I am Kind, I am Smart, I am Important: The Effects of Self-Affirmation on Women's Confrontational Responses to Sexism

Kathryn L. Roderick
University of Maine, kathryn.roderick@maine.edu

Follow this and additional works at: https://digitalcommons.library.umaine.edu/etd

Part of the Social Psychology Commons

Recommended Citation
Roderick, Kathryn L., "I am Kind, I am Smart, I am Important: The Effects of Self-Affirmation on Women's Confrontational Responses to Sexism" (2022). Electronic Theses and Dissertations. 3665.
https://digitalcommons.library.umaine.edu/etd/3665

This Open-Access Thesis is brought to you for free and open access by DigitalCommons@UMaine. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of DigitalCommons@UMaine. For more information, please contact um.library.technical.services@maine.edu.
I AM KIND, I AM SMART, I AM IMPORTANT: THE EFFECTS OF SELF-AFFIRMATION ON WOMEN’S CONFRONTATIONAL RESPONSES TO SEXISM

By

Kathryn L. Roderick

B.S. University of Lynchburg, 2019

A THESIS

Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Arts (in Psychology)

The Graduate School
The University of Maine
August 2022

Advisory Committee:

Shannon McCoy, Associate Professor of Psychology, Advisor

Jordan LaBouff, Associate Professor of Psychology

Mollie Ruben, Assistant Professor of Psychology
I AM KIND, I AM SMART, I AM IMPORTANT: THE EFFECTS OF SELF-AFFIRMATION ON WOMEN’S CONFRONTATIONAL RESPONSES TO SEXISM

By Kathryn L. Roderick

Thesis Advisor: Dr. Shannon McCoy


The present research sought to examine the effects of self-affirmation on women’s confrontation approaches following exposure to sexism. However, before examining the effects of self-affirmation, we needed a way to measure confrontation approaches. In Study 1, participants read either prototypical HATE or CARE confrontations and assessed the response. Researchers then conducted an exploratory factor analysis to create a novel scale that can assess approaches to confrontation. Study 1 also established the reliability of the HATE and CARE subscales ($\alpha = .956$ and $\alpha = .929$, respectively). Study 2 sought to establish the construct validity of the new measure. Study 2 implemented the same procedure as Study 1 and a confirmatory factor analysis revealed that while there is room to improve the new measure, the model fit is not necessarily bad. Finally, Study 3 explored the effects of self-affirmation on women’s confrontation approaches after exposure to sexism. After manipulating self-affirmation, women participated in an imagined scenario where they responded to a male colleague making sexist comments. We hypothesized that (1) self-affirmed women would directly confront the sexism less than non-affirmed women, (2) self-affirmed women would have lower HATE scores
than non-affirmed women, (3) self-affirmed women would have higher CARE scores than non-affirmed women, (4) self-affirmed women would have lower perceived responsibly to confront compared to non-affirmed women, and (5) self-affirmed women would have lower negative state affect than non-affirmed women. Results of the statistical analyses supported Hypothesis 1. Implications for the findings of the three studies are discussed.
DEDICATION

To Mom, Dad, Ross, and Susan, without whom this would not have been possible.
ACKNOWLEDGEMENTS

I would like to thank Dr. Shelby Helwig for inspiring my research in the confrontation domain, Dr. Shannon McCoy for her statistical expertise and guidance throughout this process, and the research assistants in the McCoy lab who worked diligently to make this research possible. Lastly, I want to thank all my former educators for shaping me into the researcher I am today.
# TABLE OF CONTENTS

DEDICATION ................................................................................................................................. ii

ACKNOWLEDGEMENTS ........................................................................................................ iii

LIST OF TABLES .......................................................................................................................... ix

LIST OF FIGURES ..................................................................................................................... x

Chapter

1. GENERAL INTRODUCTION ................................................................................................. 1
   - Defining Confrontation ........................................................................................................ 1
   - The Confronting Prejudiced Responses (CPR) Model ...................................................... 2
   - Approaches to Confrontation: Techniques to Avoid ......................................................... 3
   - Approaches to Confrontation: Techniques to Employ ..................................................... 5
   - The Gap in Existing Literature .......................................................................................... 5
   - Self-Affirmation .................................................................................................................. 6
   - The Present Research ........................................................................................................ 7

2. STUDY 1: EXPLORATORY FACTOR ANALYSIS ................................................................. 8
   - Introduction ....................................................................................................................... 8
   - Method ............................................................................................................................... 9
     - Participants ..................................................................................................................... 9
     - Procedure ...................................................................................................................... 9
     - Measure ....................................................................................................................... 10
   - Results ............................................................................................................................. 11
     - Data Screening .............................................................................................................. 11
3. STUDY 2: CONFIRMATORY FACTOR ANALYSIS ................................................................. 24
   Introduction ................................................................................................................. 24
   Method ......................................................................................................................... 25
     Participants ............................................................................................................. 25
     Procedure ............................................................................................................... 25
   Results ....................................................................................................................... 25
   Data Screening .......................................................................................................... 25
   Confirmatory Factor Analysis .................................................................................. 26
     Did Model 1 Fit the Data Well? ............................................................................. 28
     Modification Index Selection Criteria .................................................................. 28
     Did Model 2 Fit the Data Well? ............................................................................. 29
   Discussion ................................................................................................................ 30

4. STUDY 3: SELF-AFFIRMATION AND CONFRONTATION ........................................ 32
   Introduction .............................................................................................................. 32
Hypotheses .................................................................................................................... 33

Hypothesis 1 .................................................................................................................... 33
Hypothesis 2 .................................................................................................................... 33
Hypothesis 3 .................................................................................................................... 34
Hypothesis 4 .................................................................................................................... 34
Hypothesis 5 .................................................................................................................... 34

Method .................................................................................................................................. 34

Participants ....................................................................................................................... 34

Procedure ......................................................................................................................... 35

Measures ................................................................................................................................ 36

Self-Affirmation Manipulation Checks ............................................................................. 36
Perceived Sexism ............................................................................................................... 36
Perceived Responsibility to Confront ............................................................................... 37
State Affect ....................................................................................................................... 37
Self-Evaluation of Confrontation Approach ................................................................ 37
Confrontation Behavior .................................................................................................. 37
Confrontation Approach ................................................................................................. 37

Results ................................................................................................................................ 38

Data Screening ................................................................................................................. 38

Manipulation Checks ....................................................................................................... 38

Self-Appraisal .................................................................................................................... 38
Value Awareness ............................................................................................................... 38
Did the Self-Affirmation Manipulation Work? ................................................................. 38
Statistical Analyses ................................................................. 39

Rate of Confrontation Behavior ................................................. 39

Creating HATE and CARE Composite Variables ................. 40

Did Confrontation Behavior Differ by Affirmation Condition? .... 40

Did Confrontation Approaches Differ by Condition? .......... 44

Did Perceived Sexism Differ by Condition? .................. 46

Does Perceived Responsibility to Confront Differ by Affirmation Condition? .................................................. 48

Did Negative State Affect Differ by Condition? .................. 49

Do Participants Perceive Their Approach to Confrontation Differently from Coders? .................................................. 51

Discussion ........................................................................... 53

5. CONCLUSIONS ...................................................................... 57

Strengths … ........................................................................... 57

Limitations ........................................................................... 59

Recommendations for Future Research ................................. 61

Future Implications .................................................................. 63

REFERENCES ........................................................................... 64

APPENDICES ........................................................................... 72

Appendix A. Studies 1 and 2 Demographic Information .......... 72

Appendix B. HATE Stimulus ................................................... 73

Appendix C. CARE Stimulus ................................................... 74

Appendix D. Initial EFA Structure Matrix ............................... 75

Appendix E. Inter-Item Correlations Between HATE and CARE Scales .......................................................... 77
Appendix F. Study 3 Demographic Information .................................................. 78
Appendix G. B-SAT Affirmation Condition .......................................................... 79
Appendix H. B-SAT No Affirmation Condition ..................................................... 80
Appendix I. PANAS-X Form ................................................................................. 81
BIOGRAPHY OF THE AUTHOR........................................................................... 82
LIST OF TABLES

Table 1. Studies 1 and 2 Demographic Information (Appendix A) .................... 72
Table 2. Confrontation Items ........................................................................... 10
Table 3. Initial EFA Structure Matrix (Appendix D) ........................................ 75
Table 4. Total Variance Explained for Initial EFA ........................................... 12
Table 5. Items for HATE and CARE Scales ...................................................... 13
Table 6. Inter-items Correlations Between HATE and CARE Scales
(Appendix E) .................................................................................................. 77
Table 7. Descriptive Statistics for HATE Composite ........................................ 15
Table 8. Descriptive Statistics for CARE Composite .......................................... 16
Table 9. Descriptive Statistics for Composites Split by Condition and Gender .... 17
Table 10. Descriptive Statistics for Efficacy Composite Split by Condition and
Gender ............................................................................................................. 20
Table 11. Fit Statistics for CFA Models .............................................................. 29
Table 12. Study 3 Demographic Information (Appendix F) ............................... 78
Table 13. Confrontation Behavior Rates .......................................................... 39
Table 14. HATE Scores Split by Confrontation Behavior ................................... 42
Table 15. CARE Scores Split by Confrontation Behavior ................................... 43
Table 16. Descriptive Statistics for Confrontation Approach Split by Condition .... 44
Table 17. Descriptive Statistics for Perceived Sexism Split by Condition .......... 47
Table 18. Relationship Between Perceived Sexism and HATE-CARE Scores ....... 48
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The Confronting Prejudiced Responses (CPR) Model</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>Mean HATE Scores for HATE and CARE Conditions Split by Gender</td>
<td>18</td>
</tr>
<tr>
<td>3</td>
<td>Mean CARE Scores for HATE and CARE Conditions Split by Gender</td>
<td>19</td>
</tr>
<tr>
<td>4</td>
<td>Mean Efficacy Scores for HATE and CARE Conditions Split by Gender</td>
<td>21</td>
</tr>
<tr>
<td>5</td>
<td>Confirmatory Factor Analysis Model 1</td>
<td>26</td>
</tr>
<tr>
<td>6</td>
<td>Confirmatory Factor Analysis Model 2</td>
<td>29</td>
</tr>
<tr>
<td>7</td>
<td>Study 3 Procedure</td>
<td>36</td>
</tr>
<tr>
<td>8</td>
<td>Confrontation Behavior Split by Condition</td>
<td>41</td>
</tr>
<tr>
<td>9</td>
<td>Mean HATE Scores Split by Confrontation Behavior</td>
<td>43</td>
</tr>
<tr>
<td>10</td>
<td>Mean CARE Scores Split by Confrontation Behavior</td>
<td>44</td>
</tr>
<tr>
<td>11</td>
<td>Mean HATE Scores Split by Condition</td>
<td>45</td>
</tr>
<tr>
<td>12</td>
<td>Mean CARE Scores Split by Condition</td>
<td>46</td>
</tr>
<tr>
<td>13</td>
<td>Perceived Responsibility to Confront Split by Condition</td>
<td>49</td>
</tr>
<tr>
<td>14</td>
<td>Mean Negative State Affect Split by Condition</td>
<td>50</td>
</tr>
<tr>
<td>15</td>
<td>Mean HATE Score Split by Perceiver</td>
<td>52</td>
</tr>
<tr>
<td>16</td>
<td>Mean CARE Score Split by Perceiver</td>
<td>53</td>
</tr>
</tbody>
</table>
CHAPTER 1
GENERAL INTRODUCTION

Envision you are a woman going about your daily life when suddenly, you overhear someone make a comment that you perceive as sexist. The remark is not directed towards you, but the rhetoric is problematic. What do you do? Do you confront the perpetrator? How do you approach the situation? What if, before hearing the comment, you were having a genuinely good day? You felt fulfilled and valued. Does that change how you approach the situation? To answer these and other questions, we need to explore what situational factors influence confrontation. However, before we can do that, we need to be able to measure confrontation.

The goal of Study 1 and Study 2 was to develop and validate a scale that analyzes approaches to confrontation. Then, we could implement that scale to help us explore what factors, such as self-affirmation, influence women’s confrontation approach following exposure to sexism (Study 3). Conflicting viewpoints frequently result in arguments that are less than productive. By exploring the nuances of confrontation, we gain insight regarding the tools and resources needed to avoid arguments and have more productive conversations.

Defining Confrontation

Researchers define confrontation as a construct in a number of ways. Most often, researchers choose to distinguish between direct and indirect confrontation. Direct confrontation refers to explicitly addressing the prejudice, whereas indirect confrontation tactics signal that there is a problem, but the prejudice is never explicitly addressed (Brett et al., 2014). For the purposes of the present research, we chose to define confrontation in a way that would capture both direct and indirect confrontations. Therefore, we define confrontation as addressing the prejudice, either overtly or covertly, to signal that there is a problem. In some analyses where we
felt it was beneficial, we distinguish between direct and indirect confrontation. In those cases, we highlight our decision to decipher the two. Furthermore, we define no confrontation in the following research as failing to anticipate saying or doing anything to address the prejudice.

**The Confronting Prejudiced Responses (CPR) Model**

The confronting prejudiced responses (CPR) model, proposed by Ashburn-Nardo and colleagues (2008), defines the steps that lead to interpersonal confrontation. The CPR model states interpersonal confrontation is motivated by (1) recognition of the event/comment as prejudiced, (2) perceiving the prejudice as an emergency that needs to be addressed, (3) feeling personally responsible to intervene, (4) deciding a course of action, and (5) deciding to act (see Figure 1).

