JUST SAY “NO” (AND MEAN IT): MEANINGFUL NEGATION AS A TOOL TO MODIFY AUTOMATIC RACIAL PREJUDICE

THESIS

Presented in Partial Fulfillment of the Requirements for
the Degree Master of Arts in the Graduate School
of The Ohio State University

By

India R. Johnson, B. A.
Graduate Program in Psychology

The Ohio State University
2009

Master’s Examination Committee:

Richard E. Petty, Advisor
Russell H. Fazio
Lisa K. Libby
ABSTRACT

The present research examines whether negation training can successfully reduce the automatic activation of racial prejudice. A pilot study showed that simply responding “no” when presented with unfavorable traits about African Americans was not effective in reducing automatic prejudice, but responding with a more meaningful “That’s wrong” was successful. To further explore this finding, in Study 1, participants were trained to either meaningfully negate prejudice-consistent information or to meaningfully negate prejudice-inconsistent information, and completed an evaluative priming measure of racial prejudice before and after training. Consistent with the Meta-cognitive Model (MCM; Petty, Briñol, & DeMarree, 2007) of attitudes, participants who engaged in meaningful negation of prejudice-consistent information significantly reduced their levels of automatic racial prejudice relative to those who negated prejudice-inconsistent information. Furthermore, in Study 2 this effect was moderated by participant’s motive to control for prejudiced reactions (MCPR), such that those participants who were high in MCPR demonstrated the greatest observed changes in racial prejudice from time 1 to time 2, relative to those participants low in MCPR. Contrary to recent research suggesting that negation training is an ineffective means to reduce automatic racial prejudice, the present research suggests that negation can indeed be a useful tool when the negation is meaningful.
Dedicated to my grandmother, Daisy Lee Nickleson

May she rest in peace
ACKNOWLEDGMENTS

I would like to thank my advisor, Richard E. Petty, for intellectual support, and for his help in correcting both my stylistic and scientific errors.

I thank Brandon Kopp for his extensive help in the early stages of this research project, and all other members of the Petty Lab for providing helpful feedback on this research.

I am grateful for all of the excellent feedback and suggestions made by the members of the Social Cognition Research Group, Group for Attitudes and Persuasion, and the Attitudes Cavalry.

I would also like to thank my mother, Lisa Johnson, and the rest of my family and friends for teaching me that failure is never an option.
VITA

April 26, 1985..........................Born – Indianapolis, IN

2007.................................B.A. Psychology, Purdue University, Indianapolis

2007-2008............................Graduate Fellow, The Ohio State University

2008-present........................Graduate Teaching Associate, The Ohio State University

FIELDS OF STUDY

Major Field: Psychology
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>ii</td>
</tr>
<tr>
<td>Dedication</td>
<td>iii</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>iv</td>
</tr>
<tr>
<td>Vita</td>
<td>v</td>
</tr>
<tr>
<td>List of Tables</td>
<td>vii</td>
</tr>
<tr>
<td>List of Figures</td>
<td>viii</td>
</tr>
<tr>
<td>Chapters:</td>
<td></td>
</tr>
<tr>
<td>1. Introduction</td>
<td>1</td>
</tr>
<tr>
<td>2. Pilot Study</td>
<td>15</td>
</tr>
<tr>
<td>Method</td>
<td>15</td>
</tr>
<tr>
<td>Results</td>
<td>21</td>
</tr>
<tr>
<td>Discussion</td>
<td>25</td>
</tr>
<tr>
<td>3. Study 1: Meaningful Negation</td>
<td>27</td>
</tr>
<tr>
<td>Method</td>
<td>28</td>
</tr>
<tr>
<td>Results</td>
<td>30</td>
</tr>
<tr>
<td>Discussion</td>
<td>32</td>
</tr>
<tr>
<td>4. Study 2: Moderation of Meaningful Negation</td>
<td>34</td>
</tr>
<tr>
<td>Method</td>
<td>36</td>
</tr>
<tr>
<td>Results</td>
<td>37</td>
</tr>
<tr>
<td>Discussion</td>
<td>43</td>
</tr>
<tr>
<td>5. General Discussion</td>
<td>45</td>
</tr>
<tr>
<td>List of References</td>
<td>55</td>
</tr>
<tr>
<td>Appendix A: Tables, Figures and Stimuli</td>
<td>59</td>
</tr>
</tbody>
</table>
### LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Mean response latencies in milliseconds as a function of prime (White vs. Black), Target (Positive vs. Negative), Training task condition (negate-consistent vs. negate-inconsistent vs. negate consistent – meaningful), and Time of measurement (Before Training vs. After Training), Pilot study</td>
</tr>
<tr>
<td>2</td>
<td>Mean response latencies in milliseconds as a function of prime (White vs. Black), Target (Positive vs. Negative), Training task condition (negate-consistent vs. negate-inconsistent), and Time of measurement (Before Training vs. After Training), Study 1</td>
</tr>
<tr>
<td>3</td>
<td>Mean response latencies in milliseconds as a function of prime (White vs. Black), Target (Positive vs. Negative), Training task condition (negate-consistent vs. negate-inconsistent), and Time of measurement (Before Training vs. After Training), Study 2</td>
</tr>
<tr>
<td>4</td>
<td>Mean response latencies in milliseconds as a function of prime (White vs. Black), Target (Positive vs. Negative), Training task condition (negate-consistent vs. negate-inconsistent), and Time of measurement (Before Training vs. After Training), and MCPR (HI vs. LO), Study 2</td>
</tr>
<tr>
<td>5</td>
<td>Mean response latencies in milliseconds as a function of prime (White vs. Black), Target (Positive vs. Negative), Training task condition (negate-consistent vs. negate-inconsistent), and Time of measurement (Before Training vs. After Training), Study 1 &amp; 2 combined.</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Percent errors made by block on training task collapsed across condition (Pilot study).</td>
<td>66</td>
</tr>
<tr>
<td>2</td>
<td>Reaction time in milliseconds for each block of the training task collapsed across condition (Pilot Study).</td>
<td>67</td>
</tr>
<tr>
<td>3</td>
<td>Mean scores of automatic racial prejudice as a function of training task (negate-consistent vs. negate-inconsistent vs. negate-consistent – meaningful). (Pilot study).</td>
<td>68</td>
</tr>
<tr>
<td>4</td>
<td>Mean scores of automatic racial prejudice as a function of training task (negate-consistent vs. negate-inconsistent). Higher scores indicate a preference for Whites relative to Blacks (Study 1).</td>
<td>69</td>
</tr>
<tr>
<td>5</td>
<td>Mean scores of automatic racial prejudice as a function of training task (negate-consistent vs. negate-inconsistent). Higher scores indicate a preference for Whites relative to Blacks. (Study 2).</td>
<td>70</td>
</tr>
<tr>
<td>6</td>
<td>Automatic racial prejudice as a function of training condition (negate-consistent vs. negate-inconsistent) and time (before vs. after training) for high MCPR (top panel) participants, $F$ (1, 40) = 5.05, $p = .03$, and low MCPR (bottom panel) participants, $F$ (1, 42) = .016, $p = .90$. (Study 2).</td>
<td>71</td>
</tr>
</tbody>
</table>
CHAPTER 1: INTRODUCTION

Negation is a central part of daily communication. Each day we encounter numerous statements that require us to understand negation in order to comprehend language. Interestingly, the mental operations employed to process and store negated information have long been of interest in the fields of philosophy (Bennett, 1984; Descartes, 1984; Spinoza, 1982) and psychology (Gilbert & Osborne, 1989; Loftus & Davies, 1984; Wyer & Unverzagt, 1985). In the field of social psychology, an explanation for just how mental systems believe has most clearly been articulated by Gilbert (1991; Gilbert, Krull, & Malone, 1990; Gilbert, Tafarodi, & Malone, 1993) in his research on the encoding of false beliefs. In articulating how mental systems believe, Gilbert has compared and contrasted two historic views on believing and understanding: the Cartesian procedure and the Spinozoan procedure.

The Cartesian procedure, first introduced by René Descartes in the early 17th century, rests on the simple notion that comprehension precedes and is separate from validity assessment (1644/1984). According to Descartes, comprehension is passive; understanding an idea is effortless and automatic. Assessment of an idea as true or false, on the other hand, was thought to be effortful. Thus, according to Descartes, determining
a proposition’s veracity is independent of comprehending that statement. Along these lines, Descartes’ reasoning suggests that it is possible to represent a proposition mentally before even assessing the veracity of the proposition. For instance, the proposition *the waterfall is purple* can be mentally represented and understood without assessing if the statement is veridical in nature. Fittingly, Gilbert (1991) likens this description to a computer. Computers are capable of representing and storing volumes of information in an unevaluated format. Likewise, Descartes’ argues that ideas can be understood without being assessed. In essence, the Cartesian procedure argues that mental systems operate in a fashion in which comprehension and assessment of mental contents are inherently independent.

In stark contrast to the Cartesian procedure, the Spinozoan procedure (1677/1982) posits that belief systems need not distinguish between beliefs and ideas, because the two are the same. Baruch Spinoza, the originator of the Spinozoan procedure in the late 17th century dismissed Descartes’ distinction between comprehension and assessment. Instead, he argued that comprehending and accepting an assertion as veridical were in fact the same mental operation. Spinoza argued that in order to understand a proposition, it’s necessary for one to implicitly accept that proposition as true and only after acceptance is it possible for subsequent invalidation to occur. Thus, in considering the proposition *the waterfall is purple*, Spinoza asserts that one must first believe the statement in order to comprehend it. In other words, comprehension and assessment
occur simultaneously – in a single step. Furthermore, it is only after the individual considers other propositions that conflict with the statement (i.e. waterfalls are blue), is he or she able to then negate this initial proposition. Essentially, according to Spinoza, to comprehend is to accept as true and only after comprehension and acceptance, a single mental operation, can invalidation, a second mental operation, take place.

