contributions may help to improve university efforts to promote open and honest dialogue about race.
   b. This research might provide a better understanding of the processes and benefits of dialogues on race, and also on what factors influence this process. This information could help plan programs and make student services better.

5. Duration/Time: Participation in the research will not require any time beyond participation in the group. Time will be provided during the dialogue groups to complete all of the forms that are part of this research. It will take approximately 30 minutes during the first session, 10 minutes at the end of each session, and 20 minutes at the end of the final session to complete the questions.

6. Statement of Confidentiality: Only the person in charge and his assistants will know your identity. In fact, the primary investigator will not know whether or not you choose to participate in the study until after grades have been reported for the semester. If this research is published, no information that would identify you will be written.

7. Right to Ask Questions: You have the right to ask questions regarding the research. All questions should be directed to D’Andre Wilson at 863-0395.

8. Compensation: There is no compensation involved with participating in this research.

Continued on Back ->
9. Voluntary Participation: You do not have to participate in this research. You can end your participation at any time by telling D'Andre Wilson. You do not have to answer any questions that you do not want to answer either on the written forms or verbally should you be asked, and agree, to be interviewed.

You must be 18 years of age or older to consent to participate in this research study. If you consent to participate in this research study and to the terms above, please sign your name and indicate the date below.

You will be given a copy of this consent form to keep for your records.

Participant Signature __________________________ Date __________________________

I, the undersigned, verify that the above informed consent procedure has been followed.

Investigator Signature __________________________ Date __________________________
INFORMED CONSENT FORM
The Pennsylvania State University

Title of Project: Dialogues on race: Process and effects (IRB#02B0041)
Principal Investigator: M. D’Andre Wilson, 242 Ritenour, University Park, PA 16802
(814) 863-0395 mdw166@psu.edu

10. Purpose of the Study: The purpose of this research is to assess the effectiveness of the “Dialogues on Race” groups at promoting open and honest conversation about race, and to explore what factors influence the process. This will be contrasted with students not participating in these groups, but having exposure to the general diversity opportunities provided at the University.

11. Procedures to be followed: Participation will be asked to answer approximately 100 questions on several survey forms once, approximately ten weeks prior to the end of the semester, and again the final week of classes for the semester. If you choose to participate, you need to sign and return this consent form, and complete and return all included materials. All materials should be returned in the included envelope. The envelope should be sealed and returned to your course instructor at the next class meeting, or, if for any reason you are unable to return the materials to the course instructor, you can simply drop the sealed envelope in intra-campus mail at the HUB information desk. To be included in the study, you must complete the research materials at both times that they are required, and return them within 5-days. If you return a signed consent form, you will automatically be included to receive a follow-up packet during the final week of classes.

12. Discomforts and Risks: There are no risks in participating in this research beyond those experienced in everyday life. Some of the questions are personal and might cause discomfort.

13. Benefits:
   a. The knowledge that your contributions may help to improve university efforts to promote open and honest dialogue about race.
   b. This research might provide a better understanding of the processes and benefits of dialogues on race, and also on what factors influence this process. This information could help plan programs and make student services better.

14. Duration/Time: Participation in the research will take approximately 30 minutes to complete the first set of survey instruments, and approximately 20 minutes at the end of the semester to complete the second set of survey instruments.

15. Statement of Confidentiality: Only the person in charge and his assistants will know your identity, and your name will not be coded and entered with your responses. If this research is published, no information that would identify you will be written.

16. Right to Ask Questions: You have the right to ask questions regarding the research. All questions should be directed to D’Andre Wilson at 863-0395.

Continued on Back->
17. Compensation: There is no guaranteed compensation for your participation in this research. However, all participants who complete and return all research materials within 5-days of receiving them (at both times required for the study) will be included in a random drawing to be held on 5/2/02. The first three participants selected will receive $50 each.

18. Voluntary Participation: You do not have to participate in this research. You can end your participation at any time by telling D’Andre Wilson. You do not have to answer any questions that you do not want to answer.

You must be 18 years of age or older to consent to participate in this research study. If you consent to participate in this research study and to the terms above, please sign your name and indicate the date below.

You will be given a copy of this consent form to keep for your records.

______________________________________  _____________________
Participant Signature     Date

I, the undersigned, verify that the above informed consent procedure has been followed.