**Figure 1**

*Confronting Prejudiced Responses (CPR) Model*

Note. Diagram adapted from Ashburn-Nardo et al. (2008).

The model authors acknowledge that not all steps are required for a confrontation to occur, nor is it a requirement that the steps occur sequentially. According to the CPR model,
before the confrontation encounter ensues, confronters must decide on a course of action (Step 4). Similar to the varying definitions of what a confrontation is, confronting individuals can choose different courses of action. Essentially, confronters can approach the confrontation in different ways.

**Approaches to Confrontation: Techniques to Avoid**

Monteith and colleagues (2019) identified four tactics that are best avoided when confronting a perpetrator of prejudice. Those tactics include hostility, aggression, threat, and extremity, often referred to collectively as HATE tactics. Avoiding HATE tactics helps minimize negative evaluations from others (Becker & Barreto, 2014), which, in turn, potentially promotes generally more successful interactions. Martinez et al. (2017) illustrated this notion in an experiment that examined the implications of confronting homophobia in a work environment. The researchers manipulated hostility (hostile vs. calm), confrontation directness (direct vs. indirect), and confronter type (ally vs. target) and randomly assigned participants to watch one of nine videos (i.e., eight conditions and one control). The videos depicted two White colleagues chatting about a gay employee who was fired. In the videos, the perpetrator makes a comment that the former employee’s termination is entirely justified because it is not appropriate to be gay in a workplace setting. The confronter’s response varied by condition. After watching the videos, participants rated their perceptions of the confronter. Martinez et al. (2017) found a main effect for hostility; participants perceived hostile confronters more negatively relative to calm confronters and people who did not confront. The results suggest that using a hostile confrontation approach can result in negative evaluations from others. Furthermore, the Martinez et al. (2017) study offers empirical support for Becker and Barreto’s (2014) claim that it is best to avoid HATE tactics if
maintaining positive evaluations from others (i) is of concern to the confronter or (ii) can promote more productive interactions.

Other common techniques in confrontation settings include responding with humor or sarcasm (Swim & Hyers, 1999). Humor differs from the other HATE techniques in that it potentially increases the likability of the confronter in the eyes of both the confronted and bystanders. While humor aids in maintaining a positive evaluation of the confronter, some scholars claim it is best avoided. The results from a study that examined the effects of humor on the perceived likability of the confronter and the effectiveness of the confrontation, conducted by Woodzicka and colleagues (2020), illustrate why humor may not be the best tactic to implement during a confrontation. Participants read a scenario in which they imagined overhearing someone make a sexist joke followed by the perpetrator’s friend confronting either with or without humor (i.e., humor vs. serious). Participants then rated likability of the confronter and effectiveness of the confrontation using a 1 (not at all) to 5 (extremely) scale. Woodzicka et al. (2020) found participants rated confronters who used humor in their confrontations as more likable but less effective than confronters who were serious in their response. Although humor as a confrontation tactic has the potential to help maintain positive evaluations of the confronter, the results yielded by Woodzicka et al. (2020) suggest humor is perceived as a less effective approach to confrontations and should be avoided.

Overall, the literature suggests techniques such as hostility, aggression, targeting/threat tactics that challenge the perpetrator’s self-image in some way, extremity, and humor are best avoided when confronting prejudice. What then is the alternative to these HATE tactics?
Approaches to Confrontation: Techniques to Employ

In addition to techniques that scholars recommend confronters avoid, there are tactics researchers believe are desirable to employ in confrontation settings. Czopp et al. (2006) recommend confronters remind the perpetrator of their egalitarian self-image. Other techniques include utilizing collectivistic language, being assertive, being respectful towards the perpetrator, examining the other party’s intentions (i.e., asking questions), and using evidence-based arguments as a vehicle to educate the perpetrator (Monteith et al., 2019; Nelson et al., 2011; Plous, 2000; Sue et al., 2019). For ease, we will refer to these confrontation techniques as CARE tactics. Note that the existing literature does recommend CARE tactics. However, the research concerning CARE techniques is not as well established as it is for the HATE techniques discussed above.

The Gap in Existing Literature

One goal of the research was to develop a measure to help assess the approach (i.e., HATE or CARE) confronters utilize in their interactions with perpetrators of prejudice. Only recently have scholars started to develop possible measures to assess people’s confrontation responses to prejudice. Chaney and Sanchez (2021) developed the Prejudice Confrontation Styles (PCS) scale, which assesses educational, argumentative, help-seeking, empathy, and humor approaches to confrontation. However, currently, no measure exists that allows researchers to distinguish between the use of productive and unproductive confrontation tactics. In this context, productivity refers to whether scholars recommend confronters employ a particular tactic.

While researchers are starting to create scales to identify and elucidate confrontation styles, research seldom explores what factors influence an individual’s chosen response during a confronting interaction. Most of the existing literature focuses on the frequency of confrontation,
situational factors that influence women’s willingness to directly confront sexism, and perceptions of confronters (Hennessey, 2018; Shelton & Stewart, 2004; Swim & Hyers, 1999; Vaccarino & Kawakami, 2020; Woodzicka & Good, 2021).

The present research sought to address a significant gap in the confrontation literature by exploring what influences the use of HATE versus CARE responses. More specifically, the research explored if antecedents to Step 4 in the CPR model, such as self-affirmation, alleviates the urge to utilize unproductive styles (i.e., HATE tactics), thus influencing women’s approaches to confrontation following exposure to sexism.

**Self-Affirmation**

Self-affirmation theory posits that an individual’s thoughts and behavior are a byproduct of one’s motivation and desire to maintain a positive self-image as moral and adaptive (Aronson et al., 1999; Steele, 1988). Steele (1988) argues it is easier to remain objective about information that threatens the self when self-affirming thoughts are made salient. As a result, self-affirmation is a potentially effective way of reducing defensive responses. Sherman and colleagues (2000) supported this notion when the researchers manipulated self-affirmation and concluded that participants who self-affirmed prior to receiving threatening medical information were more receptive to the medical messages compared to individuals who did not self-affirm. Additional research suggests that affirming the self bolsters resources that can stifle the reflexive need to enhance the self following threat exposure (Schmeichel & Vohs, 2009; Sherman & Cohen, 2006). Thus, self-affirmation might be a useful coping mechanism to help buffer the self against threatening information.

What then constitutes as threatening information? Social identity theory, proposed by Tajfel and Turner (1979), posits that identification with a group influences one’s well-being,
namely because the group is viewed as part of the psychological self (McCoy & Major, 2003). Thus, an attack on an individual within a group is felt by those who identify with said group. For example, a woman might respond negatively to hearing sexist comments, even if those comments are targeted at another woman, because of her identification with her gender group. Barreto and Ellemers (2005) determined that women typically react with anger and hostility when exposed to blatant sexism, even when the sexism is not directly targeted at the female participant. If self-affirming reduces defensive responses, as Steele (1988) suggests, then it is plausible that self-affirming before exposure to sexism can alter women’s reactions. Specifically, we would expect self-affirming to help curb the hostile reactions Barreto and Ellemers (2005) saw. One goal of the proposed research is to explore how self-affirming prior to exposure to sexist comments influences the confrontation approaches of women.

The Present Research

The goals of the present research were twofold. First, develop and validate a novel scale to help assess confrontation approaches (Study 1 and Study 2). Second, implement the novel scale and explore antecedents to Step 4 in the CPR model that may alleviate the urge to utilize unproductive (i.e., HATE) approaches when confronting a perpetrator (Study 3). In sum, the proposed research sought to address a significant gap in the confrontation literature by exploring what influences the use of HATE versus CARE responses. More specifically, the research investigated if self-affirmation influences women’s approaches to confrontation following exposure to sexism.
CHAPTER 2

STUDY 1: EXPLORATORY FACTOR ANALYSIS

Introduction

One criticism gaining more traction in the psychological sciences is that measurement practices are substandard (Borsboom, 2006; Scheel et al., 2021). In order to obtain meaningful results in a study, we first need to ensure the measure itself has meaning, especially in instances where the variable of interest is latent, as is the case with most psychological research. To achieve a higher level of quality, Flake and colleagues (2017) recommend that authors make a more conscious effort to validate original scales.

Study 3 seeks to explore how self-affirmation influences approaches to confrontation. However, before we explore the causal relationship between self-affirmation and confrontation responses, we first need a way to measure the confrontation approach. We are aware of only one other scale developed to measure confrontation approaches, the Prejudice Confrontation Styles (PCS) Scale (Chaney & Sanchez, 2021). While the scale helps assess educational, argumentative, help-seeking, empathy, and humor approaches to confrontation, we feel it does not encapsulate all the tactics confrontation researchers recommend and discourage in the literature. Thus, in Study 1, we sought to create and validate a new scale to assess confrontation approaches.

Exploratory factor analysis (EFA) is the preferred method in scale development because the statistical approach helps identify underlying latent dimensions (Simms & Watson, 2007). Researchers recommend starting with an overinclusive item pool when constructing a scale and selecting items for an EFA (Clark & Watson, 2019). This recommendation stems from the notion that exploratory factor analysis can help identify which items to drop due to lack of relevance to the theoretical construct, but it cannot identify missing content. After running an exploratory
factor analysis and selecting the most relevant items, researchers can assess internal consistency with Cronbach’s alpha. Cronbach’s alpha provides information regarding how well the included items measure the same construct, so the higher the alpha, the more reliable the scale items (Tavako & Dennick, 2011).

The goal of Study 1 was to develop and validate a scale to help assess the approach confronters use when addressing prejudice. An exploratory factor analysis helped determine which items to include on the finalized scale. Reliability (i.e., internal consistency) was also established in the study. Additional analyses were included to explore binary gender differences among the variables and gain insight regarding how effective participants perceive the two distinguished approaches to confrontation.

Method

Participants

A total of 300 University of Maine students participated in the study (see Table 1 in Appendix A for demographic information). All participants were at least 18 years of age and recruited via the SONA system. Per Comrey and Lee’s (2013) recommendation, a sample size of 300 is considered ‘good’ for a factor analysis. Students were compensated with one research credit for their participation in the study.

Procedure

After completing informed consent, participants read a scenario. In the presented scenario, two members of a hiring committee were deciding between two candidates (Robert and Rebecca) for a lab manager position when one of the committee members comments, “Like most women, Rebecca would probably be too emotional to handle a competitive lab environment. To be a researcher, you need to be focused and rational, and not having those qualities means she
probably won’t have strong ideas to contribute.” The scenario was adopted from a study conducted by Helwig et al. (2022) and was previously effective in eliciting confrontation responses from participants. Participants were then randomly assigned to read either a stereotypical HATE (Appendix B) or CARE (Appendix C) response to the comment. We developed HATE and CARE stimuli using actual confrontation responses from participants in other studies conducted in the McCoy lab (Helwig & McCoy, 2022a). After reading the response, participants answered a series of surveys and were debriefed. The entire study was completed on the Qualtrics survey site.

**Measure**

After reading the response, participants answered 71 survey items about the confrontation on a 1 (*Strongly Disagree*) to 9 (*Strongly Agree*) Likert-type scale. Participants responded to how strongly they felt the words described the confrontation. Prior to data collection, the items were classified into HATE and CARE categories based on discussions in the confrontation literature, which outlined suggestions on which approaches to implement and avoid when confronters address prejudice (Czopp et al., 2006; Monteith et al., 2019). See Table 2 for items and pre-data collection HATE-CARE classification.

**Table 2**

<table>
<thead>
<tr>
<th>Confrontation Items</th>
<th>HATE</th>
<th>CARE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abrasive</td>
<td>Hostile</td>
<td>Agreeable</td>
</tr>
<tr>
<td>Accusatory</td>
<td>Humorous</td>
<td>Amiable</td>
</tr>
<tr>
<td>Aggressive</td>
<td>Immature</td>
<td>Assertive</td>
</tr>
<tr>
<td>Antagonistic</td>
<td>Impassive</td>
<td>Assured</td>
</tr>
<tr>
<td>Argumentative</td>
<td>In-your-face</td>
<td>Bold</td>
</tr>
<tr>
<td>Attacking</td>
<td>Ineffective</td>
<td>Careful</td>
</tr>
<tr>
<td>Childish</td>
<td>Juvenile</td>
<td>Civil</td>
</tr>
<tr>
<td>Combative</td>
<td>Malicious</td>
<td>Collected</td>
</tr>
<tr>
<td>Confrontational</td>
<td>Mean</td>
<td>Compassionate</td>
</tr>
</tbody>
</table>
Results

Data Screening

Before analysis, respondents were excluded if they were (i) identified as a straight-liner (i.e., the participant had a standard deviation equal to zero on the 71 confrontation items) or (ii) if they failed an attention check embedded within the confrontation items. After excluding straight-liners and inattentive respondents, total sample size was $N = 235$ ($N = 124$ in HATE condition, $N = 111$ in CARE condition). No additional outliers were identified during data inspection. While the initial goal was to have a sample size totaling 300 to achieve ‘good’ status as recommended by Comrey and Lee (2013), 200 participants is considered ‘fair’ and acceptable for a sample when conducting a factor analysis, so the exploratory analysis proceeded as planned.