Can one really understand the proposition *waterfalls are purple* without believing that water is purple and not blue? Gilbert and colleagues (1993) sought to determine if such was the case by empirically testing which of these two mental systems truly characterized the comprehension and belief of ideas. In a series of experiments, participants were exposed to false information about a criminal defendant. Half of the participants were placed under cognitive load during exposure to the information (Gilbert et al., 1993; Experiments 1 and 2). For the participants under cognitive load (vs. no load), the false information was more likely to be believed and subsequently used in making consequential judgments about the defendant. Based on these findings, Gilbert concluded the Spinozan procedure appeared to be a more accurate portrayal of mental operations. That is, the manipulation of cognitive load prevented participants from correctly negating or unbelieving the false information, suggesting that to believe is implicit to comprehension.

Importantly, the ideas put forth by Spinoza and the work of Gilbert and colleagues are informative regarding the mental operations underlying negation. Spinoza (1982)
articulated a useful metaphor in depicting how negation is mentally represented. He described the mind as a library that utilized a tagged and untagged system to assess its books, analogous to mental contents stored in the mind. In Spinoza’s view, because comprehending and believing are synonymous, books are represented before they are assessed as true or false. Furthermore, as a consequence of the particular tagging system that is used to denote the outcome of the assessment, a new book that appeared without a tag – a believed yet unassessed idea – looks exactly like a belief assessed and determined to be true. Analogously, ideas determined to be veridical in nature through a rational assessment process are represented in the mind in precisely the same way as ideas that have simply been comprehended; only ideas that are judged to be false and unaccepted, are given a special tag. In essence, Spinoza suggests that negation requires us to tag our mental contents as false.

Meta-cognitive Model (MCM) of Attitudes

The idea of “tagging” our mental contents as false is made explicit in the Meta-cognitive Model (MCM; Petty, Briñol & DeMarree, 2007) of attitudes. The MCM builds on the writings of Spinoza and the more recent empirical research by Gilbert (1991) and others (Mayo, Yaavoc, & Burnstein, 2004). In brief, the MCM asserts that people can tag their evaluative associations as true or false, or held with varying degrees of confidence. The MCM contends that people can hold oppositely-valenced evaluations of attitude objects. Specifically, according to the MCM, an attitude object can be associated with
both positive and negative evaluations. The feature of the MCM that gives the model its name is the assumption that people can then tag these evaluative associations as true or false, similar to tagging books in a Spinozan library.

Critically, meta-cognition refers to thinking about one’s thoughts or thought processes (Petty, Briñol, Tormala, & Wegener, 2007). The meta-cognitive associations in the MCM can be represented in various ways such as yes/no, confidence/doubt, true/false, accept/reject, and so forth. This is consistent with Spinoza’s suggestion that we can tag our mental contents with a validity assessment. However, the MCM goes beyond the simple tag-untagged system described by Spinoza and the on-line validation process described by other judgment models such as Kruglanski’s (1989) theory of lay epistemics and Gawronski and Bodenhausen’s (2006) APE model of implicit and explicit attitude change. Specifically, in the MCM, these meta-cognitions, or tags, can be stored for later retrieval just as the evaluations themselves can be stored. And, just as the stored evaluations can vary in their strength of association with the attitude object (Fazio et al., 1995), so too can the validity tags vary in their strength of association with the evaluation (Petty & Briñol, 2009). There are two implications of this. First, the more a person rehearses a validity assessment, the more likely it is that the validity assessment will be retrieved and impact expressed judgment when the evaluation itself is made accessible (i.e., is retrieved). Second, because rehearsing a validity tag necessarily involves
rehearsing the evaluation itself, the stronger the validity tag, the stronger the evaluation itself is likely to be in terms of its accessibility.

The MCM has many potential implications in the domain of stereotyping and prejudice. The automatic operation of stereotypes and prejudice has been shown to be a robust effect in the social psychological literature. For instance, research examining automatic prejudice has revealed that most Whites have automatic negative associations to Blacks, and that most men and women automatically associate males and females with stereotypic attributes. In fact, effects such as these have been documented in over 100 studies (Blair, 2002). The relative ease of revealing these automatic associations, the strength and prevalence of the effects, and the growing evidence that such associations predict and influence behavior have had a profound influence on how researchers view stereotyping and prejudice (Greenwald, Poehlman, Uhlmann, & Banaji, 2009).

One of the most prominent early ideas from this research was the notion that because stereotypes and prejudice are automatic, these biases are inevitable and their influence nearly impossible to avoid (Bargh, 1999; Devine, 1989). That is, there is an assumption that automatic processes are inflexible and impervious to the perceiver’s intentions and goals. Many researchers argue that these automatic biases represent people’s true attitudes that are deep seated, consistent across time and situations, and resistant to external pressures and strategic processes (Banaji & Hardin, 1996; Greenwald et al., 2009).
The notion of the inevitability of prejudice clashes with the implications of the MCM. That is, the MCM holds that if people initially hold a negative evaluation of a particular group, they can also learn, over time, to negate that association meaning that the prejudice can be corrected automatically (see also Wegener & Petty, 1997; Maddux, Barden, Brewer, & Petty, 2005). On the other hand, if such negation is not possible, people would not be able to correct for their prejudice by attempting to deny it. They could only potentially overcome prejudice by learning to associate positive attributes with the stigmatized group.

Social psychological research has demonstrated that although nearly everyone knows the content of stereotypes and shows some degree of automatic prejudice, not all individuals endorse these beliefs or use them in judgments of others (Devine, 1989). This could be because some individuals “tag” these beliefs or attitudes as invalid. If people can tag their evaluative associations as true or false, or being held with varying degrees of confidence, they should presumably be able to tag their negative evaluative associations (prejudice) as false. Indeed, recent research in social psychology has argued that this might be the case – that people can say no to their prejudice. Specifically, in recent research by Kawakami and colleagues (2000), participants were trained to negate stereotype-consistent information. For instance, participants would see a Black face and a trait word related to the stereotype of Blacks such as lazy or streetwise, and would be instructed to hit the space bar to indicate “NO”! Participants practiced this negation
training extensively; for a total of 160 trials. In addition, both prior and immediately following training, all participants completed a person categorization task (Blair & Banaji, 1996) which assessed the speed with which participants classified Black and White faces following the subliminal presentation of positive and negative trait words. Kawakami and colleagues found that the automatic activation of prejudice was reduced following extensive training on the negation of stereotype-consistent information (see Kawakami, Dovidio, Moll, Hersmen, & Russin, 2000).

Using the MCM as a springboard to examine the work of Kawakami and colleagues, one can view the training as a mechanism to increase the link between the “NO” or negation tag, and the associated evaluation. Furthermore, consistent with the predictions of the MCM, as the strength in the link between the “NO” tag and the evaluative association increases, the likelihood that the no tag will be retrieved along with the evaluation also increases. This could explain the observed reduction in automatic activation of prejudice following the training. In essence, Kawakami and colleagues’ work might be viewed as providing some initial evidence that the application of the MCM to the domain of stereotyping and prejudice could prove beneficial in reducing automatic prejudice.

The work of Kawakami and colleagues, however, is not without its limitations. As noted previously, Kawakami and colleagues (2000) demonstrated that extensive training in the negation of stereotypic associations facilitated a reduction in the automatic
activation of prejudice. However, the study utilized a within-subjects design - participants actually engaged in the negation of stereotype-consistent information and the affirmation of stereotype-inconsistent information in one procedure, thus making it unclear which was driving the observed reduction in prejudice. Gawronski and colleagues (see Gawronski, Deutsch, Mbirkou, Seibt, & Strack, 2008) realized this shortcoming, and extended the work of Kawakami et al.’s (2000).

To determine which of the two forms of training was driving the reduction in prejudice, Gawronski and colleagues adopted a between-subjects design, in which participants either engaged in affirmation or negation training. Participants were trained to either negate prejudice-consistent information (i.e. saying “NO” whenever Black was paired with a negative trait) or affirm prejudice-inconsistent information (e.g. saying “YES” whenever Black was paired with a positive trait). Based on their findings, Gawronski and colleagues (2008) posit that it is only the affirmation of prejudice-inconsistent information that facilitates the reduction of automatic prejudice. Furthermore, they add that continually practicing negation would not necessarily stop negative traits from coming to mind and that negation can, in fact, strengthen associations and make the link between the category and trait stronger. This is because in Gawronski

1 Contrary to participants’ negating or affirming positive and negative traits associated with the group stereotype of Blacks and Whites as they did in Kawakami et al., in the Gawronski et al. research, participants were only presented with negative traits stereotypically associated with Blacks and positive traits stereotypically associated with Whites.
et al.’s conceptual framework, negation is not stored so all that happens is that
the negative evaluative association increases in strength. However, the MCM suggests that
continual practice in negating unfavorable traits can lead to a stored negation tag and thus
*should* at some point lead the person to immediately think “no” when such traits are
activated. Essentially, Gawronski and colleagues (2008) are arguing that people do not
take into account whether a “YES” or a “NO” is associated with the activated trait. If
invalidity tags are not stored, then indeed either validation or invalidation would only
strengthen the initial association. According to the MCM, however, although it may be
the case that affirming prejudice-inconsistent information is more effective than negating
prejudice-inconsistent information, negation should not be a meaningless activity
altogether.

Will people who have practiced negating unfavorable traits be faster to reject
them (i.e., think “no” after *hostile* comes to mind in the presence of a Black person)
thereby correcting themselves (and perhaps others)? Is there any value in using negation,
or does it, ironically, do more harm than good as Gawronski and colleagues (2008) posit?
The present research sought to answer these questions.