______________________________________  _____________________
Investigator Signature     Date

**Contact Information for Drawing:**

______________________________________  _________________________
Name (please print clearly)    Local Phone Number

______________________________________
Local Address
INFORMED CONSENT FORM

The Pennsylvania State University

Title of Project: Dialogues on race: Process and effects (IRB#02B0041)
Principal Investigator: M. D’Andre Wilson, 242 Ritenour, University Park, PA 16802
(814) 863-0395 mdw166@psu.edu

19. Purpose of the Study: The purpose of this research is to assess the effectiveness of the “Dialogues on Race” groups at promoting open and honest conversation about race, and to explore what factors influence the process.

20. Procedures to be followed: Participation in this portion of the research will require you to participate in an individual semi-structured interview with the primary investigator. The interview will last approximately ½ hour and will focus on gaining a better understanding of your personal views and experiences as a participant in the dialogue groups. The interview will be audio-taped for later analysis.

21. Discomforts and Risks: There are no risks in participating in this research beyond those experienced in everyday life. Some of the questions are personal and might cause discomfort.

22. Benefits:
   a. The knowledge that your contributions may help to improve university efforts to promote open and honest dialogue about race.
   b. This research might provide a better understanding of the processes and benefits of dialogues on race, and also on what factors influence this process. This information could lead to improvements to the dialogue on race groups and also help plan other programs and make student services better.

23. Duration/Time: Participation in this portion of the research will take approximately 30 minutes.

24. Statement of Confidentiality: Only the person in charge and his assistants will know your identity. If you are receiving course credit for participation in the dialogue group, agreeing or refusing to participate in this portion of the study will have no effect on your grade for the course. In fact, final grades for the course will be calculated prior to the interview. If this research is published, no information that would identify you will be included. Audio taped recordings of the interview will be stored by the primary investigator in a locked cabinet except when the transcriptionist has possession of them, during those times she will be responsible for keeping the tapes in a locked and secured location. Only the primary investigator and the persons hired to transcribe the tapes will have access to them. Tapes will be transcribed using first names only as identifiers. Names will be replaced with identification codes, matching the other study materials, during the process of coding the data. All interview tapes will be cut-up and incinerated not more than one year following their recording.

25. Right to Ask Questions: You have the right to ask questions regarding the research. All questions should be directed to D’Andre Wilson at 863-0395.

Continued on Back ->
26. Compensation: Participants interviewed for this portion of the research will be compensated $5 for their time.

27. Voluntary Participation: You do not have to participate in this interview or research study. You can end your participation at any time by telling D’Andre Wilson. You do not have to answer any questions that you do not want to answer.

You must be 18 years of age or older to consent to participate in this research study. If you consent to participate in this research study and to the terms above, please sign your name and indicate the date below.

You will be given a copy of this consent form to keep for your records.

______________________________________  _____________________
Participant Signature     Date

I, the undersigned, verify that the above informed consent procedure has been followed.

______________________________________  _____________________
Investigator Signature
APPENDIX C

PERSONAL DATA SHEET / DEMOGRAPHIC QUESTIONAIRRE
Demographic Questionnaire

Last 4-digits SS# : __________

Please answer all questions by responding with the answer that best reflects your current situation.

1. What is your age? _______

2. What is your gender?
   1. Male  2. Female

3. Which of the following best describes your race, ethnicity, or cultural background?
   7. Other: ______________

4. What is your class standing?
   1. 1st Year  2. 2nd Year  3. 3rd year  4. 4th year  5. 5th year
   6. 6th year or more  7. Graduate Student

5. What is your major area of study? ____________________________
APPENDIX D

*Miville-Guzman Universality-Diversity Scale* (M-GUDS; Miville, 1992)
Miville-Guzman Universality-Diversity Scale (M-GUDS)

The following items are made up of statements using several terms which are defined below for you. Please refer to them throughout the rest of the questionnaire.

Culture refers to the beliefs, values, traditions, ways of behaving, language of any social group. A social group may be racial, ethnic, religious, etc.

Race or racial background refers to a sub-group of people possessing common physical or genetic characteristics. Examples include White, Black, American Indian.

Ethnicity or ethnic group refers to specific social group sharing a unique cultural heritage (i.e., customs, beliefs, language, etc.). Two people can be of the same race (e.g., White), but be from different ethnic groups (e.g., Irish-American, Italian-American).