Exploratory Factor Analysis

The objective of the exploratory factor analysis was to craft a condensed HATE-CARE scale to help identify the type of confrontation approach confronters use when addressing prejudice. While 71-items were initially included, we never intended for the final scale to include all items. Thus, we conducted the exploratory analysis with the goals of (i) condensing the number of items required for a two-factor HATE-CARE scale and (ii) determining which items were most important to include.
It was unclear whether the latent factors of HATE and CARE would be negatively correlated (i.e., is HATE simply the absence of CARE and vice versa?). As a result, we opted to use an oblimin rotation as opposed to a varimax in the analysis because orthogonal methods such as varimax restrict correlations between latent variables, whereas oblique rotation methods like oblimin allow correlations between factors (Brown, 2009). Initially, the exploratory factor analysis revealed the 71-items loading onto eight factors (see Table 3 in Appendix D). Cumulatively, the first three factors accounted for roughly 62.96% of the initial eigenvalue variance, with the first factor accounting for 49.62% of the variance (see Table 4). Looking at the structure matrix (Table 3), it was apparent that items did not load as cleanly onto factors four through eight. We confirmed this observation when examining the total variance explained (see Table 4 below).

### Table 4

*Total Variance Explained for Initial EFA*

<table>
<thead>
<tr>
<th>Factor</th>
<th>Initial Eigenvalues</th>
<th>% of Variance</th>
<th>Cumulative %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>49.618</td>
<td>49.618</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>7.737</td>
<td>57.354</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>5.603</td>
<td>62.957</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>2.719</td>
<td>65.676</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>2.088</td>
<td>67.764</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1.903</td>
<td>69.668</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>1.617</td>
<td>71.285</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>1.456</td>
<td>72.741</td>
<td></td>
</tr>
</tbody>
</table>

*Note.* Cumulative percent does not go to 100 because initial EFA input criteria indicated a value of 25 for the model iteration. Only the eight factors relevant to the discussion are shown.

The percent of variance explained decreased to about 2.72% in factor four and continued to decrease for all factors that followed. When evaluating the structure matrix (Table 3), we noticed that all the items that loaded onto factor one were pre-classified as HATE items. The
items that loaded onto factors two and three were pre-classified CARE items. Because several of the items loaded onto multiple factors at the .40 level, we felt comfortable moving forward with a two-factor goal.

**Creating HATE and CARE Scales.** After the initial exploratory factor analysis, we selected items to shorten the overall length of the HATE-CARE scale. We considered several influences when establishing retention criteria for items. First, as is standard in the field, only items with factor loadings greater than .40 were considered for the final scale (Bandalos & Finney, 2010; Lloret-Segura et al., 2014). Next, we aimed to have the scales be of equal length, so we kept that in mind when selecting items. Lastly, we consulted the existing confrontation literature to help mold the subscales (see Czopp et al., 2006 and Monteith et al., 2019 for discussions on which tactics are best implemented and avoided when confronting a perpetrator of prejudice). Table 5 below shows the items chosen for the final scale and their factor loadings.

**Table 5**

*Items for HATE and CARE Scales*

<table>
<thead>
<tr>
<th>Item</th>
<th>HATE Scale ($\alpha = .956$) Factor Loading</th>
<th>CARE Scale ($\alpha = .929$) Factor Loading</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attacking</td>
<td>.941</td>
<td>Civil</td>
</tr>
<tr>
<td>Hostile</td>
<td>.902</td>
<td>Thoughtful</td>
</tr>
<tr>
<td>Targeting</td>
<td>.861</td>
<td>Respectful</td>
</tr>
<tr>
<td>Hateful</td>
<td>.849</td>
<td>Collected</td>
</tr>
<tr>
<td>Extreme</td>
<td>.825</td>
<td>Rational</td>
</tr>
<tr>
<td>Abrasive</td>
<td>.810</td>
<td>Careful</td>
</tr>
<tr>
<td>Combative</td>
<td>.778</td>
<td>Understanding</td>
</tr>
<tr>
<td>Forceful</td>
<td>.768</td>
<td>Genuine</td>
</tr>
<tr>
<td>Threatening</td>
<td>.766</td>
<td>Compassionate</td>
</tr>
<tr>
<td>Accusatory</td>
<td>.735</td>
<td>Dignified</td>
</tr>
</tbody>
</table>

*Note.* Factor loadings are from the exploratory analysis that was conducted after pulling items that satisfied selection criteria.
We then conducted another exploratory factor analysis to assess the factor loadings of the final 20 items that we chose. Again, because we felt there was a possibility that the latent variables of HATE and CARE might be negatively related, we implemented an oblimin rotation to allow for correlations between the factors. The exploratory factor analysis revealed items loading onto two factors (see Table 5 above). The decision to use an oblique rotation method was warranted, seeing the factor correlation matrix revealed the factors were moderately correlated at $r = -.604$.

HATE accounted for 58.29% of the initial eigenvalue variance, whereas CARE accounted for 9.74% of the initial eigenvalue variance. The exploratory factor analysis also produced chi-squared goodness-of-fit test output. In an exploratory factor analysis, it is preferable that the goodness-of-fit test is not significant because a significant result suggests the model does not fit the data, and there may be another factor that can be extracted (Thompson, 2004). Examination of the goodness of fit analysis revealed a significant value $[\chi^2(151, N = 235) = 433.53, p < .001]$. However, this was not a major concern because for two reasons. First, the chi-squared value reduced substantially from the initial exploratory factor analysis $[\chi^2(1945, N = 235) = 2719.54, p < .001]$, suggesting that the two-factor model did not fit as poorly as the original eight-factor model. Second, statisticians argue that the chi-squared goodness-of-fit test is extremely sensitive to sample size and, therefore, may not be the best or most accurate approach when assessing model fit (Tong & Bentler, 2013).

**Reliability.** Next, we calculated inter-item correlations between HATE and CARE to confirm there were no redundant items. Analysis revealed no redundant items, with all correlations $<|.79|$ (see Table 6 in Appendix E for all correlations). Cronbach’s alpha values for HATE and CARE revealed strong reliability within the two factors, $\alpha = 0.956$ and $\alpha = 0.929$, respectively.
respectively. A value of .80 is a widely accepted level of internal consistency (Nunnally, 1978; Streiner, 2003). Since both HATE and CARE had alpha values greater than .90, we felt the scale had sufficient internal consistency.

**Creating HATE and CARE Composite Variables.** We calculated the HATE composite from averaging the score across the 10 HATE items. Because reliability was strong for the HATE items ($\alpha = 0.956$), participants needed to respond to a minimum of eight HATE items to calculate a mean composite score. Higher HATE composite scores indicate greater use of HATE tactics in a confrontation. Originally, items were rated using a 1 (*Strongly Disagree*) to 9 (*Strongly Agree*) Likert-type scale. In this case, a value of 5 indicates the midpoint of the scale. Descriptive statistics for the HATE composite can be found in Table 7.

**Table 7**

*Descriptive Statistics for HATE Composite*

<table>
<thead>
<tr>
<th></th>
<th>Minimum</th>
<th>Maximum</th>
<th>$M$</th>
<th>$SD$</th>
<th>$n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>HATE Condition</td>
<td>1</td>
<td>9</td>
<td>6.44</td>
<td>1.41</td>
<td>124</td>
</tr>
<tr>
<td>CARE Condition</td>
<td>1</td>
<td>7.40</td>
<td>2.82</td>
<td>1.47</td>
<td>111</td>
</tr>
<tr>
<td>Overall</td>
<td>1</td>
<td>9</td>
<td>4.73</td>
<td>2.31</td>
<td>235</td>
</tr>
</tbody>
</table>

Next, we calculated the CARE composite using the same technique as the HATE composite. Again, because reliability was strong across the CARE items ($\alpha = 0.929$), a minimum of eight responses to the CARE items was required to calculate the composite score. Similar to the HATE composite, items were originally rated using a 1 (*Strongly Disagree*) to 9 (*Strongly Agree*) Likert-type scale. In this case, a value of 5 indicates the midpoint. Additionally, higher CARE composite scores indicate greater use of CARE confrontation techniques. Descriptive statistics for the CARE composite can be found in Table 8.
Table 8

Descriptive Statistics for CARE Composite

<table>
<thead>
<tr>
<th>Condition</th>
<th>Minimum</th>
<th>Maximum</th>
<th>M</th>
<th>SD</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>HATE Condition</td>
<td>1</td>
<td>8.20</td>
<td>3.68</td>
<td>1.44</td>
<td>124</td>
</tr>
<tr>
<td>CARE Condition</td>
<td>1</td>
<td>9</td>
<td>6.45</td>
<td>1.51</td>
<td>111</td>
</tr>
<tr>
<td>Overall</td>
<td>1</td>
<td>9</td>
<td>4.99</td>
<td>2.02</td>
<td>235</td>
</tr>
</tbody>
</table>

Relationship Between HATE and CARE Composite Variables. The exploratory factor analysis revealed a moderate, negative correlation between the HATE and CARE factors. We then ran a series of correlations to explore the relationship between the HATE-CARE composites within stimulus condition. For participants in the HATE stimulus condition, there was a significant, moderate, negative correlation between composite scores ($r = -.42, p < .001$). There was a similar relationship between composite scores for participants in the CARE stimulus condition ($r = -.51, p < .001$). While there is a moderate correlation between HATE and CARE, the relationship does not appear strong enough to definitively state that the constructs are reciprocals of one another. Thus, it is possible that HATE may not simply be the absence of CARE and vice versa.

Do HATE and CARE Composites Differ by Condition? After creating the composite variables, we ran a series of independent samples t-tests to confirm that participants in the HATE stimulus condition had higher HATE scores, and participants in the CARE stimulus condition had higher CARE scores. Analysis of the HATE composite scores confirmed that participants in the HATE stimulus condition ($M = 6.44, SD = 1.41$) had significantly higher scores on the HATE composite compared to participants in the CARE stimulus condition ($M = 2.82, SD = 1.47$); $t(233) = 19.23, p < .001, d = 2.51$, large. Similarly, analysis of the CARE composite scores confirmed that participants in the CARE stimulus condition ($M = 6.45, SD = 1.51$) had
significantly higher CARE composite scores relative to participants in the HATE stimulus condition ($M = 3.68, SD = 1.44$); $t(233) = -14.36, p < .001, d = 1.88$, large.

**Does Gender Influence Perceived Confrontation Response?** Next, we ran a series of $2 \times 2$ condition (HATE stimulus vs. CARE stimulus) factorial ANOVAs on the HATE and CARE composites to determine if gender has any influence on confrontation response ratings. We are aware that including both the independent samples t-test and the factorial ANOVAs may be redundant. However, we felt it necessary to include the independent samples t-test because it allows those who do not identify with the gender binary to be included in the analysis. Gender nonconforming participants and participants who did not specify gender were excluded from the factorial analyses due to insufficient sample size. Additionally, we acknowledge that the study was not designed with the intention of examining gender differences in perceived confrontation response. Ultimately, we opted to explore the relationship for any noteworthy trends. Also note that the sample size among the groups is unequal (see Table 9 below), so all results should be interpreted with caution.

**Table 9**

*Descriptive Statistics for Composites Split by Condition and Gender*

<table>
<thead>
<tr>
<th>Condition</th>
<th>Gender</th>
<th>$n$</th>
<th>HATE Composite</th>
<th>CARE Composite</th>
</tr>
</thead>
<tbody>
<tr>
<td>HATE</td>
<td>Men</td>
<td>80</td>
<td>6.55</td>
<td>1.36</td>
</tr>
<tr>
<td></td>
<td>Women</td>
<td>30</td>
<td>6.40</td>
<td>1.37</td>
</tr>
<tr>
<td>CARE</td>
<td>Men</td>
<td>67</td>
<td>2.90</td>
<td>1.42</td>
</tr>
<tr>
<td></td>
<td>Women</td>
<td>31</td>
<td>2.60</td>
<td>1.64</td>
</tr>
</tbody>
</table>

*Note.* Composite scores can range between 1 and 9, therefore a value of 5 indicates the midpoint.

Analysis of the HATE composite revealed no significant interaction between gender and condition, $F (1, 204) = 0.134, p = .715, \eta^2_p = .001$. Main effects revealed no significant effect of gender, $F (1, 204) = 1.04, p = .31, \eta^2_p = .005$, and a significant effect of condition, $F (1, 204) =$
292.78, $p < .001$, $\eta^2_p = .589$. Thus, it appears that gender had no influence on how participants perceive HATE confrontation responses (see Figure 2).

**Figure 2**

*Mean HATE Scores for HATE and CARE Conditions Split by Gender*

![Chart showing mean HATE scores for HATE and CARE conditions split by gender, with error bars indicating +/- standard error mean.]

*Note.* Error bars indicate +/- standard error mean.

Analysis of CARE composites yielded different results. There was no main effect of gender, $F(1, 204) = 0.264, p = .608$, $\eta^2_p = .001$. There was a main effect of condition, $F(1, 204) = 191.77, p < .001$, $\eta^2_p = .585$. However, the effect of condition is qualified by a significant interaction effect, $F(1, 204) = 4.91, p = .028$, $\eta^2_p = .024$. To further explore the results, the omnibus analysis was followed up with exploration of simple main effects.