*Overview of the Present Research*

The goal of the present research was to examine if negation training could be used
to reduce individuals’ automatic racial prejudice. Previous research by Kawakami and
colleagues (2000) and the implications of the MCM suggest that negation training could
be a useful tool to reduce racial prejudice. However, the work of Gawronski et al. raised a serious challenge to this point of view. The Gawronski et al. research indicates that negation will only make unfavorable associations stronger, resulting in an ironic rebound effect (2008; Wegner, 1994). According to Gawronski, how an individual assesses the information, (i.e. affirming or negating) is not important. Instead, the simple pairing of a trait with a category member, such as Black with poor, will increase the strength of the association thereby enhancing automatic negative evaluations.

Based on Gawronski’s reasoning, one would predict that *negating* consistent information is the equivalent of *affirming* consistent information and that *affirming* inconsistent information is the equivalent of *negating* inconsistent information. Thus, one would expect that training an individual to negate unfavorable traits, as in Gawronski et al. (2008), should result in an increase in automatic racial prejudice, whereas training an individual to negate favorable traits should result in a decrease in prejudice, given that the training is analogous to affirming favorable traits. However, based on the implications of the MCM, we predict the exact opposite.

In the present research, we sought to extend the work of Gawronski and colleagues by pitting the negation of prejudice-consistent information against the negation of prejudice-inconsistent information. If, as Gawronski argues, it is simply about pairing the trait with the category member, then one would expect to replicate his findings for those individuals who negate prejudice-consistent information. If, however,
the continual practice in negation is resulting in these evaluative associations being tagged as false as we believe, then we predict that individuals who negate prejudice-consistent information will show a decrease in automatic racial prejudice. Conversely, for those individuals who negate prejudice-inconsistent information (i.e. negating Black paired with rich), we predict that these individuals will demonstrate an increase in automatic racial prejudice following negation training.

In addition to delineating the value of negation training, the present research also hoped to address a reason why the Gawronski et al. research failed to uncover any effective negation. One reason, perhaps, is that the negation used was not strong enough, or not meaningful enough. The MCM (Petty et al., 2007) posits that as the strength of the link between a negation tag and an associated evaluation becomes stronger, the tag becomes more likely to be retrieved along with the association itself. There are at least two ways to increase the strength of a negation tag. One method would be to increase the number of negation trials. Alternatively, strength of negation could be enhanced by making the negation more meaningful. That is, it is possible that some forms of negation could be more potent in increasing the strength of the link between the negation tag and the associated evaluation. For example, a simple “no” is plausibly less meaningful and impactful than “That’s wrong!” Assuming that a more meaningful negation, such as “That’s wrong!” facilitates an increase in the strength of the link between the negation tag and the association more so than a simple “no,” we would expect that meaningful
negation would be more impactful. To explore this possibility, in addition to using a simple negate-consistent and a negate-inconsistent condition in a pilot study, the research also included a negate-consistent – meaningful condition to empirically test if meaningful negation would show an enhanced impact above and beyond simple negation.

In the pilot study, participants were randomly assigned to either a simple negate-consistent, a simple negate-inconsistent or a negate-consistent–meaningful condition, and their automatic attitudes towards Blacks and Whites were assessed before and after training using an evaluative priming task (Fazio et al, 1995). If a simple “no” is sufficient to produce a negation effect, then individuals who negated consistent information should show a decrease in automatic racial prejudice, whereas those who negated inconsistent information should show an increase in prejudice. However, if a simple no is insufficient to produce negation, then either no effects would be observed or it is possible that the effects obtained by Gawronski et al. would be obtained. If a simple “no” is insufficient to produce negation, then perhaps the meaningful negation condition would prove effective.

Because the pilot study revealed an effect for meaningful negation but not simple negation, in the next two studies of this thesis only meaningful negation conditions were used. In Study 1 using only meaningful negation, two conditions were examined -- negate prejudice-consistent and negate prejudice-inconsistent traits. We predicted that individuals who negate prejudice-consistent information in a meaningful way would
show a decrease in automatic racial prejudice, whereas those who negate prejudice-inconsistent information in a meaningful way would show an increase in automatic racial prejudice.

In Study 2, we wanted to replicate the results of Study 1 and in addition examine the role of individuals’ Motivation to Control for Prejudice Responding (MCPR; Dunton & Fazio, 1997) on the effectiveness of the negation training task. Together, these three studies allow us to better understand the value of meaningful negation training as a tool to change automatic racial attitudes. The studies also provide data relevant to the question of whether or not negation is simply an on-line process as some models imply (e.g., Gawronski & Bodenhausen, 2006) or whether negations can be stored for later retrieval as other models suggest (e.g., Petty et al., 2007).
CHAPTER 2: PILOT STUDY

Method

Participants and Design

Participants were 115 undergraduates at the Ohio State University, including 62 females (53.9%) and 53 males (46.1%). The average age of participants was 19 years old ($M = 19.28$ and $SD = 2.42$). Participants were recruited through the Psychology department’s subject pool. Participants received partial course credit to take part in a study examining “learning and memory.”

The design of the pilot study was adapted from Gawronski and colleagues (2008), and employed a 2 (time of measurement: before vs. after training task) x 3 (training task: negate prejudice-consistent vs. negate prejudice-inconsistent vs. negate prejudice-consistent–meaningful) mixed-model design with the first variable as a within-subjects factor and the second a between-subjects factor.

Procedure

Upon arrival, participants were greeted by the experimenter and instructed to have a seat at a personal computer station of their choice. All instructions for subsequent procedures were presented to participants on their computer screens.

All participants viewed the following overview on their computer screens shortly after being seated:
“In today’s experiment we are going to ask you to perform several tasks related to the black/white cultural stereotype. Many people know what the basic elements of the stereotype are, and just because one knows about the stereotype, that does not mean they endorse its contents (for example, I know there is a stereotype that African-Americans are better than whites when it comes to athletics but I believe there are no racial differences when it comes to athletic skill). Today’s experiment is concerned with the content of the stereotype, not your endorsement of it.”

Participants were subsequently instructed that today’s experiment would consist of two tasks. Task One, the evaluative priming measure, was described as a categorization task that would take place in two parts. Participants were instructed they would complete the first block of Task One now, and the remaining block after completing Task Two. Task Two, the training task, was described as a task that required participants to either negate stereotype-consistent or stereotype-inconsistent information. Participants were informed they would receive more specific details about this task after completing the first block of Task One.

Thus, the basic procedure consisted of participants completing the evaluative priming measure twice -- prior to and immediately after the training task. Once participants completed both halves of Task One (i.e. the priming measure) and Task Two (i.e. the training task), they were thanked and thoroughly debriefed.
**Training Task**

Instructions presented to participants differed depending on their assigned condition. Participants assigned to the *negate prejudice-consistent* condition were instructed to respond *NO* whenever they saw a face-trait combination that was consistent with expressing unfavorable prejudice toward Blacks or favorable prejudice toward Whites (e.g., a Black face paired with “lazy” or a White face paired with “friendly”). Participants were told that they could indicate a NO response by pressing the space bar on their computer keyboard. The specific instructions were:

“The following task is concerned with the cultural stereotype of blacks and whites. As you probably know, blacks are often portrayed negatively whereas whites are often portrayed positively. This, however, is a cultural stereotype that may or may not be true. In the following task, you will be presented with pictures of blacks and whites. In addition, you will be presented with stereotypically negative traits of blacks and stereotypically positive traits of whites. Your task is to respond “NO!” each time you see a combination that is CONSISTENT with the cultural stereotype of blacks and whites. Specifically, you are asked to respond “NO!” with the space bar each time you see a picture of a BLACK man and a NEGATIVE trait or a picture of a WHITE man and a POSITIVE trait. Please attend particularly to combinations that are CONSISTENT with the
cultural stereotype of blacks and whites! For combinations that are inconsistent
with the cultural stereotype of blacks and whites, you do not have to do anything.
Again, please respond “NO!” with the space bar each time you see a
combination that is CONSISTENT with the cultural stereotype of blacks and
whites. Please try to respond as quickly and accurately as possible!”

For participants who were in the negate prejudice-inconsistent training condition,
you were told to respond NO whenever they saw a face-trait combination that was
prejudice-inconsistent (e.g., a Black face paired with “friendly” or a White face paired
with “lazy”), and do nothing for combinations that were prejudice-consistent.

Participants in a third condition, negate prejudice-consistent–meaningful,
received identical instructions to the negate prejudice-consistent training condition, with
one exception. Instead of being instructed to simply respond NO to prejudice-consistent
face-trait combinations, these participants were asked to respond “That’s wrong”! As
noted earlier, this third condition was added to see if making negation more meaningful
would reduce automatic evaluations of prejudice above and beyond just responding NO
to prejudice-consistent face-trait combinations.

Participants in all three conditions were then presented with a total of 200 face-
trait pairings. These pairings included 50 combinations of each of: (a) a Black face paired
with a positive trait word, (b) a Black face paired with a negative trait word, (c) a White
face paired with a positive trait word, and (d) a White face paired with a negative trait
word. Pairings of Black versus White faces and positive versus negative trait words was individually randomized by the computer for each participant. A total of 10 Black male faces and 10 White male faces were used as stimuli. Photos of both Blacks and Whites pictured an expressionless target from the neck up, on a dark blue background (see Appendix A).

For each trial, participants were first presented with a Black or a White face that appeared at the top of the screen. After 500 ms, a positive or negative trait word appeared at the bottom of the screen. Trait words were either related to a positive evaluation of White people or a negative evaluation of Black people (see Appendix A for full list of trait words). When participants correctly pressed the spacebar, the stimuli disappeared and the next trial started. If participants incorrectly pressed the spacebar in response, the stimuli were replaced by the message “ERROR!” which appeared in the middle of the screen for 1500 milliseconds. If participants did not respond to a given combination, the stimuli disappeared after 2500 ms and the next trial started. The time between trials for all responses was 1000 milliseconds. All training tasks consisted of five blocks of 40 trials each, resulting in a total of 200 training trials. After participants completed each block, they were asked to take a moment to relax before proceeding and to hit the space bar when they were ready to continue the task.