Country refers to groups that have been politically defined; people from these groups belong to the same government (e.g., France, Ethiopia, United States). People of different races (White, Black, Asian) or ethnicities (Italian, Japanese) can be from the same country (United States).

Instructions: Please indicate how descriptive each statement is of you by filling in the number corresponding to your response. This is not a test, so there are no right or wrong, good or bad answers. All responses are anonymous and confidential.

1. _____ I am interested in knowing people who speak more than one language.
2. _____ It deeply affects me to hear persons from other countries describe their struggles of adapting to living here.
3. _____ I attend events where I might get to know people from different racial backgrounds.
4. _____ I feel a sense of connection with people from different countries.
5. _____ I am not very interested in reading books translated from another language.
6. _____ Knowing about the experiences of people of different races increases my self understanding.
7. _____ I sometimes am annoyed at people who call attention to racism in this country.
8. _____ Knowing someone from a different ethnic group broadens my understanding of myself.
9. _____ Knowing how a person differs from me greatly enhances our friendship.
10. _____ I don’t know too many people from other countries.
11. _____ I place a high value on being deeply tolerant of others’ viewpoints.
12. _____ It’s really hard for me to feel close to a person from another race.
13. _____ It grieves me to know that many people in the Third World are not able to live as they would choose.
14. _____ I would like to join an organization that emphasizes getting to know people from different countries.
15. _____ In getting to know someone, I try to find out how I am like that person as much as how that person is like me.
16. _____ When I hear about an important event (e.g., tragedy) that occurs in another country, I often feel as strongly about it as if it had occurred here.
17. _____ It’s hard to understand the problems that people face in other countries.
18. _____ I can best understand someone after I get to know how he/she is both similar and different from me.
19. _____ I often feel irritated by persons of a different race.
20. _____ It does not upset me if someone is unlike myself.
21. _____ I would like to know more about the beliefs and customs of ethnic groups who live in this country.
22. _____ It’s often hard to find things in common with people from another generation.
23. _____ When I listen to people of a different race describe their experiences in this country, I am moved.
24. _____ I often feel a sense of kinship with persons from different ethnic groups.
25. _____ I would be interested in participating in activities involving people with disabilities.
26. _____ Knowing about the different experiences of other people helps me understand my own problems better.
27. _____ Persons with disabilities can teach me things I could not learn elsewhere.
28. _____ I am often embarrassed when I see a person with disabilities.
29. _____ I am only at ease with people of my race.
30. _____ I would like to go to dances that feature music from other countries.
31. _____ For the most part, events around the world do not affect me emotionally.
32. _____ Placing myself in the shoes of a person from another race is usually too tough to do.
33. _____ I often listen to the music of other cultures.
34. _____ If given another chance, I would travel to different countries to study what other cultures are like.
35. _____ I have friends of differing ethnic origins.
36. _____ Knowing how a person is similar to me is the most important part of being good friends.
37. _____ It is important that a friend agrees with me on most issues.
38. _____ In getting to know someone, I like knowing both how he/she differs from me and is similar to me.
39. _____ Getting to know someone of another race is generally an uncomfortable experience for me.
40. _____ I would be interested in taking a course dealing with race relations in the United States.
41. _____ Becoming aware of experiences of people from different ethnic groups is very important to me.
42. _____ I am interested in learning about the many cultures that have existed in this world.
43. _____ I am interested in going to exhibits featuring the work of artists from different minority groups.
44. _____ I feel comfortable getting to know people from different countries.
45. _____ I have not see many foreign films.
APPENDIX E

THE COLOR-BLIND RACIAL ATTITUDES SCALE

(CoBRAS; Neville, Lilly, Lee, Duran, and Brown, 2000)
Directions. Below is a set of questions that deal with social issues in the United States. Using the 6-point scale, please give your honest rating about the degree to which you personally agree or disagree with each statement. Please be as open and honest as you can; there are no right or wrong answers. Record your response to the left of each item.