We followed up the omnibus ANOVA with pairwise comparisons to help investigate the simple main effects. Looking at gender within condition, there was not a significant difference between men and women’s CARE composite scores in the HATE condition [$F(1, 204) = 7.935$, $p = .053$, $\eta^2_p = .018$] or the CARE condition [$F(1, 204) = 2.998, p = .233$, $\eta^2_p = .007$; see Figure
The significant interaction effect in the omnibus ANOVA may be driven by some type of error resulting from the unequal sample sizes, but we cannot be sure of this without additional research in the future.

**Figure 3**

*Mean CARE Scores for HATE and CARE Conditions Split by Gender*

![Bar chart showing mean CARE scores for HATE and CARE conditions split by gender.](image)

*Note.* Error bars indicate +/- standard error mean.

**How Effective do Participants Perceive HATE and CARE Confrontation**

**Approaches?** In addition to the overall HATE-CARE approach, we investigated how effective participants perceived the confrontation response. We had no *a priori* hypotheses related to perceived efficacy for the confrontation response but chose to explore this variable, seeing as it is often discussed in the confrontation literature.

Using items that were not pulled to create the HATE and CARE scales, we created a composite variable for efficacy. We calculated the efficacy composite by averaging across the following items: effective, productive, constructive, ineffective (reverse scored), unproductive (reverse scored), and destructive (reverse scored). Since reliability was strong for the efficacy
items \((\alpha = 0.926)\) participants needed to respond to a minimum of five items to calculate a mean composite score (see Table 10 for descriptive statistics). Higher efficacy composite scores indicate greater perceived effectiveness of the confrontation response. Originally, items were rated using a 1 (Strongly Disagree) to 9 (Strongly Agree) Likert-type scale. In this case, a value of 5 indicates the midpoint of the scale.

**Table 10**

*Descriptive Statistics for Efficacy Composite Split by Condition and Gender*

<table>
<thead>
<tr>
<th>Condition</th>
<th>Gender</th>
<th>(n)</th>
<th>(M)</th>
<th>(SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HATE</td>
<td>Men</td>
<td>80</td>
<td>4.32</td>
<td>1.91</td>
</tr>
<tr>
<td></td>
<td>Women</td>
<td>30</td>
<td>4.41</td>
<td>1.68</td>
</tr>
<tr>
<td></td>
<td>Collapsed</td>
<td>110</td>
<td>4.35</td>
<td>1.84</td>
</tr>
<tr>
<td>CARE</td>
<td>Men</td>
<td>67</td>
<td>7.24</td>
<td>1.34</td>
</tr>
<tr>
<td></td>
<td>Women</td>
<td>31</td>
<td>7.37</td>
<td>1.76</td>
</tr>
<tr>
<td></td>
<td>Collapsed</td>
<td>98</td>
<td>7.28</td>
<td>1.48</td>
</tr>
</tbody>
</table>

Following the creation of the composite variable, we conducted a 2 gender (men vs. women) x 2 condition (HATE stimulus vs. CARE stimulus) factorial ANOVA to simultaneously examine differences in perceived efficacy between condition and gender. Similar to the factorial analysis above, gender nonconforming participants and participants who did not specify gender were excluded from the analysis due to insufficient sample size. Again, the sample size among the groups is unequal (see Table 10 above), so all results should be interpreted with caution.

The analysis revealed\(^1\) no significant interaction between gender and condition, \(F(1, 204) = 0.007, p = .933, \eta^2_p = .000\). Main effects revealed no significant effect of gender, \(F(1, 204) = 0.192, p = .662, \eta^2_p = .001\), and a significant effect of condition, \(F(1, 204) = 130.149, p < .001, \eta^2_p = .389\). Thus, it appears that gender had no influence on how effective participants

---

\(^1\) Failed Levene’s Test of homogeneity of variance \((p = .008)\). As ANOVAs are relatively robust to this violation, data interpretation proceeded as normal (Ito, 1980).
perceived the confrontation responses. Furthermore, participants in the CARE stimulus condition perceived the confrontation responses a significantly more effective than participants in the HATE stimulus condition (see Figure 4).

**Figure 4**

*Mean Efficacy Scores for HATE and CARE Conditions Split by Gender*

![Bar chart showing mean efficacy scores for HATE and CARE conditions split by gender.

Note. Error bars indicate +/- standard error mean.*

**Discussion**

Results from the initial exploratory factor analysis revealed items loading onto eight factors. After applying selection criteria, the final 20-items of the HATE-CARE scale loaded onto two factors. The obtained alpha values for HATE ($\alpha = 0.956$) and CARE ($\alpha = 0.929$) suggest that the developed scale has sufficient internal consistency. While the two factors were negatively correlated, the correlations are not strong enough to claim that HATE and CARE are simply endpoints of the same construct. Thus, it is plausible that the two are separate constructs.

We then conducted a series of factorial ANOVAs to determine if there were any gender differences in the scales. It is critical to note that the analyses were conducted for exploratory
purposes, so there were no \textit{a priori} hypotheses. Additionally, Study 1 was not designed with the intent to explore gender differences, and the sample sizes between groups were unequal, so the obtained results are interpreted with caution. First, participants in the HATE stimulus condition did have higher HATE composites relative to participants in the CARE stimulus condition. Furthermore, it does not appear that gender had any influence over perceived confrontation response, in that men did not perceive more HATE tactics in the confrontation than women or the contrary. This trend was similar for the CARE composite. Participants in the CARE stimulus condition had higher CARE composite scores compared to participants in the HATE stimulus condition. While the omnibus ANOVA did reveal a significant interaction between gender and condition, analysis of simple main effects revealed no gender differences. Thus, women did not perceive more CARE tactics in the confrontations than men.

Taken together, the results suggest there are no gender differences in how the different approaches to confrontation are perceived, which further leads us to believe that this scale can be used by anyone regardless of gender. However, as noted above, Study 1 was not designed with the purpose of exploring gender differences. Furthermore, due to insufficient sample size, the analyses only consider gender as defined by the binary (i.e., our analyses only accounted for those who identified as men or women). Thus, additional research is needed to determine if it is appropriate to implement the novel scale with populations that do not identify with or conform to the gender binary.

Lastly, we chose to explore how effective participants perceive the different types of confrontation approaches as it is a topic garnering more attention within the confrontation literature. The results suggest that regardless of gender, participants in the CARE stimulus condition perceived the confrontation response as more effective than participants in the HATE
stimulus condition. It is acknowledged within the confrontation literature that there are numerous ways to define the success or effectiveness of a confrontation interaction (Becker & Barreto, 2019; Brett et al., 2014; Czopp et al., 2006; Drury & Kaiser, 2014; Haslett & Lipman, 1997; Hyers, 2007). It is important to note that the present study does not clearly define efficacy. Rather, exploration of the efficacy variable was included to gain a more descriptive understanding of how participants view different approaches to confrontation. Consequently, additional research is required to develop a more comprehensive understanding of the effectiveness of these two approaches to confrontation.

Study 1 established internal consistency for the novel HATE-CARE scale and offered some preliminary insight regarding gender differences in perceptions of confrontation responses. While it is good practice to establish internal consistency for novel measures, as Scheel and colleagues (2021) point out, it is also a minimum requirement for researchers. Researchers should strive to reach a more rigorous standard to ensure the quality of novel scales in psychological research (Borsboom, 2006; Clark & Watson, 2019; Scheel et al., 2021). In an effort to do that, we followed up the exploratory factor analysis in Study 1 with a confirmatory factor analysis in Study 2 with the goal of establishing construct validity for the novel scale.
CHAPTER 3

STUDY 2: CONFIRMATORY FACTOR ANALYSIS

Introduction

In social psychology, researchers rarely do any scale validation on author-developed scales beyond reporting Cronbach’s alpha (Flake et al., 2017). The lack of concern for establishing reliable and valid measurements is a problem because the psychological sciences often deal with latent constructs that are not easily observed. If the scale itself is not measuring the underlying theoretical concept it intends to, researchers cannot make meaningful conclusions regarding the data. Therefore, it is vital to establish the link between the meaning and the measurement before implementing the measurement in a study (Barrett, 2005). It is also critical to note that there are many procedures to validate scales (for an overview, see Flake et al., 2017; Kimberlin & Winterstein, 2008; Markus & Borsboom, 2013). Keeping in mind the importance of scale validation and the vast array of procedures available, researchers should be methodical in their efforts to validate original scales. Thus, not all forms of scale validation can occur in a single study.

In Study 1, we established the reliability of the HATE-CARE scale. Reliability refers to the stability of the findings. If a measure is reliable, it should yield similar results over multiple trials (Altheide & Johnson, 1994). Reliability, specifically Cronbach’s alpha is commonly reported for author developed scales (Flake et al., 2017), likely due to ease of analysis and reporting. However, sound scales are not only stable, but they also measure what the researchers intend to measure (i.e., quality scales are valid; Kimberlin & Winterstein, 2008). Researchers can, and should, establish the construct validity of original scales. Construct validity is often established over multiple studies, and it refers to the extent to which a scale measures the
theoretical concept it was developed to measure (Hays & Reeve, 2008). A common approach used to help establish construct validity is conducting a confirmatory factor analysis (Byrne, 1998; DiStefano & Hess, 2005). The aim of Study 2 was to conduct a confirmatory factor analysis on a new sample to help (i) verify the factor structure and (ii) establish the construct validity of the scale developed in Study 1.

Method

Participants

A total of 223 University of Maine students participated in the study (see Table 1 in Appendix A for demographic information). All participants were at least 18 years of age and recruited via the SONA system. Per Comrey and Lee’s (2013) recommendation, a sample size of 200 is considered ‘fair’ for a factor analysis. Students were compensated with one research credit for their participation in the study.

Procedure

The procedure for Study 2 was identical to the one outlined in Study 1.

Results

Data Screening

The same data screening procedure used in Study 1 was applied. Before analysis, respondents were excluded if they were (i) identified as a straight-liner or (ii) if they failed an attention check embedded within the confrontation items. Additionally, because some fit statistics require that there is no missing data among participants, participants with missing values in the relevant observed variables (i.e., the 20 selected HATE-CARE items) were excluded from analyses. After excluding participants based on the aforementioned criteria, the total sample size was $N = 186$ ($N = 95$ in HATE condition, $N = 91$ in CARE condition). No
additional outliers were identified during data inspection. While Comrey and Lee (2013) recommend a minimum sample size of 200 to achieve ‘fair’ status, Thompson (2004) argues that in cases with 10 or more structure coefficients all around .40, a minimum sample size of 150 is required. The finalized HATE-CARE scale at the end of Study 1 was 20-items, with no factor loadings < .56. (see Table 5 in Chapter 2). Thus, a sample size of 186 is sufficient for a confirmatory factor analysis.

**Confirmatory Factor Analysis**

Using AMOS version 28, we conducted a confirmatory factory analysis to assess the model fit. Figure 5 shows the correlational paths for Model 1. Fit statistics for Model 1 can be found in Table 11. Per Kline’s (2015) recommendation, we report several fit statistics below.

**Figure 5**

*Confirmatory Factor Analysis Model 1*

Similar to the exploratory factor analysis, a nonsignificant chi-square statistic is desirable for a confirmatory factor analysis, as it suggests the data fits the model well (Stapleton, 1997).
The analysis yielded a significant result \( \chi^2(168) = 437.49, p < .001 \). While such a result might suggest the data does not fit the model well, scholars acknowledge that because the chi-square statistic is so sensitive to sample size. It is not always clear if statistical significance is indicative of poor model fit or simply a result of sample size (Stapleton, 1997; Tong & Bentler, 2013). Therefore, examination of additional fit statistics is required to make conclusions regarding the model fit.

Next, we examined the Comparative Fit Index (CFI). CFI is another statistical measure of model fit that adjusts for sample size and thus, does not have the same issues as the chi-square statistic (Rigdon, 1996). CFI values can range from zero to one, with larger values indicating a better fit. Previous standards defined a CFI value > .90 as indicative of an acceptable model fit. However, more recently, scholars have pushed to increase the CFI threshold to > .95 (Hu & Bentler, 1999). CFI for Model 1 is slightly below the .95 criteria for a good model fit with CFI = .912.

After analyzing CFI, we examined the root mean squared error of approximation (RMSEA). RMSEA is an absolute fit statistic where a value of zero indicates the best possible fit (Kline, 2015). A common cutoff criteria value scholars cite is RMSEA < .06 (Hooper et al., 2008; Hu & Bentler, 1999; Newsom, 2015). For Model 1, RMSEA = .093, 90% CI [.082-.104]. Because RMSEA > .06, we cannot confidently claim the model fits well. Lastly, we explored the standardized root mean squared residual (SRMR). An SRMR value equal to zero indicates a perfect fit, most scholars argue that SRMR < .08 is sufficient (Cho et al., 2020; Hooper et al., 2008; Hu & Bentler, 1999). For Model 1, SRMR = .0687. Therefore, according to the SRMR the model fits well.
Did Model 1 Fit the Data Well? While not all model fit indices suggested a good fit for Model 1, fit statistics like CFI and RMSEA were not far from achieving the desirable cutoff values. Model 1 even met the criteria for SRMR to be considered a good fit. Taken together, the fit of Model 1 can be improved, but it is not egregious. After examining the fit for Model 1, we assessed the modification indices and added covariance paths to the model to explore how doing so influenced model fit.