*Automatic Evaluation*
To assess automatic evaluations of Blacks and Whites, we utilized a subliminal evaluative priming task (see Fazio, Jackson, Dunton & Williams, 1995) as it was adapted by Gawronski and colleagues (2008). In this task, each trial began with a fixation cross (“+”) which was presented for 1 second in the center of the screen. Immediately afterwards, the prime word “black” or “white” was presented for 15 ms and followed by a masking stimulus (“XXXXX”) for 250 milliseconds. The masking stimulus was then replaced by a positive or negative target word in the center of the screen which remained until participants responded.

Participants were instructed to press the left-hand key (“A”) as quickly as possible when they saw a positive word and the right-hand key (“L”) when they saw a negative word. Each of the 40 target words was presented twice with each of the two prime words, resulting in a total of 160 trials. In order to assure independence of automatic prejudice and automatic stereotyping at the measurement level, positive and negative nouns were used as target words (e.g., candy, dirt) rather than the positive and negative trait words (e.g., lazy, friendly) used in the training task. Order of trials was randomized individually for each participant. Incorrect responses were indicated with the word “ERROR!” appearing in red ink for 1 second in the middle of the screen. The inter-trial interval for incorrect and correct responses was 1 second. Participants completed the evaluative priming task twice; once immediately before training and once immediately after training.
Consistent with Gawronski et al. (2008), the priming measure was scored by creating an index that reflected the overall automatic preference for Whites over Blacks. Raw latencies were used to create the index. The overall index were calculated by first subtracting the mean response latency to positive words after White primes from the mean response latency to positive words after Black primes (i.e. higher scores indicate stronger activation of positivity for White as compared to Black), and by subtracting the mean response latency to negative words after Black primes from the mean response latency to negative words after White primes (i.e. higher scores indicate stronger activation of negativity for Black as compared to White). Negativity scores were then added to positivity scores, resulting in an index of automatic preference for Whites over Blacks with higher scores indicating a stronger preference for Whites over Blacks.

Results

Training Task

To ensure that participants were successfully completing the training task, both participants’ errors and reaction times across all five training blocks were examined. An error was defined as either negating a combination on a trial in which the participant was to give no response, or by failing to hit the space bar for combinations which were to be negated. For example, for those participants’ assigned to the negate-consistent or the negate-consistent – meaningful condition, an error would be classified as either hitting
the space bar for combinations that were inconsistent with the cultural automatic evaluation of Blacks and Whites, or failing to hit the space bar for combinations that were consistent with cultural attitudes. Given that each block consisted of 40 trials in which participants were required to hit the space bar in response, participants could potentially make anywhere from zero to 40 errors per block. For each trial block, participants randomly received 20 face trait-combinations that were to be negated by hitting the space bar and 20 face-trait combinations that required no response. Consequently, errors for the consistent block were summed and divided by the total number (i.e., 20) and subsequently converted to a percent. The same was done for inconsistent trials for each block. These two percentages were later averaged to create the total percentage of errors per training block.

The mean percentage of errors made for each block was then submitted to a 3 (training task) x 5 (training block) mixed-model ANOVA, which revealed a significant main effect of block, $F (4, 444) = 29.20, p < .0001$, with participants making fewer errors each subsequent block, from block 1 to block 5. Errors made across all five learning blocks, collapsed across condition, are illustrated in Figure 1. As depicted in the figure, the most errors were made in the first block and they tapered off dramatically after that. Indeed, there were no significant differences in errors in Blocks 2 – 5. In addition, the number of errors did not differ by experimental condition, $F (8, 444) = .86, p = .86$. 

22
suggesting that across condition all subjects learned the task equally well. Finally, there was no interaction between training task and training block.

Participants’ reaction times across blocks showed a similar pattern. A 3 (training task) x 5 (training block) mixed-model ANOVA assessing reaction times for each block revealed a significant main effect of block, $F(4, 444) = 39.34, p < .0001$, with participants’ reaction times decreasing with each subsequent block. No differences in reaction times by experimental condition were found, $F(8, 444) = .462, p = .88$, nor was an interaction obtained. Thus, reaction times across blocks were collapsed across condition. As depicted in Figure 2, responses were slowest in the first block and became much faster after that. Indeed, there were no significant differences in response times in Blocks 2 – 5. In essence, participants made fewer errors and responded faster in each subsequent block indicating that they were able to successfully complete the training task.

*Evaluative Priming*

Across participants, relatively few errors were made in judging adjectives as positive or negative (4.4% at time 1; 5.0% at time 2), and these trials were excluded from further analyses. The data also exhibited few outliers, defined as response times less than 300 ms or greater than 1,000 ms. Response times less than 300 ms indicate that participants responded more quickly than is considered physically possible (indicated a blind guess), so these trials were dropped from the analyses. Response times greater than
1000 ms usually indicate temporary attention lapses. These responses were discarded so that they would not overly influence the results. Consequently, prior to analyses, outliers were excluded by discarding responses lower than 300 ms (0.12% at time 1; 0.6% at time 2) or higher than 1000 ms (7.8% at time 1; 11.76% at time 2). These procedures are common in analyzing data from evaluative priming tasks (Fazio et al., 1995; Gawronski et al., 2008; Neely, 1977). A 3 (training task) x 2 (time of measurement) mixed-model ANOVA with errors made on block 1 of the training task (err1) as a covariate revealed a significant 2-way interaction of time and condition, $F(1, 110) = 4.36, p < .01$, as pictured in Figure 3.\(^2\) To specify this interaction in terms of the present hypotheses, differences in racial prejudice from time 1 to time 2 were assessed by condition. As expected due to random assignment, participants did not differ by condition in their automatic racial evaluation at time 1, $F(2, 114) = 2.47, p = .08$. Likewise, participants also did not differ by condition at time 2, $F(2, 114) = .817, p = .445$.

However, a paired-samples t-test looking within condition from time 1 to time 2 demonstrated a significant decrease in automatic racial prejudice for the negate-consistent–meaningful training condition, $t(35) = 3.60, p < .001$. In contrast, paired-samples t-tests for the negate-consistent, $t(38) = -.15, p = .878$, and the negate-

\(^2\) Errors made on block 1 (i.e. err1) were chosen as a covariate because participants’ error rate was highest for this block of the training task (see Figure 1). However, using errors on any of the subsequent blocks or all errors combined produces the same results as the presented findings.
inconsistent condition, \( t (39) = -0.56, p = .57 \), revealed no significant differences in automatic racial prejudice from time 1 to 2. ³

**Discussion**

The present findings are interesting in two regards. First, contrary to work by Gawronski and colleagues (2008), participants in the negate-consistent training showed no change in levels of automatic racial prejudice from time 1 to 2. Furthermore, those participants who negated prejudice-inconsistent face-trait combinations also showed no change in automatic racial prejudice.

Second, and of greater interest, the negate-consistent–meaningful condition showed significant learning effects. Relative to participants who only responded “NO” to prejudice-consistent face-trait combinations, those participants who responded “That’s wrong” did show a significant reduction in automatic racial prejudice from time 1 to time 2. These findings suggest that negation – when meaningful – can be a potential tool to reduce automatic prejudice.

³ Reaction times were also submitted to 2 (time: before vs. after) x 2 (valence: positive vs. negative) x 2 (race: black vs. white) x 3 (Condition: Negate-consistent vs. Negate-inconsistent vs. Negate-consistent – meaningful), with the first three factors as within-subjects variables, and the last a between-subjects variable, mixed-model ANOVA. A significant 4-way interaction of Time x Valence x Race x Condition \( F (2, 112) = 3.61, p = .030 \), emerged. Likewise, similar analyses conducted for Studies 1 and 2 also revealed a significant 4-way interaction, \( F (1, 53) = 8.34, p = .006 \), and \( F (1, 86) = 3.94, p = .05 \) respectively. The 4-way interaction was the only significant effect to consistently emerge across all three studies. See Tables 1-5 for specific scores for each study.
Although the findings of the present study are promising, the pilot study is not without its limitations. Specifically, although participants’ level of automatic racial prejudice did not differ at time 1 according to conventional levels of significance ($F (2, 114) = 2.47, p = .08$), the value was trending in that direction. Thus, although the negate-consistent–meaningful training condition significantly differed in levels of automatic racial prejudice from pre to post-training, it is unclear based on the present findings if this condition demonstrates a clear advantage relative to the negate-consistent condition or whether there was a failure of random assignment. In order to address these concerns, Study 1 was conducted.

The key aim of this study was to replicate the finding that meaningful negation can result in reduced prejudice. Furthermore, this study included a condition in which meaningful negation of prejudice-inconsistent pairings was used. Because the pilot study only examined meaningful negation of prejudice-consistent information, it is possible that simply indicating “That’s wrong” in the context of a study on race would lead to reduced prejudice. However, if the negation is being applied to the specific face-trait pairs presented, then meaningful negation of prejudice-inconsistent information should actually increase levels of automatic prejudice.
Study 1 was motivated by the findings of the pilot study. Specifically, we sought to examine the impact of meaningful negation training on levels of automatic racial prejudice. The present research again examines negation-consistent and negation-inconsistent training; however, in the present research the negation is always meaningful. Specifically, the negate instructions for the negate-consistent-meaningful condition in the previous study are used for both training conditions since no effects were observed when the negation was minimal (i.e., “no” instead of “That’s wrong”).