<table>
<thead>
<tr>
<th>Item</th>
<th>Statement</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Everyone who works hard, no matter what race they are, has an equal chance to become rich.</td>
</tr>
<tr>
<td>2</td>
<td>Race plays a major role in the type of social services (such as type of health care or day care) that people receive in the U.S.</td>
</tr>
<tr>
<td>3</td>
<td>It is important that people begin to think of themselves as Americans and not African American, Mexican American or Italian American.</td>
</tr>
<tr>
<td>4</td>
<td>Due to racial discrimination, programs such as affirmative action are necessary to help create equality.</td>
</tr>
<tr>
<td>5</td>
<td>Racism is a major problem in the U.S.</td>
</tr>
<tr>
<td>6</td>
<td>Race is very important in determining who is successful and who is not.</td>
</tr>
<tr>
<td>7</td>
<td>Racism may have been a problem in the past, but it is not an important problem today.</td>
</tr>
<tr>
<td>8</td>
<td>Racial and ethnic minorities do not have the same opportunities as White people in the U.S.</td>
</tr>
<tr>
<td>9</td>
<td>White people in the U.S. are discriminated against because of the color of their skin.</td>
</tr>
<tr>
<td>10</td>
<td>Talking about racial issues causes unnecessary tension.</td>
</tr>
<tr>
<td>11</td>
<td>It is important for political leaders to talk about racism to help work through or solve society’s problems.</td>
</tr>
<tr>
<td>12</td>
<td>White people in the U.S. have certain advantages because of the color of their skin.</td>
</tr>
<tr>
<td>13</td>
<td>Immigrants should try to fit into the culture and adopt the values of the U.S.</td>
</tr>
<tr>
<td>14</td>
<td>English should be the only official language in the U.S.</td>
</tr>
<tr>
<td>15</td>
<td>White people are more to blame for racial discrimination in the U.S. than racial and ethnic minorities.</td>
</tr>
<tr>
<td>16</td>
<td>Social policies, such as affirmative action, discriminate unfairly against White people.</td>
</tr>
<tr>
<td>17</td>
<td>It is important for public schools to teach about the history and contributions of racial and ethnic minorities.</td>
</tr>
<tr>
<td>18</td>
<td>Racial and ethnic minorities in the U.S. have certain advantages because of the color of their skin.</td>
</tr>
<tr>
<td>19</td>
<td>Racial problems in the U.S. are rare, isolated situations.</td>
</tr>
<tr>
<td>20</td>
<td>Race plays an important role in who gets sent to prison.</td>
</tr>
</tbody>
</table>
APPENDIX F

THE QUICK DISCRIMINATION INDEX

(QDI; Ponterotto et al., 1995)
Social Attitude Survey

Last 4-digits SS# : ___________

Please respond to all items in the survey. Remember there are no right or wrong answers.

The survey is completely anonymous; do not put your name on the survey. Record your response to the left of each item.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly</td>
<td>Disagree</td>
<td>Not sure</td>
<td>Agree</td>
<td>Strongly</td>
<td></td>
</tr>
<tr>
<td>Disagree</td>
<td>Agree</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. _____ I do think it is more appropriate for the mother of a newborn baby, rather than the father, to stay with the baby (not work) during the first year.
2. _____ It is as easy for women to succeed in business as it is for men.
3. _____ I really think affirmative action programs on college campuses constitute reverse discrimination.
4. _____ I feel I could develop an intimate relationship with someone from a different race.
5. _____ All Americans should learn to speak two languages.
6. _____ It upsets (or angers) me that a woman has never been president of the United States.
7. _____ Generally speaking, men work harder than women.
8. _____ My friendship network is very racially mixed.
9. _____ I am against affirmative action programs in business.
10. _____ Generally, men seem less concerned with building relationships than women.
11. _____ I would feel O.K. about my son or daughter dating someone from a different race.
12. _____ It upsets (or angers) me that a racial minority person has never been president of the United States.
13. _____ In the past few years there has been too much attention directed toward multicultural or minority issues in education.
14. _____ I think feminist perspectives should be an integral part of the higher education curriculum.
15. _____ Most of my close friends are from my own racial group.
16. _____ I feel somewhat more secure that a man rather than a woman is currently president of the United States.
17. _____ I think it is (or would be) important for my children to attend schools that are racially mixed.

18. _____ In the past few years there has been too much attention directed towards multicultural or minority issues in business.

19. _____ Overall, I think racial minorities in America complain too much about racial discrimination.

20. _____ I feel (or would feel) very comfortable having a woman as my primary physician.

21. _____ I think the president of the United States should make a concerted effort to appoint more women and racial minorities to the country’s Supreme Court.

22. _____ I think White people’s racism toward racial minority groups still constitutes a major problem in America.