Modification Index Selection Criteria. Kline (2015) states that adding covariance paths to a model should improve the predicted overall fit of said model in cases where the modification index is large. Because we had no a priori assumptions regarding modification indexes and we did the following analysis solely for exploratory purposes, we arbitrarily chose a value. The modification index needed a minimum value of 10 for us to consider adding the covariance path. Additionally, paths can only be added among the error terms within the indicators of each latent variable (Kline, 2015). For instance, a path can exist between the error terms of two HATE items or the error terms of two CARE items, but a path cannot be added between an error term of a HATE item and an error term of a CARE item. Keeping the selection criteria in mind, we added eight paths to the model (see Figure 6).
Note that it is not a common practice to add as many paths as we chose. Adding covariance paths to the model may help the overall fit of the model, but it does so only for that specific data set, which in turn, limits the generalizability of the model (West et al., 2012). See Table 11 for Model 2 fit statistics.

Table 11

<table>
<thead>
<tr>
<th></th>
<th>Chi-square goodness-of-fit</th>
<th>RMSEA</th>
<th>90% CI</th>
<th>SRMR</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\chi^2$</td>
<td>$df$</td>
<td>$p$</td>
<td>CFI</td>
<td>RMSEA</td>
</tr>
<tr>
<td>Model 1</td>
<td>437.49</td>
<td>168</td>
<td>&lt; .001</td>
<td>.912</td>
<td>.093</td>
</tr>
<tr>
<td>Model 2</td>
<td>371.55</td>
<td>161</td>
<td>&lt; .001</td>
<td>.931</td>
<td>.084</td>
</tr>
</tbody>
</table>

Note. AIC = Akaike Information Criterion.

**Did Model 2 Fit the Data Well?** Similar to Model 1, the chi-square analysis was significant for Model 2 [$\chi^2(161) = 371.55, p < .001$], suggesting a poor fit. We followed up by exploring other fit indexes. CFI < .95, thus Model 2 did not quite achieve the threshold for a good fit. While RMSEA decreased in Model 2, RMSEA = .084, 90% CI [.073-.095], it was still
greater than the desired .06 value required for a good fit. However, Model 2 maintained an SRMR < .08, so according to SRMR, Model 2 fit well.

**Discussion**

While the data fit Model 2 slightly better than Model 1, the conclusions made regarding the fit statistics did not differ between the two models. In this case, we would recommend Model 1 because while adding covariance paths to models may improve the overall model fit, it limits the generalizability of the model by tailoring it too much to the data provided in a single sample (West et al., 2012). Additionally, when two models fit the data similarly, as is the case above, the parsimony principle states that simpler model is preferred (Kline, 2015).

In Model 1, we see HATE and CARE have a strong, negative correlation ($r = -.80$). Such a result lends support to the notion that HATE and CARE might be two endpoints of a single construct. However, we interpret this with caution and cannot make any definitive conclusions regarding the relationship between HATE and CARE because (i) we did not design Study 2 with this purpose of exploring the relationship between the subscales in mind, and (ii) Study 1 offered competing results. Additional research is required in order to determine whether or not HATE and CARE are reciprocals of each other.

Overall, though the fit of Model 1 does not meet the statistical criteria to be defined as good, the fit is not poor enough for us to discard the scale as a whole. It is possible that if we replicated the study with a new sample, the fit indices would improve. Furthermore, improving the scale may be a matter of making edits to individual items rather than constructing an entirely new model. For instance, looking at the correlation coefficients in Model 1, we see that the individual HATE items correlate well to the latent variable. Some of the CARE items (e.g., dignified, compassionate, genuine) do not correlate as strongly with the latent variable as other
items on that scale. In future research, we might consider editing the CARE scale, as items such as dignified, compassionate, and genuine, may not be capturing the underlying construct of CARE as well as we previously thought. Also note, HATE is better established within the confrontation literature, so it makes sense that we would be quicker to identify the individual components of that latent variable compared to CARE.

Recall the goal of Study 2 was to establish the construct validity of the HATE-CARE scale. Model 1 did not fit the data when examining the chi-square, CFI, and RMSEA fit indexes. However, according to the SRMR fit index, the model fit the data well. We feel that Study 2 has established some degree of construct validity for the novel scale, though future studies should work to revise the CARE scale to strengthen the overall scale. Because the fit of Model 1 was not egregious in Study 2, we felt it was appropriate to implement the scale developed in Study 1 as a measure in Study 3.
CHAPTER 4

STUDY 3: SELFF-AFFIRMATION AND CONFRONTATION

Introduction

According to the Confronting Prejudiced Responses (CPR) model outlined by Ashburn-Nardo and colleagues (2008), witnesses of prejudice face five general hurdles before ultimately deciding whether to confront the prejudice in some way (see Figure 1 in Chapter 1). One shortcoming of the model is that it only indicates whether an individual is likely to confront the prejudice and does not explore _how_ a person will confront (i.e., the model does not distinguish between confrontation approaches). We were interested in exploring how antecedents to a prejudiced encounter might influence one’s chosen confrontation approach. More specifically, we sought to investigate how self-affirming before witnessing sexism would influence women’s approaches to confrontation.

Recall from our discussion in Chapter 1 that self-affirmation can bolster oneself against threatening information, which, in turn, has the potential to reduce defensive responses. If self-affirming protects individuals from threatening information, then it is plausible the process can affect how individuals interpret various situations. For example, a self-affirmed woman might feel less inclined to appraise sexist comments as an emergency that needs to be addressed (Step 2 of the CPR model), which might further influence said women’s decision to confront the perpetrator as well as her chosen approach to confrontation. Note, Ashburn-Nardo and colleagues (2008) claim that (i) not all steps of the CPR model must be addressed for a confrontation to occur, and (ii) the steps are not required to occur sequentially. With this in mind, we have no reason to believe that if self-affirming does, in fact, alter the appraisal of the prejudiced event as an emergency, then it will automatically result in no confrontation every
time. So, we sought to investigate exactly how an antecedent like self-affirmation would influence confrontation.

Generally, the literature suggests that self-affirmation, as a construct, is relatively easy to manipulate. Steele (1988) argues that making values salient is sufficient to self-affirm individuals. Value scales, questionnaires (Napper et al., 2009; Sherman et al., 2000), and value essays (Armitage & Rowe, 2011; Crocker et al., 2008) are all valid approaches to manipulating self-affirmation. More recently, Zhu and Yzer (2019) developed the brief scale affirmation task (B-SAT), in which the researchers were successful at manipulating self-affirmation with an 11-item questionnaire. For ease of use in a virtual setting, the present study utilized the B-SAT.

Study 3 had two primary goals. First, we aimed to investigate how antecedents, such as self-affirmation, influence women’s confrontation behaviors and their approaches to confrontation following exposure to sexism. Second, we sought to implement the novel scale developed and established in Studies 1 and 2 to assess approaches to confrontation.

**Hypotheses**

**Hypothesis 1.** First, we hypothesized that confrontation behavior would differ by condition. We predicted that self-affirmed women would have lower rates of direct confrontation relative to women in the no affirmation condition.

**Hypothesis 2.** Next, we were interested in investigating confrontation approaches. We hypothesized that self-affirmation would influence women’s approaches to confrontation following exposure to sexism. More specifically, we predicted that self-affirmed women would use fewer HATE tactics (i.e., have lower scores on the HATE subscale) relative to women in the no self-affirm condition.
Hypothesis 3. In a similar vein, we suspected that self-affirmation would help to buffer against more defensive responses, resulting in women in the self-affirmation condition utilizing more CARE techniques (i.e., have higher scores on the CARE subscale) than women in the no affirmation condition.

Hypothesis 4. Next, we were interested in exploring if self-affirmation affects women’s perceptions of sexism as an emergency that needs to be addressed (Step 2 of the CPR model; Ashburn-Nardo et al., 2008), which in turn, might affect feelings regarding their personal responsibility to intervene (Step 3 of CPR model). More specifically, we thought that if self-affirmation buffers the women against threats, they may not appraise the sexism as an emergency and feel less personal responsibility to confront the sexism. Therefore, we predicted women in the no affirmation condition would have higher perceived responsibility to confront compared to women in the self-affirmation condition.

Hypothesis 5. Lastly, we wanted to gain some insight regarding the role of state affect and obtain results that could help guide future research. Since it is thought that self-affirming can buffer against threatening comments, we hypothesized women in the self-affirmation condition would have lower scores on the negative state affect subscale of the PANAS-X after exposure to sexism relative to women in the no affirmation condition.

Method

Participants

A total of 109 women from the University of Maine participated in the study (see Table 12 in Appendix F for demographic information). All participants were at least 18 years of age and recruited via the SONA system. Students were compensated with one research credit for
their participation in the study. Only women were sampled because we were interested in investigating the effects of self-affirmation on women’s confrontation responses to sexism.

**Procedure**

The entire study was conducted virtually using the Qualtrics survey system. After providing informed consent, participants were randomly assigned to one of two self-affirmation conditions. We manipulated self-affirmation with an edited version\(^2\) of the brief scale affirmation task (B-SAT; Zhu & Yzer, 2019). In the self-affirmation condition, participants rated agreement with statements regarding the self on a 0 (Strongly Disagree) to 6 (Strongly Agree) scale (Appendix G). In the no affirmation condition, participants completed the same questionnaire, but instead of answering questions about themselves, participants answered the questions about a celebrity (e.g., David Beckham) on a 0 (Strongly Disagree) to 6 (Strongly Agree) scale (Appendix H).

Following the manipulation, we presented participants with an imagined scenario. The scenario, adapted from Swim and colleagues (2003), is as follows. We asked participants to imagine they were at lunch with a group of coworkers when one of the male coworkers asked the other men if they were interested in joining a basketball team. Participants then learned that their male colleague did not want women to join the team because he assumed women would be more interested in the color of the uniforms than how the game was played. Immediately after reading the scenario, we instructed participants to imagine joining the conversation and write their imagined response in an open-ended item. The response page required participants to remain on the page for a minimum of five minutes before continuing with the survey to encourage longer, 

\(^2\) Note the original brief scale affirmation task by Zhu and Yzer (2019) is 11-items. One item: “I must stand up for what I believe in, even in the face of strong oppositions” was omitted to avoid potentially priming participants to confront sexism, seeing as we designed this study to explore the influence of self-affirmation on confrontation following exposure to sexism.
more detailed responses. Participants were allowed to go over the five-minute period if they needed to. After responding to the imagined scenario, participants completed a series of self-report surveys and were debriefed. See Figure 7 for a pictural representation of the procedure.

**Figure 7**

*Study 3 Procedure*

![Study 3 Procedure Diagram](image)

**Measures**

**Self-Affirmation Manipulation Checks.** Self-appraisal was measured as a manipulation check (Napper et al., 2009; Zhu & Yzer, 2019). We instructed participants to think back to the task they completed before the imagined scenario and respond to four statements on a bipolar scale ranging from -3 to 3. Items included, “The task made me think about... negative aspects of myself/positive aspects of myself; things that are not important to me/things that are important to me; things I don’t like about myself/ things I like about myself; things I don’t value about myself/things I value about myself”.

Value awareness was also measured and used as a manipulation check (Napper et al., 2009; Zhu & Yzer, 2019). Participants responded to two items on a 0 (Strongly Disagree) to 6 (Strongly Agree) scale. Items included “The task made me aware of who I am” and “The task made me aware of my values.” The two manipulation checks were counterbalanced to ensure that responses on one measure did not influence the responses on the other.

**Perceived Sexism.** Five items using a 0 (Strongly Disagree) to 6 (Strongly Agree) scale measured perceived sexism (Bastart et al., 2021). Two items assessed the perpetrator as sexist
and prejudiced (e.g., “[the perpetrator is] a prejudiced person/ a sexist person”). Two items assessed the perpetrator’s behavior as sexist and prejudiced (e.g., “[the perpetrator’s behavior is] sexist/ prejudiced”). The final item assessed the perception of the whole event (e.g., “the outcome was due to discrimination”).

**Perceived Responsibility to Confront.** Perception of personal responsibility to confront was measured with a single item (e.g., “How much responsibility do you feel to confront these comments when they occur?”) on a 0 (*None*) to 6 (*A lot*) scale.

**State Affect.** Participants completed the Positive and Negative Affect Expanded Form Scale (PANAS-X; Watson & Clark, 1999). For the purposes of this study, we modified the language to ask about state affect (Appendix I).

**Self-Evaluation of Confrontation Approach.** Using the 20-item scale developed in Study 1, participants coded their perceptions of their response to the imagined scenario on a 1 (*Strongly Disagree*) to 9 (*Strongly Agree*) scale.

**Confrontation Behavior.** Trained coders, blind to the experimental condition, used the written responses to assess whether participants directly, indirectly, or did not confront the perpetrator. For the purposes of the present study, direct confrontation was defined as verbally addressing the male colleague on the basis that his argument is sexist or occurs due to the candidate’s gender (Helwig et al., 2022). Indirect confrontation was defined as addressing the event as problematic or unfair, but not explicitly addressing sexism as the cause of the event (Brett et al., 2014). In this study, no confrontation occurred when the participant did not anticipate saying or doing anything to address the sexism.

**Confrontation Approach.** Trained coders, blind to the experimental condition, assessed the written responses using the HATE-CARE scale developed in Study 1. In doing so, we were
able to give each participant a HATE composite score and a CARE composite score to help identify how participants who chose to address the sexism approached the confrontation.