Using meaningful negation as a tool, we predict that when participants engage in the meaningful negation of prejudice-consistent information, they will show a decrease in levels of automatic prejudice, replicating the pilot study. Conversely, for those participants who engage in the meaningful negation of prejudice-inconsistent information, they will show an increase in levels of automatic prejudice. To the extent that this occurs, it supports the notion that negation is specific to the information negated and it is not the case that simply responding “That’s wrong” in the context of a study on race is sufficient to reduce automatic prejudice. The present research sought to test these ideas.
Method

Participants and Design

Participants were fifty-five Ohio State University undergraduates, including 31 females (56.4%) and 24 males (43.6%). The average age of participants was 19 years old ($M =19.43$ and $SD =1.24$). Participants were recruited through the Psychology Department’s subject pool and all participants received partial course credit for their involvement in the study.

The experiment consisted of a 2 (time of measurement: before vs. after training) x 2 (training task: negate-consistent vs. negate-inconsistent) mixed-model design with the first variable representing a within-subjects factor and the second a between-subjects factor. The basic procedure was the same as that for the pilot study. That is, all participants responded to the evaluative priming measure on race prior to the negation training task and then again after it.

Training Task

Similar to the pilot study, all participants were greeted by the experimenter upon arrival and seated at a computer station. Instructions were presented on the screen for all participants. Instructions for the training task were similar to the pilot study. However, in both conditions participants were instructed to negate information by thinking “NO! THAT'S WRONG”! In the negate-consistent training condition, participants received the following instructions:
“The following task is concerned with testing your knowledge of cultural stereotypes. Today, you will be presented with pictures of blacks and whites. In addition, you will be presented with stereotypically negative traits of blacks and stereotypically positive traits of whites. Your task is to respond “NO! THAT’S WRONG!” each time you see a combination that is CONSISTENT with the cultural stereotype of blacks and whites. Specifically, you are asked to respond “NO! THAT’S WRONG!” with the space bar each time you see a picture of a BLACK man and a NEGATIVE trait or a picture of a WHITE man and a POSITIVE trait. Please attend particularly to combinations that are CONSISTENT with the cultural stereotype of blacks and whites! For combinations that are inconsistent with the cultural stereotype of blacks and whites, you do not have to do anything. Again, please respond “NO! THAT’S WRONG!” with the space bar each time you see a combination that is CONSISTENT with the cultural stereotype of blacks and whites. Please try to respond as quickly and accurately as possible!”

Instructions for the negate-inconsistent training were nearly identical. However, in the negate-inconsistent training condition participants were instructed to respond “NO! THAT’S WRONG!” to prejudice-inconsistent combinations and to not respond to prejudice-consistent combinations. The general procedure and number of trials were identical to the pilot study.
**Automatic Evaluation**

Identical to the pilot study, to assess automatic evaluations of Blacks and Whites, we used a subliminal evaluative priming task (see Fazio, Jackson, Dunton & Williams, 1995) that was adapted from Gawronski et al. (2008). The general procedure and number of trials were identical to the pilot study. Again, participants completed the priming task both immediately before and after training.

**Results**

*Training Task*

As in the pilot study, to ensure that participants were successfully completing the training task, both participants’ errors and reaction times across all five training blocks were examined. Errors for the training task were tallied and percent errors per block were calculated identical to that of the pilot study.

Consistent with the pilot study, a 2 (training task) x 5 (training block) mixed-model ANOVA revealed a significant main effect of block, $F(4, 212) = 19.63, p < .0001$, with participants making fewer errors each subsequent block, from block 1 to block 5. Also consistent with the pilot study, there were no significant differences in errors in Blocks 2 – 5. In addition, the number of errors did not differ by experimental condition, $F(4, 212) = .49, p = .73$, suggesting that across condition all subjects learned the task equally well. Nor was there an interaction between training task and training block.
As in the pilot study, participants’ reaction times across block were subjected to a 2 (training task) x 5 (training block) mixed-model ANOVA assessing reaction times for each block. This analysis revealed a significant main effect of block, $F(4, 208) = 27.53, p < .0001$, with participants’ reaction times decreasing with each subsequent block. Again, also consistent with the pilot study, no differences in reaction times by experimental condition were found, $F(4, 208) = 1.132, p = .34$, nor was there a training task by training block interaction. In essence, participants made fewer errors and responded faster in each subsequent block indicating that they were able to successfully complete the training task.

*Evaluative Priming*

Across participants, relatively few errors were made in judging adjectives as positive or negative (4.6% at time 1; 5.1% at time 2), and these trials were excluded from further analyses. Similar to the pilot study, prior to analyses, outliers were excluded by discarding response times lower than 300 ms (0.18% at time 1; 0.5% at time 2) and those higher than 1000 ms (6.8% at time 1; 12.8% at time 2).

A 2 (training task) x 2 (time of measurement) mixed-model ANOVA with errors made on block 1 of the training task (err1) as a covariate revealed a significant 2-way interaction of time and condition, $F(1, 52) = 6.94, p < .01$, as pictured in Figure 4. As expected, participants did not differ by condition in their automatic racial evaluation at
time 1, $F(1, 54) = .123, p = .727$. In contrast, participants differed significantly in
automatic racial evaluation at time 2, $F(1, 54) = 11.35, p < .001$.

To specify this interaction in terms of the present hypotheses, differences in racial
prejudice from time 1 to time 2 were assessed by condition. As in the pilot study, paired
sample t-tests were used for this purpose. Focusing only on the negate-consistent
condition, paired samples t-tests revealed a non-significant decrease in automatic racial
prejudice from time 1 to time 2, $t(29) = 1.392, p = .17$, though the observed reduction
was in the predicted direction. Conversely, when examining the negate-inconsistent
condition, a paired samples t-test revealed a significant increase in automatic racial
prejudice from time 1 to time 2, $t(24) = -2.278, p = .032$, also in the predicted direction.

**Discussion**

Consistent with our predictions, this study showed that what people were negating
was important. That is, those participants who negated prejudice-consistent information
tended to decrease their levels of automatic racial prejudice. Conversely and also
consistent with our predictions, those participants who negated prejudice-inconsistent
information significantly increased in their levels of automatic racial prejudice from time
1 to time 2.

Similar to the negate-consistent (meaningful) condition in the pilot study, the
negation of prejudice-consistent information tended to produce learning effects.
Furthermore, the addition of the negate-inconsistent condition also resulted in learning
effects but this was manifested as a significant increase in racial prejudice. This finding suggests that simply saying “That’s wrong!” in the context of a study about racial attitudes is not sufficient to lead to a reduction in automatic racial prejudice. Given that participants who negated prejudice consistent information tended to decrease their automatic racial prejudice whereas those who negated prejudice-inconsistent information increased their automatic racial prejudice, these findings suggest that the negation is being applied specifically to the face-trait pair combinations.
CHAPTER 4: STUDY 2, MODERATION OF MEANINGFUL NEGATION

The results of Study 1 were promising, and suggest that meaningful negation may indeed be a useful tool to reduce automatic racial prejudice. However, because of the new addition of the negate-inconsistent condition and the marginal reduction in prejudice found in the negate-consistent condition, we conducted a final study to replicate these results.

Additionally, Study 2 was conducted to identify if participants’ individual differences in their motivation to control prejudicial responses might contribute to changes in automatic racial prejudice following the negation training task. Specifically, we assessed participants’ individual differences on the Motivation to Control Prejudiced Reactions Scale (MCPRS; Dunton & Fazio, 1997). We examined whether individuals who scored high on this scale (i.e. highly motivated to control for prejudicial responses; high MCPR), were driving the observed changes in levels of automatic racial prejudice relative to those participants low in their motivation to control for prejudice. One reason those high in MCPR might be driving these effects is because individuals high in MCPR are most likely to monitor their prejudice and might therefore be most sensitive to prejudice training. However, it is also possible to predict the opposite result; that the meaningful negation training may only be effective for individuals low in MCPR, and those who are high in MCPR may be resistant to the training task either because these
individuals follow their own internal guides or because, at a minimum, the training requires those in the negate-inconsistent condition to negate prejudice-inconsistent information.

Considering both of these possibilities, it could be that those high in MCPR would be most affected by the negate-consistent training since that is compatible with their wishes but least affected by the negate-inconsistent training because that is inconsistent with their desires (see also, Plant & Devine, 1998). Prior research has shown that those high in MCPR are especially likely to show reduced racial bias in contexts in which the possibility of race bias is salient. For instance, in one study Maddux, Barden, Brewer and Petty (2005) examined the interaction of context and MCPR on automatic evaluative responses towards Blacks and Whites. Interestingly, they found that Whites high in MCPR demonstrated an automatic outgroup bias in favor of Blacks over Whites in contexts in which the possibility of racial bias was salient (e.g., a photo of Blacks versus Whites in prison). Furthermore, this result was driven by an automatic inhibition of negative responses towards Blacks as if those high in MCPR were automatically correcting for their prejudice because of prior practice in doing so (Wegener & Petty, 1997; see also, Livingston, 2008). In sum, a clear prediction is that those high in MCPR
will be more influenced by the negate-consistent training than those low in MCPR.\(^4\) It is less clear, however, how they will respond to the negate-inconsistent training.

**Method**

*Participants and Design*

Participants were eighty-eight Ohio State University undergraduates, including 50 females (56.8\%) and 38 males (43.2\%). The average age of participants was 19 years old ($M = 19.07$ and $SD = 1.56$). Participants were recruited through the Psychology Department’s subject pool and all participants received partial course credit for their involvement in the study.

The experiment consisted of a $2 \times 2$ (time of measurement: before vs. after training) $\times$ (training task: negate-consistent vs. negate-inconsistent) mixed-model design with the first variable representing a within-subjects factor and the second a between-subjects factor. The basic procedure was identical to that of Study 1. That is, all participants responded to the evaluative priming measure prior to the negation training task and then again after it. In addition to the two experimental manipulations, in Study 2, all participants completed the Motivation to Control Prejudiced Reactions Scale (MCPRS;  

---

\(^4\) This effect should hold unless those high in MCPR are already so practiced in correcting for prejudice that they cannot be trained further. We expected this not to be the case because Maddux et al. (2005) found that the automatic correction effect only occurred in specific situations in which the possibility of racial bias was salient (e.g., prison context), but not in a control context as used in the current research. That is, from the Maddux et al. study, it does not appear that those high in MCPR are overly practiced in controlling their automatic prejudice in neutral contexts.
Dunton & Fazio, 1997) shortly after completing the priming task for the second time. Demographic information was collected and participants were subsequently debriefed, thanked, and dismissed.