23. _____ I think the school system, from elementary school through college, should encourage minority and immigrant children to learn and fully adopt traditional American values.

24. _____ If I were to adopt a child, I would be happy to adopt a child of any race.

25. _____ I think there is as much female physical violence towards men as there is male physical violence toward women.

26. _____ I think the school system, from elementary school through college, should promote traditional American values as well as the values representative of the culturally diverse students in the class.

27. _____ I believe that reading the autobiography of Malcolm X would be of value.

28. _____ I would enjoy living in a neighborhood consisting of a racially diverse population (i.e., Asians, Blacks, Hispanics, Whites).

29. _____ I think it is better if people marry within their race.

30. _____ Women make too big of deal out of sexual harassment issues in the workplace.
APPENDIX G

SESSION EVALUATION QUESTIONAIRRE
Participant Postsession Questionnaire

Last 4-digits SS# : _________ Date : _________

This session was:
Bad 1 2 3 4 5 6 7 Good
Safe 1 2 3 4 5 6 7 Dangerous
Difficult 1 2 3 4 5 6 7 Easy
Valuable 1 2 3 4 5 6 7 Worthless
Shallow 1 2 3 4 5 6 7 Deep
Relaxed 1 2 3 4 5 6 7 Tense
Unpleasant 1 2 3 4 5 6 7 Pleasant
Full 1 2 3 4 5 6 7 Empty
Weak 1 2 3 4 5 6 7 Powerful
Special 1 2 3 4 5 6 7 Ordinary
Rough 1 2 3 4 5 6 7 Smooth
Comfortable 1 2 3 4 5 6 7 Uncomfortable

Right now I feel:
Happy 1 2 3 4 5 6 7 Sad
Angry 1 2 3 4 5 6 7 Pleased
Moving 1 2 3 4 5 6 7 Still
Uncertain 1 2 3 4 5 6 7 Definite
Calm 1 2 3 4 5 6 7 Excited
Confident 1 2 3 4 5 6 7 Afraid
Wakeful 1 2 3 4 5 6 7 Sleepy
Friendly 1 2 3 4 5 6 7 Unfriendly
Slow 1 2 3 4 5 6 7 Fast
Energetic 1 2 3 4 5 6 7 Peaceful
Involved 1 2 3 4 5 6 7 Detached
Quiet 1 2 3 4 5 6 7 Aroused

Please briefly describe the exchange or interaction that you feel was the most significant for you personally during this session.
APPENDIX H

COURSE EVALUATION SHEET
Dialogues On Race  
Questionnaire - Spring 2002

Your responses to this survey are intended to help make changes / improvements to the dialogues on race groups. This questionnaire will not be reviewed until after final grades are submitted for the semester. Thank you for your assistance.

Last 4-digits SS#: ________________________ Date:__________________________

1. Did your dialogue group meet your expectations?  
   Fully: 1  Somewhat: 2  Not At All: 5
   Comments:

2. Was your group facilitated in a manner that encouraged open and honest dialogue?  
   Fully: 1  Somewhat: 2  Not At All: 5
   Comments:

3. Would you recommend participation in a similar group to a friend?  
   Definitely: 1  Maybe: 3  No Way: 5
   Comments:

4. Did your participation in this group help you to explore/challenge your views and/or biases regarding race and race relations?  
   Very Much: 1  Somewhat: 2  Not At All: 5
   Comments:

5. Did the assigned readings provide additional insight toward your understanding of race and race relations?  
   Very Much: 1  Somewhat: 2  Not At All: 5
   Comments:

6. Did the assigned readings help you gain greater empathy for the experiences of others?  
   Very Much: 1  Somewhat: 2  Not At All: 5
   Comments:

7. How do you feel the assigned readings facilitated your group’s ability to engage in open and honest dialogue?  
   Very Much: 1  Somewhat: 2  Not At All: 5
   Comments:
8. What did you like most about your experience in the dialogue group?