Results

Data Screening

Before analysis, inattentive participants (i.e., if they failed an attention check embedded within the survey items) were excluded. After excluding inattentive respondents, total sample size was $N = 96$ ($N = 53$ in the affirmation condition, $N = 43$ in the no affirmation condition).

Manipulation Checks

Self-Appraisal. A composite self-appraisal score was created by averaging the participants’ responses to the four appraisal items. After creating the composite, we conducted an independent samples t-test to compare mean self-appraisal ratings between conditions. The test was not significant, $t(94) = .082, p = .935, d = .017$, small. Participants in the affirmation condition ($M = 1.06, SD = 1.22$) did not significantly differ in their self-appraisal scores compared to participants in the no affirmation condition ($M = 1.08, SD = 1.16$).

Value Awareness. Like self-appraisal, a composite value awareness score was created by averaging the participants’ responses to the two value awareness items. We ran another independent samples t-test to explore the differences between conditions in value awareness. Again, the test was not significant, $t(94) = -0.625, p = .534, d = .128$, small. Participants in the affirmation condition ($M = 3.75, SD = 1.31$) did not have significantly higher value awareness scores relative to the no affirmation condition ($M = 3.56, SD = 1.77$).

Did the Self-Affirmation Manipulation Work? Results from the manipulation checks suggest that the self-affirmation manipulation did not work as intended. However, it is important to note that the study was substantially underpowered. Prior to data collection, we conducted a
sensitivity analysis for an independent samples t-test using G*Power (Faul et al., 2007) to estimate how large a sample was required to detect a small effect size ($d = .35$, see Epton et al., 2015 and Sweeney & Moyer, 2015 meta-analyses for discussion of effect size variability in self-affirmation manipulations). After inputting the following parameters: $\alpha$ (two-sided) = .05, $d = .35$, power = .80, allocation ration N2/N1 = 1, G*Power suggested a total sample size of 260 with 130 participants in each condition. Thus, it is possible that the affirmation manipulation worked, but we are simply underpowered and unable to detect the effect. We continued analyses as planned to explore the utility of the HATE-CARE measure, but results concerning affirmation condition differences should be interpreted with caution.

**Statistical Analyses**

**Rate of Confrontation Behavior.** Three trained coders classified participant responses as either direct, indirect, or no confrontation. To assess interrater reliability, we used Cohen’s kappa, as it is an appropriate reliability measure for nominal variables (Hallgren, 2012). Because Cohen’s kappa only makes comparisons between two raters at a time, we ran every possible combination, and the two raters with the strongest kappa value were used ($\kappa = .773, p < .007$). Landis and Koch (1977) classify kappa values ranging from 0.61 to 0.80 as having ‘substantial agreement’. In cases where the raters did not agree, we referred to the third coder’s response for a tiebreaker. Confrontation rates can be found in Table 13.

**Table 13**

*Confrontation Behavior Rates*

<table>
<thead>
<tr>
<th>Behavior</th>
<th>No Affirmation</th>
<th>Affirmation</th>
<th>Collapsed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Percent</td>
<td>n</td>
<td>Percent</td>
</tr>
<tr>
<td>Confronted</td>
<td>93.1</td>
<td>40</td>
<td>78.8</td>
</tr>
<tr>
<td>Direct</td>
<td>79.1</td>
<td>34</td>
<td>57.6</td>
</tr>
<tr>
<td>Indirect</td>
<td>14.0</td>
<td>6</td>
<td>21.2</td>
</tr>
<tr>
<td>No Confrontation</td>
<td>6.9</td>
<td>3</td>
<td>21.2</td>
</tr>
</tbody>
</table>
Creating HATE and CARE Composite Variables. The scale created in Study 1 was implemented in Study 3. Trained research assistants read the participants’ imagined responses and rated items using the same 1 (Strongly Disagree) to 9 (Strongly Agree) Likert-type scale used in Study 1. After responding to the individual items, a composite score was created for HATE by averaging across the 10-items. A HATE composite score was created for each research assistant. Next, we calculated interrater reliability to determine the extent to which the coders agreed on the confrontation approach. Reliability among the coders was strong ($\alpha = .842$). A value of .80 is often considered the threshold for acceptability (Lombard et al., 2002; Neuendorf, 2002). Once we determined that interrater reliability was acceptable, we created the HATE composite by averaging HATE scores across coders.

The same procedure outlined above was used to create the CARE composite. Again, reliability among coders reached an acceptable level ($\alpha = .806$). The CARE composite was then created by averaging CARE scores across the coders.

Did Confrontation Behavior Differ by Affirmation Condition? Because confrontation behavior is a categorical variable in this study, we conducted a 2 condition (no affirmation vs. affirmation) x 3 confrontation behavior (direct vs. indirect vs. no confrontation) chi-square test of association to explore if confrontation behaviors differed by condition. The results of the omnibus test were significant, $\chi^2(2, N = 95) = 6.171, p = .046$, suggesting that confrontation behavior differed by affirmation condition.

While we deciphered between direct and indirect confrontation during coding, we were more interested in direct confrontation behaviors. To further explore the results of the omnibus test, we followed up with a 2 condition (no affirmation vs. affirmation) x 2 confrontation behavior (direct vs. no confrontation) chi-square test of association. The analysis yielded
significant results, \( \chi^2(1, N = 77) = 4.859, p = .028. \) Considering the rates of conformation (Table 13), it appears that women in the self-affirmation condition directly confronted the perpetrator less than the participants in the no affirmation condition, which is consistent with Hypothesis 1 (see Figure 8).

**Figure 8**

*Confrontation Behavior Split by Condition*

After the chi-squared analyses revealed significant results, we conducted a series of Kruskal-Wallis K independent samples analyses to explore how HATE and CARE scores differ by confrontation behavior (i.e., direct, indirect, no confrontation). Rather than conducting a one-way ANOVA, we opted for its nonparametric analog, as it is more appropriate in cases where sample size differs substantially by group. Results of the omnibus analysis were significant, \( H(2) = 20.557, p < .001, \) suggesting that HATE scores did differ by confrontation behavior. We followed up the omnibus analysis with a series of Mann-Whitney tests (i.e., the nonparametric
analog of an independent samples t-test) to help us identify which groups had significantly different HATE scores (see Table 14).

Table 14

<table>
<thead>
<tr>
<th>Confrontation Behavior</th>
<th>n</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct</td>
<td>63</td>
<td>3.44</td>
<td>1.25</td>
</tr>
<tr>
<td>Indirect</td>
<td>18</td>
<td>2.92</td>
<td>0.91</td>
</tr>
<tr>
<td>No Confrontation</td>
<td>14</td>
<td>1.91</td>
<td>0.75</td>
</tr>
</tbody>
</table>

Post hoc analyses\(^3\) revealed direct confronters had significantly higher HATE scores compared to women who did not confront, \(U(N_{\text{direct}} = 63, N_{\text{no confrontation}} = 14) = 108, z = -4.399, p < .001\). Similarly, indirect confronters had significantly higher HATE scores than women who did not confront \(U(N_{\text{indirect}} = 18, N_{\text{no confrontation}} = 14) = 50, z = -2.888, p = .004\). Lastly, the results revealed no significant difference in HATE scores between direct and indirect confronters, \(U(N_{\text{direct}} = 63, N_{\text{indirect}} = 18) = 432, z = -1.534, p = .125\) (see Figure 9).

\(^3\) Bonferroni correction applied to account for inflated Type I error. Comparing to \(\alpha = .017\).
Next, we repeated the statistical procedure outlined above to explore differences in CARE scores by confrontation behavior (see Table 15).

**Table 15**

*CARE Scores Split by Confrontation Behavior*

<table>
<thead>
<tr>
<th>Confrontation Behavior</th>
<th>n</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct</td>
<td>63</td>
<td>5.09</td>
<td>1.15</td>
</tr>
<tr>
<td>Indirect</td>
<td>18</td>
<td>4.91</td>
<td>1.23</td>
</tr>
<tr>
<td>No Confrontation</td>
<td>14</td>
<td>4.98</td>
<td>0.99</td>
</tr>
</tbody>
</table>

Results from the omnibus Kruskal-Wallis test were not significant, $H(2) = 0.714, p = .70$. It appears that CARE scores did not meaningfully differ by confrontation behavior in any way (see Figure 10).
Did Confrontation Approaches Differ by Condition? To determine if confrontation approach differed by affirmation condition, we conducted a series of independent samples t-tests. Descriptive statistics can be found in Table 16. In addition to inattentive participants, we excluded participants who did not confront the perpetrator from the following analyses.

**Table 16**

*Descriptive Statistics for Confrontation Approach Split by Condition*

<table>
<thead>
<tr>
<th>Condition</th>
<th>n</th>
<th>M</th>
<th>SD</th>
<th>Range</th>
<th>M</th>
<th>SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Affirmation</td>
<td>40</td>
<td>3.12</td>
<td>1.12</td>
<td>1.33-7.5</td>
<td>5.29</td>
<td>0.95</td>
<td>2.33-7.57</td>
</tr>
<tr>
<td>Affirmation</td>
<td>41</td>
<td>3.54</td>
<td>1.26</td>
<td>1.03-8.1</td>
<td>4.82</td>
<td>1.32</td>
<td>1.57-7.37</td>
</tr>
</tbody>
</table>

First, we compared HATE scores between conditions. Keep in mind that because sample size is low, results should be interpreted with caution. The results obtained from the independent samples t-test were not significant, $t(79) = -1.595, p = .115, d = .35$, small (see Figure 11).
HATE approaches to confrontation did not differ by affirmation condition. Therefore, the results did not support our second hypothesis.

**Figure 11**

*Mean HATE Scores Split by Condition*

![Graph showing mean HATE scores split by condition](image)

*Note.* Error bars indicate +/- standard error mean.

Next, we explored whether CARE scores differed by condition. The independent samples t-test did not yield significant results\(^4\), \(t(72.910) = 2.266, p = .070, d = .408\), small. Participants’ CARE scores did not significantly differ by condition (see Figure 12). Furthermore, the results do not support our third hypothesis that self-affirmed women would implement more CARE tactics in their confrontations than women who did not self-affirm.

---

\(^4\) Failed to meet homogeneity of variance assumption \((p = .016)\). Reported statistics are adjusted for equal variances not assumed.
Did Perceived Sexism Differ by Condition? Before examining if women’s perceived responsibility to confront differed by affirmation condition, we wanted to explore if perceived sexism differed by condition. We had no *a priori* hypotheses regarding this question. We chose to include the analysis to offer some additional insight into our analysis of perceived responsibility to confront. We speculated that perceived responsibility to confront would differ by condition because women in the affirmation condition would appraise the sexism as less threatening and, therefore, feel less inclined to address it. However, we were not convinced that self-affirmed women would perceive less sexism than women in the no affirmation condition. It is plausible that women, regardless of affirmation status, perceive the sexism, but self-affirmed women simply appraise the threat of the sexism differently.

*Note.* Error bars indicate +/- standard error mean.
Perceived sexism of the perpetrator was calculated by averaging two items ($\alpha = .833$). The independent samples t-test was not significant $t(94) = 1.041, p = .30, d = .214$, small. Women in the no affirmation condition did not perceive the perpetrator as any more sexist than women in the affirmation condition.

Perceived sexism of the perpetrator’s behavior was calculated by averaging two items ($\alpha = .841$). The independent samples t-test revealed no significant results, $t(94) = 0.783, p = .436, d = .161$, small. It appears that perceived sexism of the perpetrator’s behavior did not differ by condition. Lastly, we examined perceived sexism of the event as a whole with a single item. The results of the independent samples t-test\(^5\) were not significant, $t(91.891) = 1.203, p = .232, d = .238$, small. Furthermore, women in the no affirmation condition did not perceive the event in its entirety as any more sexist than women in the affirmation condition. Deceptive statistics for the above analyses can be found in Table 17.

**Table 17**

*Descriptive Statistics for Perceived Sexism Split by Condition*

<table>
<thead>
<tr>
<th>Condition</th>
<th>n</th>
<th>Perpetrator</th>
<th>Perpetrator’s Behavior</th>
<th>Overall Event</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$M$</td>
<td>$SD$</td>
<td>$M$</td>
</tr>
<tr>
<td>No Affirmation</td>
<td>43</td>
<td>4.56</td>
<td>1.36</td>
<td>4.97</td>
</tr>
<tr>
<td>Affirmation</td>
<td>53</td>
<td>4.25</td>
<td>1.46</td>
<td>4.76</td>
</tr>
</tbody>
</table>

Recall that perceptions of sexism were measured using a 0 (*Strongly Disagree*) to 6 (*Strongly Agree*) scale. Thus, a value of 3 is the midpoint. Regardless of condition, women rated the perpetrator himself, his behavior, and the overall event as sexist above the midpoint.

\(^5\) Failed to meet homogeneity of variance assumption ($p = .014$). Reported statistics are adjusted for equal variances not assumed.
After exploring differences between the affirmation conditions, we conducted a series of correlations, collapsed across the affirmation conditions, to explore how perceptions of sexism relate to HATE and CARE scores (see Table 18).