Training Task

All instructions and procedures for the training task were identical to that of Study 1. Participants either negated prejudice-consistent or prejudice-inconsistent face-trait combinations, depending on their randomly assigned condition.

Automatic Evaluation

Identical to the previous two studies, a subliminal evaluative priming task was used to assess automatic evaluations of Blacks and Whites (see Fazio et al., 1995) that was adapted from Gawronski et al. (2008). The general procedure, number of trials, and scoring were also identical to the two previous studies. Again, participants completed the priming task both immediately before and after training.

Results

Training Task

Consistent with the pilot study and Study 1, to ensure that participants were successfully completing the training task, both participants’ errors and reaction times across all five training blocks were examined. Errors for the training task were calculated identical to the two previous studies.
A 2 (training task) x 5 (training block) mixed-model ANOVA revealed a significant main effect of block, \( F(4, 344) = 14.83, p < .0001 \), with participants making fewer errors each subsequent block, from block 1 to block 5. Again, there were no significant differences in errors in Blocks 2 – 5. In addition, the number of errors did not differ by experimental condition, \( F(4, 344) = .77, p = .54 \), nor was there a two-way interaction, suggesting that across conditions all subjects learned the task equally well.

Participants’ reaction times across block were also subjected to a 2 (training task) x 5 (training block) mixed-model ANOVA assessing reaction times for each block. This analysis revealed a significant main effect of block, \( F(4, 336) = 23.72, p < .0001 \), with participants’ reaction times decreasing with each subsequent block. Again, also consistent with the previous studies, no differences in reaction times by experimental condition were found, \( F(4, 336) = 1.887, p = .11 \), nor was there a two-way interaction. Similar to the previous studies, participants made fewer errors and responded faster in each subsequent block indicating that they were able to successfully complete the training task.

*Evaluative Priming*

Across participants, relatively few errors were made in judging adjectives as positive or negative (4.0% at time 1; 5.6% at time 2), and these trials were excluded from further analyses. Prior to analyses, outliers were excluded by discarding response times
lower than 300 ms (0.16% at time 1; 0.9% at time 2) or higher than 1000 ms (8.2% at time 1; 12.34% at time 2).

A 2 (training task) x 2 (time of measurement) mixed-model ANOVA with errors made on block 1 of the training task (err1) as a covariate revealed a significant 2-way interaction of time and condition, $F (1, 85) = 3.93, p < .05$, as pictured in Figure 5. As expected, participants did not differ by condition in their automatic racial evaluation at time 1, $F (1, 87) = 1.11, p = .29$. Interestingly, despite the overall interaction reaching conventional levels of significance, participants also did not significantly differ in automatic racial prejudice at time 2, $F (1, 87) = 2.207, p = .14$, though the trend was in that direction.

To specify this interaction in terms of the present hypotheses, differences in racial prejudice from time 1 to time 2 were assessed by condition. Paired samples t-tests were again used. Focusing only on the negate-consistent condition, a paired sample t-test revealed a marginally significant decrease in automatic racial prejudice from time 1 to time 2, $t (46) = 1.78, p = .08$. A paired-samples t-test for the negate-inconsistent condition, $t (40) = -1.06, p = .29$ revealed no significant differences in automatic racial prejudice from time 1 to 2.

Interestingly, Study 1 produced a stronger increase in prejudice in the negate-inconsistent condition than it did a decrease in prejudice in the negate-consistent condition. Study 2 had an opposite pattern. To obtain greater reliability, the data from
Studies 1 and 2 were combined in a 2 (time) x 2 (condition) x 2 (study) analysis of variance. As expected, participants did not differ by condition in their automatic racial evaluation at time 1, $F(1, 142) = 1.097, p = .29$. In contrast, however, participants significantly differed in their level of prejudice at time 2, as evidenced by a significant two-way interaction of time and condition, $F(1, 138) = 11.016, p < .001$, with no main or interaction effects involving the study variable. Furthermore, examination of within condition differences from time 1 to time 2, resulted in a statistically significant decrease for the negate-consistent condition, $t(76) = 2.27, p = .02$, and a significant increase for the negate-inconsistent condition, $t(65) = -2.25, p = .02$. Thus, we conclude that both directions of effect are reliable and were absent in any given study due to insufficient power to detect effects.\(^5\)

Motivation to Control Prejudiced Reactions

Participants’ Motivation to Control Prejudiced Reactions (MCPR) scores were calculated according to Dunton & Fazio, 1997 to examine if individual differences on this measure moderated the results just reported. The mean score for our sample was $M = \ldots$

\(^5\) Differences in racial prejudice (reaction time) scores as a function of valence and race (i.e. black-positivity, white-positivity, black-negativity, and white-negativity) from time 1 to time 2 were calculated for the combined data set from studies 1 and 2. A 2 (time) x 2 (condition) x 2 (study) ANOVA was run on each component (e.g. black-positivity) separately and no significant effects emerged. However, utilizing paired samples t-tests for each component by condition revealed two marginally significant effects, such that those who were in the negate-inconsistent condition were slower to associate White with negativity, $t(65) = -1.94, p = .06$ at time 2 than time 1, and slower to associate Black with positivity, $t(65) = -1.84, p = .07$, at time 2 than time 1. No significant effects emerged for the negate-consistent condition.
4.24, $SD = .589$ and as expected, mean scores did not differ significantly by condition, $t(86) = -.795, p = .92$. In order to assess if MCPR differentially affected participants’ changes in automatic racial prejudice from time 1 to time 2, condition was dummy coded (i.e., negate-consistent was coded as 1 and negate-inconsistent coded as -1), as was time (i.e., time 1 was coded as 1 and time 2 was coded as -1) and participants’ MCPR scores were centered and treated as a continuous variable.

Participants’ raw automatic racial prejudice scores pre- and post- training were submitted to a Time x Condition x MCPR regression analysis, using hierarchical linear modeling. In addition to the two way Time X Condition interaction noted earlier (see Figure 5), a marginally significant 3-way interaction of Time x Condition x MCPR, ($B = -32.09), t = -1.73, p = .08, also emerged. To understand the nature of this 3-way interaction, automatic racial prejudice at time 1 and time 2 by condition were investigated for people who were relatively high and low in MCPR. To keep MCPR as a continuous variable, these analyses were performed as a regression, with MCPR examined at one standard deviation above and below the mean. Decomposing this 3-way interaction into two, 2-way interactions of Time x Condition for each level of MCPR, we find a significant effect of Time x Condition among those people high in their MCPR ($B = -40.18), t (84) = -2.63, p = .01, but not among those low in MCPR ($B = -2.36), t (84) = -0.15, p = .87.
To further examine and depict the influence of MCPR on participants’ racial prejudice scores and to understand the three-way interaction, a median split was performed on MCPR and a 2 (training task) x 2 (time of measurement) mixed-model ANOVA with errors made on block 1 of the training task (err1) as a covariate was applied to high MCPR ($M \geq 4.30$) and low MCPR ($M < 4.29$) participants, separately. Consistent with the regression analyses just reported, among high MCPR participants, we found a significant two-way interaction of time and condition, $F (1, 40) = 5.05, p = .03$, as pictured in the top panel of Figure 6. Conversely, among low MCPR participants, the two-way interaction of time and condition was non-significant, $F (1, 42) = .016, p = .90$ as pictured in the bottom panel of Figure 6.

Within condition differences from time 1 to time 2 were also assessed separately for high and low MCPR participants. As in the pilot study and Study 1, paired samples t-tests were used. Among those participants high in their MCPR who were in the negate-consistent condition, we find a significant decrease in racial prejudice from time 1 to time 2, $t (21) = 2.06, p = .05$. However, for those participants’ who were also high in MCPR but were in the negate-inconsistent condition, we find a tendency for an increase in racial prejudice from time 1 to time 2, $t (20) = -1.58, p = .12$. Conversely, when examining those participants’ who are low in their MCPR, we find for those participants who negated consistent information, there were no differences in racial prejudice from time 1 to time 2, $t (24) = .60, p = .55$. A similar pattern emerged for those who negated
inconsistent information and were low in MCPR. Specifically, for those participants who were in the negate-inconsistent condition and low in their MCPR, there were no significant differences in their racial prejudice from time 1 to time 2, \( t (19) = .44, p = .65 \).

Thus, it appears that those individuals who are high in their MCPR are driving the observed changes in racial prejudice scores from time 1 to time 2.

**Discussion**

Consistent with the results of the previous two studies, the present study again demonstrated that meaningful negation can be a useful tool to alter levels of automatic racial prejudice. Specifically, and as predicted, those participants who engaged in meaningful negation of prejudice-consistent information demonstrated a decrease in levels of automatic racial prejudice, whereas those who engaged in meaningful negation of prejudice-inconsistent information tended to show an increase in levels of automatic racial prejudice. Furthermore, this effect was moderated by participants’ motivation to control for prejudicial responding as measured by the MCPR scale (Dunton & Fazio, 1997). Specifically, those participants who are high in their motivation to control for prejudiced responding appeared to be driving the observed changes in automatic racial prejudice from time 1 to 2, as evidenced by the significant effect of condition only for these individuals. Together with Study 1, the findings of Study 2 suggest not only that negation can indeed be a useful means to reduce racial prejudice, but that such training is
especially useful among those persons who are most motivated to control prejudicial responding.
CHAPTER 5: GENERAL DISCUSSION

The primary goal of these studies was to examine if meaningful negation training could serve as an effective tool to modify automatic racial attitudes. Prior work had suggested that simple negation of automatic prejudice was ineffective in reducing it and might even make things worse (Gawronski et al., 2008). However, consistent with the Meta-cognitive model of attitudes (Petty, Briñol, & DeMarree, 2007), we predicted that training individuals to negate prejudice-consistent information – saying “That’s wrong!” to a Black target paired with unfavorable traits, such as lazy – should result in a decrease in automatic racial prejudice as measured by an evaluative priming task (Fazio et al, 1995). In support of this idea, and contrary to previous conceptualization and research, in three studies we demonstrated that prejudice reduction training using meaningful negation could result in a decrease in automatic racial prejudice.