9. What did you like least about your experience in the dialogue group?

10. How has participation in this dialogue group affected you personally?

11. What suggestions do you have for improving these groups?

12. Please list other classes or organized activities in which you participated during the past nine weeks which also focused on issues of race or race relations.
APPENDIX I

BEYOND THIS STUDY:

A HISTORY AND UPDATE OF THE DIALOGUES ON RACE PROGRAM
DIALOGUES ON RACE: BEYOND THIS STUDY

This section provides background and update information on the program that was the focus of this study, Dialogues on Race. The program was started at the Pennsylvania State University in the Spring 2001 semester to bring racially and ethnically diverse students together to look at the ways culture and race affects their views of the world, and their views of others. Primarily under the auspices of the Center for Counseling and Psychological Services (CAPS), three groups were offered to twenty-two students during the first semester. Four undergraduate seniors, a CAPS intern, and a CAPS graduate assistant (GA) facilitated the groups. As a result of the first semester’s experience, it was determined that the program would be better served if all of the facilitators had some background training in counseling or psychology, and were more experienced. Starting in the fall of 2001, the program was pulled completely under CAPS. This was possible in part, as a result of temporary funding from the Equal Opportunity Planning Committee (EOPC) of the University. That semester all groups were led by CAPS interns, GA’s, or graduate students from Counseling or Clinical psychology, hired by CAPS. Group size was increased to twelve students, and six groups were run. The next semester the program received an additional endorsement from The Department of Counselor Education, Counseling Psychology, and Rehabilitation Services in the form of an offer to provide students course credit for their participation. This was also the semester that this research study was conducted. Since that time, participants have continued to receive course credit, and the facilitator training and supervision has become more of a priority. Demand for participation has continued to grow, and the
number of qualified facilitators is typically the main limit to the number of groups offered each semester.

It became increasingly obvious that facilitating these groups required additional training and supervision beyond that provided by what students were receiving in their various graduate programs. Students were consistently reporting that many of the issues and emotions being generated in their groups also represented issues that they themselves needed to process further. As a result, as a separate graduate level course, *Facilitating Intergroup/Intercultural Contact*, was developed. Group supervision for the Dialogues on Race group facilitators also includes what is essentially a process group for the facilitators, along with standard supervision activities.

Another component of the program that premiered in the Spring of 2004, is a collaborative effort with Shaver’s Creek, a nature center owned and operated by the University. This component was initiated by Sandy Newes, a CAPS intern from 2002-2003, as part of her outreach specialty. It involves a day long culturally focused training experience at Shaver’s Creek. The experience at Shaver’s Creek adds additional perspective and training on experiential learning for the graduate students in the *Facilitating Intergroup/Intercultural Contact* course, and the participants in the *Dialogues on Race* groups gain a culturally focused teambuilding experience that helps their groups bond and gain cohesion more quickly.

In the Spring of 2005 we invited “Straight Talks” to come into a session so that Dialogues related to sexual orientation could be more here and now without putting undo pressure on group members who might identify as gay or lesbian.
Feedback regarding the effectiveness of the overall Dialogues Program has been very positive thus far. Students participating in the groups have overwhelmingly endorsed the program by recommending participation to friends, asking to participate for a second semester, and commenting on the ways in which their participation has challenged them and promoted their personal growth and understanding. Graduate students receiving training, experience, and supervision through the Dialogues Program have reported the opportunity as a unique addition to their other graduate training. They have consistently reported that it has helped them in their personal and professional growth, and is of interest to potential employers. The hope is for this program to continue to evolve, improve, and potentially expand, through ongoing evaluation and programmatic research.
VITA

Marvin D’Andre Wilson

**Education**

Doctor of Philosophy in Counseling Psychology, May 2006
The Pennsylvania State University, University Park, PA

Master of Arts in Counseling Psychology and Social Psychology, May 1995
Ball State University, Muncie, IN

Bachelor of Science in Computer Science, August 1989
Minors: Psychology & Business
Ball State University, Muncie, IN

**Counseling Experience**

Member of Senior Staff – The Pennsylvania State University
Counseling & Psychological Services – August 1999 to Present

Partner/co-facilitator – CHOICES: Alternatives to Violence, Centre County, PA
County contract to work with court mandated perpetrators of domestic violence – November 2000 to January 2001

Pre-Doctoral Intern - The Pennsylvania State University
Counseling & Psychological Services – August 1998 to August 1999

Counseling Practicum & Advanced Practicum - The Pennsylvania State University

Counseling Practicum - The Pennsylvania State University

Counseling Practicum - The Pennsylvania State University

Pre-Master's Intern – Family Services Society, Inc., Marion, IN – Aug 1994 to May 1995

Counseling Practicum – Ball State University, Muncie, IN
University Counseling Clinic – January 1994 to August 1994