Table 18

Relationship Between Perceived Sexism and HATE-CARE Scores

<table>
<thead>
<tr>
<th>Subscale</th>
<th>n</th>
<th>Perpetrator</th>
<th>Perpetrator’s Behavior</th>
<th>Overall Event</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>r</td>
<td>p</td>
<td>r</td>
</tr>
<tr>
<td>HATE</td>
<td>95</td>
<td>.213</td>
<td>.038</td>
<td>.303</td>
</tr>
<tr>
<td>CARE</td>
<td>95</td>
<td>-.027</td>
<td>.798</td>
<td>-.092</td>
</tr>
</tbody>
</table>

Looking at the general trends of the correlations, it appears that higher perceived sexism in all domains (i.e., perception of the perpetrator as sexist, perception of the perpetrator’s behavior as sexist, and perception of the overall event as sexist) is related to higher scores on the HATE subscale. While the magnitude of the correlations are weak (|.10| < r’s < |.40|, Schober et al., 2018), they are all significant (all p values < .05). Conversely, perceived sexism does not appear to be meaningfully correlated with scores on the CARE subscale.

Does Perceived Responsibility to Confront Differ by Affirmation Condition? After establishing that women, regardless of condition, perceived the events in the imagined scenario as sexist, we were interested in exploring if self-affirmation influences personal perceived responsibility to confront in any way. To determine if there were differences in women’s perceived responsibility to confront, we conducted an independent samples t-test. The results were not significant, t(94) = 1.968, p = .052, d = .404, small. On average, self-affirmed women (M = 4.02, SD = 1.42) did not perceive significantly less responsibility to confront the sexism than women who were not self-affirmed (M = 4.56, SD = 1.22). Therefore, the results did not support our fourth hypothesis (see Figure 13). Note that the effect size is on the upper end of small, approaching medium (Cohen, 1988). Therefore, it is possible that self-affirmation does
affect women’s perceived responsibility to confront, but the analysis was underpowered due to the low sample size which may have resulted in a Type I error (i.e., failing to detect an effect when one exists).

**Figure 13**

*Perceived Responsibility to Confront Split by Condition*

<table>
<thead>
<tr>
<th></th>
<th>No Affirmation</th>
<th>Affirmation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceived Responsibility to Confront (± SEM)</td>
<td>4.5 ± 0.2</td>
<td>4.0 ± 0.3</td>
</tr>
</tbody>
</table>

Note. Error bars indicate +/- standard error mean.

After exploring differences between the affirmation conditions, we conducted a series of correlations, collapsed across experimental condition to explore how perceived responsibility to confront relates to the HATE and CARE scores. The Pearson correlation revealed no significant relationship between perceived responsibility to confront and scores on the HATE subscale ($r = .152, p = .140$). Results were similar for the correlation between perceived responsibility to confront and scores on the CARE subscale ($r = .159, p = .124$). Thus, it does not appear there is any predictive utility between the HATE-CARE score and perceived responsibility to confront.

**Did Negative State Affect Differ by Condition?** Initial visualization of the data suggested the distribution of negative affect scores, a sum composite score calculated per Watson
and Clark’s (1999) guidelines, is slightly positively skewed. Skewness = .620 for the no affirmation condition, and skewness = .601 for the affirmation condition. Bryne (2010) and Hair et al. (2010) claim that so long as skewness falls within the -2 to 2 range, the data can be considered normal. As a result, parametric analyses proceeded as intended. We conducted an independent samples t-test to determine if there were any differences in negative state affect between conditions. While participants in the no affirmation condition ($M = 12.42, SD = 10.13$) did have a higher mean negative state affect score compared to participants in the affirmation condition ($M = 11.34, SD = 9.26$), the difference was not significant $t(94) = 0.544, p = .294, d = .112$, small (see Figure 14). Furthermore, the data did not support our final hypothesis.

**Figure 14**

*Mean Negative State Affect Split by Condition*

![Mean Negative State Affect Split by Condition](image)

*Note.* Error bars indicate +/- standard error mean.

To gain additional insight into the utility of the HATE-CARE scale, we conducted correlational analyses between the two subscales and negative state affect. The relationship between HATE scores and negative state affect was not significant ($r = -.012, p = .908$). Results
were similar for scores on the CARE subscale ($r = -.011, p = .916$). Moreover, it seems that state affect and HATE-CARE scores are not meaningfully related.

**Do Participants Perceive Their Approach to Confrontation Differently from Coders?** We were also interested in exploring whether participants’ perceptions of their confrontation approach differed from the coders’ perceptions of the confrontation. We had no *a priori* hypotheses regarding this question and chose to include it solely for exploratory purposes. To examine differences in HATE scores between participants (i.e., confronters) and coders, we conducted a dependent samples t-test. Again, only participants who confronted the sexism were included in the analyses. Additionally, only about half the participants rated their confrontation response using the HATE-CARE scale. As a result, the sample size is low, and the results should be interpreted with caution. To avoid further reduction of the sample size, we opted to explore differences in perceived confrontation approach collapsed across experimental conditions.

Analysis revealed no significant differences in HATE scores between participants ($M = 3.56, SD = 1.79$) and coders ($M = 3.44, SD = 1.23$), $t(39) = 0.460, p = .648, d = .073$, small (see Figure 15). Coders did not perceive confrontation responses as any more HATE-filled than the confronter or vice versa.
Furthermore, a correlational analysis confirmed that there was a significant, moderate correlation between confronters’ self-evaluative HATE scores and coders’ HATE scores ($r = .451$, $p = .003$).

The same analytical procedure outlined above was used to compare CARE scores. Again, only participants who confronted the sexism were included in the analysis. The results obtained from the dependent samples t-test were not significant, $t(39) = 0.283$, $p = .778$, $d = .045$, small. The confronters ($M = 5.11$, $SD = 1.72$) did not perceive their responses as any more CARE-filled than the coders ($M = 5.04$, $SD = 1.21$; see Figure 16).
Figure 16

Mean CARE Score Split by Perceiver

Note. Error bars indicate +/- standard error mean.

Again, we ran a Pearson’s correlation to further explore the relationship between confronters’ self-evaluative CARE scores and coders’ CARE scores. The analysis yielded significant results ($r = .448, p = .004$). There was a moderate, positive correlation between confronters’ self-evaluative CARE scores and coders’ CARE scores.

Discussion

Overall, the results from Study 3 failed to support the hypotheses, apart from our first hypothesis that women in the self-affirmation condition would have lower rates of direct confrontation compared to women in the no affirmation condition. Support for the first hypothesis suggests that self-affirmation may buffer individuals against threatening information. Buffering the threat of sexism, might, in turn, affect women’s appraisal of the situation (Step 2 of the CPR model), and further influence their decision to confront (Step 3 of the CPR model).
However, we caution making such a conclusion as manipulation checks revealed a failure to effectively manipulate self-affirmation.

Analysis of confrontation behaviors also revealed that the paradigm, as it presently exists, encourages confronting. Regardless of condition, 85.2% of women confronted the sexism in this study. This value is higher than actual confrontation behaviors in previous studies (see Study 1 in Rattan & Dweck, 2010; Study 1 in Swim & Hyers, 1999) and consistent with anticipated confrontation behaviors in the confrontation literature (see Study 2 in Swim & Hyers, 1999).

It is plausible that failure to support our other hypotheses is a result of ineffectively manipulating self-affirmation, rather than self-affirmation not influencing the dependent variables of interest. We chose to implement the B-SAT for two reasons. First, surveys are easily implemented manipulations when conducting online studies, so we felt it would serve us well in Study 3. Second, because the B-SAT has a complimentary survey for the no affirmation condition, it allowed us to maintain consistency of the manipulation across conditions. We would not have been able to achieve said consistency had we used a survey for the affirmation condition and a journey essay for the comparison condition, a procedure that several self-affirmation manipulations studies cite using (for a review see Napper et al., 2009; Zhu & Yzer, 2019).

It is unclear if the self-affirmation manipulation failed as a result of the low sample size or if the no affirmation condition failed to truly be a no affirmation condition. Asking participants to evaluate the values of a celebrity such as David Beckham may unintentionally affirm the responding participant in cases where (i) she identifies in some way with the celebrity of interest (e.g., maybe she is a soccer player like David Beckham), or (ii) she reflects on her personal values when answering questions regarding someone else’s values (Zhu & Yzer, 2019).
It seems that comparison condition is a more appropriate title than no affirmation condition with the B-SAT manipulation. In future studies, researchers should consider either implementing a different self-affirmation manipulation or increasing the sample size to ensure sufficient power is reached to detect any notable effects.

Furthermore, while Study 3 failed to detect any differences between the self-affirmation conditions, the research was still beneficial. In Study 3, we implemented the HATE-CARE scale developed in Studies 1 and 2, illustrated the utility of the novel scale in an experimental setting, and gained insight regarding the appropriateness of the use of the scale in the paradigm outlined above. In relation to the predictive utility of the scale, the analyses suggest that HATE scores were significantly correlated with perceptions of sexism in multiple domains. Future researchers should consider exploring this relationship further. Researchers should also consider exploring the relationship of HATE-CARE scores with other variables. Additionally, the results revealed no significant differences between coders’ perceived HATE and CARE scores and confronters’ perceived HATE and CARE scores. These results lend support to the idea that the scale is reliable, regardless if a person is using it to assess their personal confrontation approach or someone else’s confrontation approach.

With respect to the appropriateness of the use of the scale in the paradigm outlined in Study 3, we noticed that while distribution of HATE and CARE scores were roughly normally distributed, the ranges were limited. We believe this might be due to short confrontation responses, making it difficult for coders to assess said responses using the novel scale. We required participants to remain on the response page for a minimum of five minutes, though they were allowed to go over if they wished. While some written responses were quite thorough, others were not as detailed. It is possible that implementation of the HATE-CARE scale is more
appropriate when responses are longer. Additionally, while we instructed participants to explain how they would respond in the imagined scenario, many detailed their feelings about the interaction and failed to explain their anticipated actions or verbal responses. It might be the case that the HATE-CARE scale is better equipped to assess actual confrontation responses and in turn, is better implemented in studies that use more ecologically valid paradigms.

Regardless of our shortcomings with manipulating self-affirmation, Study 3 offered valuable information concerning the utility of the HATE-CARE scale and helped to further establish the reliability of the scale.
CHAPTER 5
CONCLUSIONS

Although we conducted three independent studies, it is critical to consider how the results of the individual studies inform our general understanding of confrontation processes. To summarize, we developed a novel scale to assess confrontation approaches (Study 1 and Study 2). We then implemented that scale in an experiment to aid in an investigation of the effects of self-affirmation on women’s confrontation responses following exposure to sexism (Study 3). Study 1 offered an opportunity to establish internal consistency of the novel scale and Study 2 confirmed the scale, as it presently exists, has some degree of construct validity. However, the weak model fit in Study 2 suggests that we can improve upon the novel HATE-CARE scale in the future. More specifically, interpretation of the correlation coefficients suggests that select items on the CARE subscale might not fully capture the latent construct of CARE and can be revised. Study 3 provided a chance to implement the novel scale and explore the utility of the scale in an experimental setting. While the self-affirmation manipulation was unsuccessful in Study 3, the study revealed potential predictive utility of the HATE-CARE scale, further established the reliability of the scale, and illustrated the importance of nonconfirmatory research as a vehicle to strengthen confirmatory research in psychology.

Strengths

While the research was unsuccessful in supporting all the proposed hypothesis, there are still many strengths to consider. A primary success of the research is it illustrates how critical nonconfirmatory research is in psychology. Critics of hypothesis testing often argue that scholars are too focused on confirmatory analyses, and that more emphasis should be placed on nonconfirmatory research to better inform researchers of the derivation chain between theory and
The idea is that conducting more exploratory research can ultimately strengthen confirmatory analyses because nonconfirmatory research helps to better establish measures, causal relationships, boundary conditions, and auxiliary assumptions (Scheel et al., 2021). We feel the three studies described above illustrate this notion well. Study 1 was entirely nonconfirmatory, yet necessary to help us measure the latent construct of interest in Study 3. While Study 2 was literally a confirmatory factory analysis, the only goal was to validate a novel scale beyond simple reliability, which is often the only extent to which author-developed scales are validated (Flake et al., 2017). In Study 3, we see how critical it is to establish the derivation chain before attempting a confirmatory hypothesis test. While we were diligent in our efforts to establish our scale, we did not take the same care to conduct our own nonconfirmatory research regarding the self-affirmation manipulation. As a result, the analyses suggest we failed to successfully manipulate self-affirmation in Study 3 and any conclusions made in that study regarding the effects of self-affirmation should be interpreted with caution.

In working to validate an author-developed scale beyond the rudimentary reliability measure, we help to establish a standard in the field of psychology, one that holds the quality of research in high regard and that the field seems to have deviated from over the years (Flake et al., 2017; Nunnally, 1975). Null hypothesis significance testing is alluring because significant results are flashy, offer notoriety, and are often published in the field. However, statisticians argue that significant p-values are not as informative as one would hope (Loftus, 1996; Wagenmakers, 2007). Jumping from idea to hypothesis test without taking care to establish the derivation chain will ultimately result in weak theoretical advances. With this in mind, we view our efforts to
establish the validity of a new measure as a step in the direction to produce meaningful research of the highest quality.