In a pilot study, we extended the work of Gawronski and colleagues (2008) and examined three negation training conditions. Two of the conditions (negate prejudice-consistent and negate prejudice-inconsistent) used the same simple negations (“no”) that proved ineffective in prior research. However, the pilot study
also included one new negation condition (negate-consistent–meaningful) in which participants were asked to deny their prejudice in a more meaningful way (“That’s wrong”). We found that it was only the meaningful negation of prejudice-consistent information that resulted in a significant decrease in automatic racial prejudice from time 1 (pre-training) to time 2 (post-training).

In Study 1, we sought to replicate this meaningful negation effect and further delineate the impact of meaningful negation training. Consequently, in Study 1 we utilized two conditions – negate-consistent and negate-inconsistent – and made negation training meaningful in both conditions. Doing so allowed us to see if the meaningful negation training was being applied specifically to the category and trait combinations (i.e. Black and lazy) or if responding “That’s wrong” in the context of a study about race alone would result in a decrease in automatic racial prejudice. Replicating the results of the pilot study, we found that participants who negated prejudice–consistent information demonstrated a decrease in their automatic racial prejudice from time 1 to time 2. In addition, we also found that those participants who negated prejudice–inconsistent information demonstrated an increase in their automatic racial prejudice from time 1 to 2. Thus, the findings of Study 1 suggest that meaningfully negating information in the context of a study about race is not enough to decrease one’s level of automatic racial prejudice. Instead, what one is negating, whether it is prejudice–consistent or prejudice–inconsistent information is important.
These findings indicate that negation is being applied directly to the category and trait pairs.

In Study 2, we replicated the findings of Study 1 as well as examined the influence of participants’ motivation to control their prejudiced responses (MCPR; Dunton & Fazio, 1997) on the effectiveness of the negation training task. We found that the influence of the negation training on participants’ change in automatic racial prejudice from time 1 to time 2 was primarily being driven by those participants who were high in their motivation to control for prejudiced responding. Those high in MCPR were both more likely to become less prejudiced after they practiced negating negative information associated with Black targets and they were more likely to become more prejudiced after negating positive information associated with Black targets. Taken together, these three studies provide initial evidence that meaningful negation training can serve as a useful tool to reduce individuals’ automatic racial prejudice – especially among those who are most motivated to avoid being prejudiced.

A second goal of the present research was to examine if negation was solely an on-line process or if negation could be stored for later retrieval. If negation is solely an on-line process with no long term memory trace, it would have implications for explicit judgments taken immediately after the negation (Garownski & Bodenhausen, 2006), but it would not have implications for automatic evaluations taken some time
after the negation has occurred. Consistent with the Meta-cognitive model (MCM; Petty et al., 2007) of attitudes, we hypothesized that a negation of prejudice could be stored for later retrieval, and the findings of the present studies suggest that this is the case. As noted earlier, given the differential change in automatic racial prejudice from time 1 to time 2 that we observed in two studies for the negate–consistent and negate–inconsistent conditions, it seems unlikely that negation is simply an on-line process as some models of attitudes argue (Gawronski & Bodenhausen, 2006). In order for the negation training to be impactful on participants’ racial prejudice, it seems likely that the negation was encoded during the training and later retrieved. Indeed, the negation should be retrieved rather quickly for it to impact an evaluative priming measure.

Thus, in addition to the present findings suggesting that meaningful negation training is an effective tool to reduce automatic racial prejudice, the present work also suggests that negations can be stored along with associations in memory for later retrieval.  

As noted earlier, the present work contrasts starkly with the work of Gawronski and colleagues (2008). Specifically, Gawronski and colleagues found that the

---

6 It is possible that during negation of prejudice training, instead of encoding an association of Blacks with a negative term and a negative term with an invalidity tag, participants engaged in an on-line translation process whereby Black were encoded with the opposite of the negative term presented (i.e., a positive association was formed). Although these mental gymnastics seem unlikely, they cannot definitively be ruled out by the current studies. Nevertheless, the current work still shows that negation training can effectively undermine automatic prejudice.
negation of prejudice–consistent information resulted in an increase in racial prejudice. They argued that it is only the affirmation of prejudice–inconsistent information that is capable of producing a decrease in automatic racial prejudice. In our replication of the Gawronski et al. (2008) negate prejudice–consistent condition, we found no changes in automatic racial prejudice from time 1 to time 2. Although this null finding should be treated with caution, we believe participants failed to demonstrate a change in their automatic attitudes from time 1 to time 2 because the negation being applied to the category and trait combinations was simply not impactful enough. Given that those participants who negated prejudice–inconsistent information also did not show any change in their automatic racial prejudice from time 1 to time 2, we believe that such is likely the case.

Even if our simple negation was just not impactful enough to result in a change of participants’ automatic racial attitudes, the question of why Gawronski et al. found a change (i.e., enhancing prejudice) with the same simple negation remains. One possibility is that our sample was somewhat higher in MCPR than that of Gawronski et al. Given that Gawronski et al. conducted their study in Canada, and our study was conducted in the United States where the issue of race and prejudice is a more salient issue, this may be the case. Indeed, in one study in which MCPR was assessed on a Canadian sample similar to that used in the Gawronski et al. study, the mean MCPR of the sample was 3.27 (see Gawronski, Peters, Brochu, & Strack, 2008) whereas in
our Study 2, the mean of the sample in Study 2 was 4.26. Although such a comparison must be treated cautiously, it is consistent with our speculation.

How could this difference account for our effects versus those observed by Gawronski et al.? First, consider the participants in the simple negate prejudice-consistent condition of the pilot study where no effect of negation was found. If the students in our studies were generally higher in their MCPR than those in the Gawronski et al. research, then it might have been the case that in our pilot study, participants relatively low in MCPR were showing the increase in prejudice that Gawronski et al. observed, but those relatively high in MCPR were showing a decrease. When collapsed across MCPR in that study, one would observe no overall effect for the simple negation of prejudice. However, when the negation was made more meaningful and impactful, then there was an overall effect of prejudice reduction training because those low in MCPR were showing no effect of negation training (perhaps because the increase in prejudice is at least being undermined by the negation) but those high in MCPR are showing an actual reduction in prejudice. In any event the current data lend some support to the idea that people can tag their prejudice as wrong and that this can reduce automatic prejudice.

In the present research participants were trained to meaningfully negate prejudice-consistent information to reduce their levels of automatic racial prejudice. Interestingly, the negation training required participants to negate both Black faces
paired with negative traits and White faces paired with positive traits. One direction for future research is to examine if the negation of both face-trait combinations is necessary in order for a significant decrease in automatic racial prejudice to be observed, and if so does this differ as a function of participants’ initial racial prejudice or motivation to control for racial prejudice. Recent work by Livingston & Drwecki (2007) found that non-racially biased individuals, as measured by the racial Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998) and Attitudes Towards Blacks (ATB; Brigham, 1993) Scale, were less likely to ascribe negativity to novel attitude objects, though racially biased and non-racially biased individuals were equally likely to ascribe positivity to novel attitude objects. Consequently, if non-racially biased individuals are less likely to assign negativity to novel attitude objects, perhaps meaningful negation training that only negated Black paired with negative traits would be enough to reduce automatic racial prejudice, but only for those participants’ that are already low in racial bias. Likewise, in the present work, we found that the observed changes in automatic racial attitudes were primarily driven by those participants’ high in their motivation to control for prejudice responding (Study 2).

Future research could also explore the best means to effectively reduce automatic racial prejudice among those who are low in their motivation to control prejudice. One potential means to heighten the motivation of low MCPR individuals is to
increase their self-regulatory demands. There is mounting evidence to suggest that one primary reason individuals engage in self-regulation is to combat the expression of prejudice (e.g., Devine, 1989; Dovidio & Gaertner, 1998; Monteith, 1993; Monteith, Ashburn-Nardo, Voils, & Czopp, 2002; Richeson & Trawalter, 1995; von Hippel, Silver, & Lynch, 2000; Vorauer, Hunter, Main, & Roy, 2000). Specifically, work by Devine, Monteith, and colleagues (e.g., Devine & Monteith, 1993) argues that the self-regulation of prejudice is critical to the prejudice reduction process. They find that confronting people who endorse egalitarian values with discrepancies between those values and their actual behavior leads to the reduction of bias in their subsequent behaviors (Monteith, 1993). Furthermore, these individual experiences with prejudice discrepancies and the resultant self-regulatory processes are thought to generate associations between cues to the potential for bias (e.g., the presence of a Black woman in a department store) and the need to monitor one’s thoughts and behaviors—that is, the need to engage in self-regulation (Monteith et al., 2002). By extrapolation, this work suggests that the presence of a Black individual before negation training, or the anticipation of an interracial interaction following training may be a sufficient cue to heighten self-regulatory demands, and increase the motivation of low MCPR individuals.

On other occasions, however, research suggests that more explicit cues to the presence of prejudice may be necessary in order to lead to a reduction in bias. Similar
to research examining the impact of prejudice-discrepancies on the subsequent expression of bias, research in which individuals are provided with feedback about their prejudice level also demonstrates that concerns about prejudice often result in the reduction of racial bias (Dutton & Lake, 1973; Dutton & Lennox, 1974). For instance, people who were led to believe that they responded in a racially biased way to a series of slides subsequently donated more money to a Black panhandler than individuals who were not led to believe that they responded to the slides with racial bias (Dutton & Lake, 1973). Considered in tandem with the work of Devine, Monteith, and colleagues, these studies suggest that prejudice feedback is likely to heighten concerns about appearing prejudiced during interracial interactions and, consequently, instigate a heightened need for self-regulation necessary for low MCPR individuals to demonstrate a decrease in automatic racial prejudice following the meaningful negation training.

Another important direction for future research is to determine how meaningful negation training influences downstream social judgments and behavior. A host of social psychological research has demonstrated the impact of prejudice and stereotypes on social judgments (Macrae, Milne, & Bodenhausen, 1994) and behavior (Wheeler & Petty, 2001), but far fewer have delineated how changes in automatic racial prejudice impact attitudes and behavior (Blair, 2002; Kawakami et al., 2000; Monteith, 1993). Given that past research has documented that racial prejudice is far
more predictive of discrimination than stereotyping (Dovidio & Gaertner, 1996; Esses, Haddock & Zanna, 1993; Stangor, Sullivan, & Ford, 1991), it is critical for future research to investigate the long-term effects of meaningful negation training on attitudes and behavior.

Conclusion

This thesis has shown that meaningful negation is an effective means to reduce automatic racial prejudice. In addition, the present findings are consistent with the idea that negations can be stored in memory for later retrieval. Although the complexity of changing automatic racial prejudice is far more complicated than the meaningful negation of prejudice–consistent information, the present research suggests that negation can play a significant part. Given the prevalence of the automatic activation of negative associations in the presence of Blacks and other minorities, the present research offers a useful tool to combat such negative associations. While understanding the underlying factors that contribute to automatic racial prejudice is far from complete, we believe that meaningful negation training can play an important role in changing such attitudes when one negates racial prejudice and means it.
LIST OF REFERENCES


APPENDIX A

TABLES, FIGURES AND STIMULI
<table>
<thead>
<tr>
<th>Prime</th>
<th>Before Training</th>
<th></th>
<th>After Training</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Black Prime</td>
<td>White Prime</td>
<td>Black Prime</td>
<td>White Prime</td>
</tr>
<tr>
<td>Neg Consistent</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>594</td>
<td>594</td>
<td>584</td>
<td>583</td>
</tr>
<tr>
<td>Negative</td>
<td>597</td>
<td>597</td>
<td>589</td>
<td>590</td>
</tr>
<tr>
<td>Neg Inconsistent</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>595</td>
<td>596</td>
<td>596</td>
<td>599</td>
</tr>
<tr>
<td>Negative</td>
<td>600</td>
<td>597</td>
<td>596</td>
<td>599</td>
</tr>
<tr>
<td>Neg Consistent (meaningful)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>582</td>
<td>580</td>
<td>582</td>
<td>586</td>
</tr>
<tr>
<td>Negative</td>
<td>594</td>
<td>598</td>
<td>583</td>
<td>579</td>
</tr>
</tbody>
</table>

*Table 1.* Mean response latencies in milliseconds as a function of prime (White vs. Black), Target (Positive vs. Negative), Training task condition (negate-consistent vs. negate-inconsistent vs. negate-consistent – meaningful), and Time of measurement (Before Training vs. After Training), Pilot study.
<table>
<thead>
<tr>
<th>Prime</th>
<th>Before Training</th>
<th></th>
<th>After Training</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Black Prime</td>
<td>White Prime</td>
<td>Black Prime</td>
<td>White</td>
</tr>
<tr>
<td>Neg Consistent</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>606</td>
<td>602</td>
<td>609</td>
<td>614</td>
</tr>
<tr>
<td>Negative</td>
<td>604</td>
<td>601</td>
<td>618</td>
<td>611</td>
</tr>
<tr>
<td>Neg Inconsistent</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>596</td>
<td>593</td>
<td>604</td>
<td>592</td>
</tr>
<tr>
<td>Negative</td>
<td>598</td>
<td>591</td>
<td>600</td>
<td>611</td>
</tr>
</tbody>
</table>

*Table 2. Mean response latencies in milliseconds as a function of prime (White vs. Black), Target (Positive vs. Negative), Training task condition (negate-consistent vs. negate-inconsistent), and Time of measurement (Before Training vs. After Training), Study 1*
Table 3. Mean response latencies in milliseconds as a function of prime (White vs. Black), Target (Positive vs. Negative), Training task condition (negate-consistent vs. negate-inconsistent), and Time of measurement (Before Training vs. After Training), Study 2

<table>
<thead>
<tr>
<th></th>
<th>Before Training</th>
<th>After Training</th>
<th>Black Prime</th>
<th>White Prime</th>
<th>Black Prime</th>
<th>White Prime</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neg Consistent</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>587</td>
<td>585</td>
<td>582</td>
<td>585</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>591</td>
<td>595</td>
<td>590</td>
<td>586</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neg Inconsistent</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>595</td>
<td>594</td>
<td>607</td>
<td>604</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>610</td>
<td>608</td>
<td>611</td>
<td>614</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### HI MCPR

<table>
<thead>
<tr>
<th>Prime</th>
<th>Black Prime</th>
<th>White Prime</th>
<th>Black Prime</th>
<th>White</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neg Consistent</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>588</td>
<td>586</td>
<td>576</td>
<td>582</td>
</tr>
<tr>
<td>Negative</td>
<td>596</td>
<td>599</td>
<td>595</td>
<td>587</td>
</tr>
<tr>
<td>Neg Inconsistent</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>586</td>
<td>589</td>
<td>603</td>
<td>593</td>
</tr>
<tr>
<td>Negative</td>
<td>602</td>
<td>598</td>
<td>602</td>
<td>606</td>
</tr>
</tbody>
</table>

### LO MCPR

<table>
<thead>
<tr>
<th>Prime</th>
<th>Black Prime</th>
<th>White Prime</th>
<th>Black Prime</th>
<th>White</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neg Consistent</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>587</td>
<td>585</td>
<td>588</td>
<td>587</td>
</tr>
<tr>
<td>Negative</td>
<td>586</td>
<td>592</td>
<td>586</td>
<td>586</td>
</tr>
<tr>
<td>Neg Inconsistent</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>605</td>
<td>599</td>
<td>613</td>
<td>617</td>
</tr>
<tr>
<td>Negative</td>
<td>620</td>
<td>618</td>
<td>620</td>
<td>625</td>
</tr>
</tbody>
</table>

*Table 4.* Mean response latencies in milliseconds as a function of prime (White vs. Black), Target (Positive vs. Negative), Training task condition (negate-consistent vs. negate-inconsistent vs. negate-consistent – meaningful), and Time of measurement (Before Training vs. After Training) and MCPR (HI vs. LO), Study 2
<table>
<thead>
<tr>
<th></th>
<th>Before Training</th>
<th></th>
<th>After Training</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Black Prime</td>
<td>White Prime</td>
<td>Black Prime</td>
<td>White Prime</td>
</tr>
<tr>
<td>Neg Consistent</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>595</td>
<td>592</td>
<td>592</td>
<td>596</td>
</tr>
<tr>
<td>Negative</td>
<td>596</td>
<td>597</td>
<td>601</td>
<td>596</td>
</tr>
<tr>
<td>Neg Inconsistent</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>595</td>
<td>593</td>
<td>606</td>
<td>600</td>
</tr>
<tr>
<td>Negative</td>
<td>606</td>
<td>602</td>
<td>607</td>
<td>614</td>
</tr>
</tbody>
</table>

*Table 5.* Mean response latencies in milliseconds as a function of prime (White vs. Black), Target (Positive vs. Negative), Training task condition (negate-consistent vs. negate-inconsistent), and Time of measurement (Before Training vs. After Training), Study 1 & 2 combined
Figure 1. Percent errors made by block on training task collapsed across condition (Pilot study).
Figure 2. Reaction time in milliseconds for each block of the training task collapsed across condition (pilot study).
Figure 3. Mean scores of automatic racial prejudice as a function of training task (negate-consistent vs. negate-inconsistent vs. negate-consistent – meaningful). (Pilot study).
Figure 4. Mean scores of automatic racial prejudice as a function of training task (negate-consistent vs. negate-inconsistent). Higher scores indicate a preference for Whites relative to Blacks (Study 1).
Figure 5. Mean scores of automatic racial prejudice as a function of training task, (negate-consistent vs. negate-inconsistent). Higher scores indicate a preference for Whites relative to Blacks. (Study 2).
Figure 6. Automatic racial prejudice as a function of training condition (negate-consistent vs. negate-inconsistent) and time (before vs. after training) for high MCPR (top panel) participants, $F(1, 40) = 5.05, p = .03$, and low MCPR (bottom panel) participants, $F(1, 42) = .016, p = .90$. (Study 2).
STIMULI

Stereotypical trait words used in the training task in the pilot study and Study 1 & 2

Trait words related to the negative stereotype of Black people: poor, dishonest, complaining, violent, shiftless, superstitious, lazy, threatening, dumb, hostile

Trait words related to the positive stereotype of White people: intelligent, successful, ambitious, industrious, educated, responsible, wealthy, ethical, smart, friendly

Positive and negative target words used in the evaluative priming task in the pilot study and Study 1 & 2

Positive target words: paradise, summer, harmony, freedom, honesty, honor, health, cheer, pleasure, heaven, friend, sunrise, love, relaxation, peace, vacation, happy, lucky, miracle, gift

Negative targets words: evil, sickness, vomit, bomb, murder, abuse, prison, death, assault, cancer, rotten, accident, grief, poison, stink, cockroach, virus, disaster, ugly, terror
Pictures of Black faces used on training task in the pilot study, and Study 1 & 2
Pictures of White faces used on training task in the pilot study, and Study 1 & 2