As previously mentioned, researchers are beginning to develop measures to assess people’s confrontation responses to prejudice (see the PCS scale by Chaney & Sanchez, 2021). The HATE-CARE scale is unique in that it is the first measure to distinguish between the implementation of tactics that confrontation scholars recommend confronters of prejudice employ and avoid. As it currently exists, there is opportunity to improve the scale. Examination of the correlation coefficients of CARE items in Model 1 in Study 2 suggests that we can improve upon the CARE subscale in the future. We expected the CARE subscale to be weaker than the HATE subscale, given that CARE is not as well established in the confrontation literature as HATE. We anticipated that revisions would be necessary to strengthen the HATE-CARE scale but feel that our research is not without merit, as it is an excellent initial attempt at scale development and validation.

**Limitations**

As with all research, the research presented was not without its limitations. Diversity of the sample was a significant limitation across all three studies. While Study 1 offered some insight regarding the generalizability of the HATE-CARE scale across men and women, we have no information concerning how those who do not identify with the gender binary perceive HATE and CARE responses. We have even less information with respect to race. The intersection of race and gender creates unique experiences for individuals (Settles et al., 2008). Those unique experiences have the potential to alter how an individual perceives different confrontation responses. Thus, future researchers should consider expanding the diversity of the sample to better establish the generalizability of the HATE-CARE scale.
The three studies also failed to explore the nuances of confrontation efficacy (i.e., productivity). The present research defined productivity in terms of whether confrontation researchers recommend confronteders implement or avoid particular techniques. Currently, in the confrontation literature, there are two distinct ways of defining confrontation efficacy. First, effective confrontations can be defined in terms of the perpetrator’s response. A confrontation is considered effective or successful if the perpetrator alters their attitudes or behaviors (Brett et al., 2014; Czopp et al., 2006; Drury & Kaiser, 2014). Alternatively, efficacy can be defined in relation to the confronter’s personal goals of the interaction. In this case, a confrontation might be considered effective if the confronter experiences reduced feelings of anger, gains a sense of closure, or feels as if personal goals were achieved (Becker & Barreto, 2019; Haslett & Lipman, 1997; Hyers, 2007). In Study 1, we determined that participants considered CARE confrontation responses to be more effective than HATE responses. However, we did not distinguish between efficacy type. There is still much work to do to help elucidate how people perceive the success, efficacy, productivity, what have you, of the different approaches to confrontation.

Study 3 also helped identify some potential limitations of the novel HATE-CARE scale. First, it may be better to implement the scale in cases where confrontation responses are longer. Second, it may also be better to utilize the scale when measuring actual confrontations as opposed to anticipated confrontations. The limited range of HATE-CARE scores in Study 3 and feedback obtained from coders highlighted the difficulty of using the scale on short confrontation responses. The results from Study 3 did confirm that confronters and coders perceived confrontation responses similarly, so we do not necessarily think it is inappropriate to use the HATE-CARE scale in cases where confrontation responses are short. Regardless, we acknowledge that the scale is potentially more accurate with lengthier responses. It is also
possible that the relationship of the HATE and CARE scores between confronter and coder would improve with longer responses, though this is entirely speculative, and additional research is required to make such a claim. Furthermore, coders reported that some participant responses lacked clarity. In cases where participants also detailed their feelings of the event, clarity of anticipated confrontation response was sometimes murky, making it difficult to accurately assess confrontations with the HATE-CARE scale. With these limitations in mind, future implementation of the HATE-CARE scale may be better in more ecologically valid experiments that encourage actual confronting behavior as opposed anticipated behavior.

**Recommendations for Future Research**

First, we suggest researchers continue to work to revise the HATE-CARE scale. The weak fit of the model in Study 2 suggests there is opportunity for improvement regarding the current scale, especially the CARE subscale, as illustrated above. Not only should researchers revise the scale, but they must also work to validate the scale. Scale validation is an ongoing process. Establishing only one type of validity or reliability of the scale and claiming it as a sufficient validation practice weakens the measure and threatens research findings in studies where the scale is used, especially in the psychological sciences, where most constructs of interest are latent. Therefore, future research should work to replicate validity findings and also implement different validation procedures (e.g., test-retest reliability, Kimberlin & Winterstein, 2008).

In addition to revising the existing HATE-CARE scale, we recommend replicating Study 3. Either an exact replication with a larger sample or a similar design with a different, possibly more effective, self-affirmation manipulation. The success of self-affirmation manipulations varies, likely because of assorted effect sizes (Epton et al., 2015; McQueen & Klein, 2006;
Sweeney & Moyer, 2015). Furthermore, future researchers may benefit from conducting more nonconfirmatory research on self-affirmation manipulation procedures to better establish the derivation chain for self-affirmation before implementing it as an independent variable in a confrontation setting. Manipulation checks on self-affirmation suggest that Study 3 failed to achieve the desired manipulation. Thus, we were unable to confidently test our hypotheses regarding the effects of self-affirmation on women’s confrontation responses following exposure to sexism. To date, little has been done to explore what antecedents to Step 4 in the CPR model (i.e., deciding a course of action, Ashburn-Nardo et al., 2008) influence confrontation behaviors and approaches. So, while we may have failed to effectively manipulate self-affirmation in Study 3, we feel there is merit to this line of investigation. Since several of our questions are driven by existing theoretical structures, it is worth revisiting this study in the future once the derivation chain is better established.

We implore future confrontation researchers to explore the nuances of confrontation efficacy, as we believe this is a fertile line of research. Not only is there much to discover about general perceptions of confrontation efficacy regarding the different approaches to confrontation, but there is so much to explore concerning efficacy and perpetrator prototypicality. Prototypicality refers to an individual’s expectations about what constitutes discrimination or prejudice based on societal norms like status differences (Barrett & Swim, 1998; Inman & Baron, 1996). In situations of sexism, men are considered prototypical perpetrators (Helwig & McCoy, 2022b). In the scenario presented to participants in Studies 1 and 2, the perpetrator and confronter were not gendered. Thus, it would be interesting to explore how the perceived efficacy of a particular confrontation approach changes based on gender combinations of the perpetrator and confronter.
Future Implications

The present research addressed a gap in the confrontation literature by working to develop and validate a novel scale to help assess the latent construct of confrontation approach. By striving to develop reliable and valid measure, we can strengthen the conclusions drawn from future prejudice confrontation research. Additionally, the three studies presented in this manuscript offer a greater understanding of confrontation approaches when confronting sexism. The information gleaned from the present research regarding the HATE-CARE scale can help guide future studies that will offer a more comprehensive understanding of the situational factors that influence different components of the CPR model. Furthermore, while the studies presented focused on sexism, similar strategies can be used to explore other forms of prejudice, such as racism or classism, in the future.
REFERENCES


Barrett, P. (2005). What if there were no psychometrics?: Constructs, complexity, and measurement. Journal of Personality Assessment, 85(2), 134-140.


Loftus, G. R. (1996). Psychology will be a much better science when we change the way we analyze data. Current Directions in Psychological Science, 5(6), 161-171.


Vaccarino, E., & Kawakami, K. (2020). In the office or at the gym: The impact of confronting sexism in specific contexts on support for confrontation and perceptions of others. *Self and Identity, 20*(7) 1-20.

70


Appendix A: Studies 1 and 2 Demographic Information

Table 1

Studies 1 and 2 Demographic Information

<table>
<thead>
<tr>
<th></th>
<th>Study 1: Exploratory Analysis (N = 235)</th>
<th>Study 2: Confirmatory Analysis (N = 186)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>19.07 (1.48)</td>
<td>19.31 (1.77)</td>
</tr>
<tr>
<td>Min-Max</td>
<td>18-28</td>
<td>18-33</td>
</tr>
<tr>
<td>Gender (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>62.6</td>
<td>55.9</td>
</tr>
<tr>
<td>Female</td>
<td>26.0</td>
<td>40.3</td>
</tr>
<tr>
<td>Gender Nonconforming</td>
<td>6.9</td>
<td>3.8</td>
</tr>
<tr>
<td>Did Not Specify</td>
<td>4.5</td>
<td>0</td>
</tr>
<tr>
<td>Race (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>86.8</td>
<td>84.9</td>
</tr>
<tr>
<td>Black or African American</td>
<td>1.7</td>
<td>1.6</td>
</tr>
<tr>
<td>American Indian or Alaska Native</td>
<td>1.3</td>
<td>0.5</td>
</tr>
<tr>
<td>Asian</td>
<td>1.3</td>
<td>2.2</td>
</tr>
<tr>
<td>Multiracial</td>
<td>3.0</td>
<td>8.1</td>
</tr>
<tr>
<td>Did Not Specify</td>
<td>5.9</td>
<td>2.7</td>
</tr>
<tr>
<td>Ethnicity (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic or LatinX</td>
<td>3.4</td>
<td>4.3</td>
</tr>
<tr>
<td>Did Not Specify</td>
<td>8.5</td>
<td>4.3</td>
</tr>
</tbody>
</table>
Appendix B: HATE Stimulus

First off, thank you for wasting 1.5 minutes of our lives with this unhelpful, misogynistic yammering. Rather than assessing the actual candidates, you created a straw-woman onto whom you projected your own deeply troubling beliefs about women. Wow, you are a sexist person. That was such utter nonsense and incorrect assumptions that I am practically speechless. You are really the worst person ever. You should really educate yourself or just shut your mouth because your ideas and justification are utter nonsense. You had literally nothing to back up why Robert would make a great candidate and stuttered your way through a sexist and pathetic speech. As such, none of it was useful. Looking at resumes, Rebecca is a more qualified candidate for this position. Not only does she have more research experience, her research experience is more relevant to the lab’s academic focus. Rebecca is experienced in social and physiological psychology while Robert is more focused on neuropsychology. Rebecca’s experience better aligns with the desired skills for the research manager role. Assuming her interview and reference check do not turn up any issues, Rebecca is the better fit for the position as described.
Appendix C: CARE Stimulus

I appreciate your input. Let's not look at their performance through the lens of male vs. female, but rather what they can bring to the table. Firstly, I would like to respond to your concerns about Rebecca. When reviewing candidates for a job, I think we need to be fair and leave behind any preconceived notions. You worry Rebecca will not have ideas, considering her successful grades and her experiences in research I think I can venture to say she is not one who has ever lacked ideas or a work-ethic. I think we're focusing too much on personalities instead of qualifications. We don't know how Rebecca or Robert will perform in a work setting, so it's unfair to judge them based on this aspect. Let's just look at the facts: Rebecca has experience with human research, while Robert has dealt more with animal testing. Not only does Rebecca have greater familiarity with human-based research, but she has an expressed interest and experience in the social psychology field, whereas Robert has more experience in neuropsychology. Given her proclivity towards social psychology, it is without doubt that Rebecca is the better option for this position.
## Appendix D: Initial EFA Structure Matrix

### Table 3

**Initial EFA Structure Matrix**

<table>
<thead>
<tr>
<th>Item</th>
<th>Factor 1</th>
<th>Factor 2</th>
<th>Factor 3</th>
<th>Factor 4</th>
<th>Factor 5</th>
<th>Factor 6</th>
<th>Factor 7</th>
<th>Factor 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attacking</td>
<td>.894</td>
<td>-.071</td>
<td>-.586</td>
<td>-.321</td>
<td>-.706</td>
<td>.533</td>
<td>-.387</td>
<td>.297</td>
</tr>
<tr>
<td>In-your-face</td>
<td>.878</td>
<td>.045</td>
<td>-.563</td>
<td>-.374</td>
<td>-.601</td>
<td>.458</td>
<td>-.289</td>
<td>.287</td>
</tr>
<tr>
<td>Aggressive</td>
<td>.853</td>
<td>-.031</td>
<td>-.600</td>
<td>-.342</td>
<td>-.696</td>
<td>.541</td>
<td>-.404</td>
<td>.326</td>
</tr>
<tr>
<td>Snappy</td>
<td>.848</td>
<td>-.108</td>
<td>-.592</td>
<td>-.398</td>
<td>-.607</td>
<td>.516</td>
<td>-.333</td>
<td>.302</td>
</tr>
<tr>
<td>Targeting</td>
<td>.823</td>
<td>-.052</td>
<td>-.553</td>
<td>-.273</td>
<td>-.607</td>
<td>.481</td>
<td>-.354</td>
<td>.271</td>
</tr>
<tr>
<td>Extreme</td>
<td>.810</td>
<td>-.212</td>
<td>-.443</td>
<td>-.411</td>
<td>-.643</td>
<td>.599</td>
<td>-.507</td>
<td>.264</td>
</tr>
<tr>
<td>Hostile</td>
<td>.808</td>
<td>-.063</td>
<td>-.544</td>
<td>-.333</td>
<td>-.761</td>
<td>.558</td>
<td>-.438</td>
<td>.462</td>
</tr>
<tr>
<td>Abrasive</td>
<td>.800</td>
<td>-.079</td>
<td>-.490</td>
<td>-.376</td>
<td>-.508</td>
<td>.461</td>
<td>-.266</td>
<td>.236</td>
</tr>
<tr>
<td>Forceful</td>
<td>.799</td>
<td>.049</td>
<td>-.460</td>
<td>-.318</td>
<td>-.424</td>
<td>.338</td>
<td>-.352</td>
<td>.091</td>
</tr>
<tr>
<td>Negative</td>
<td>.769</td>
<td>-.185</td>
<td>-.620</td>
<td>-.337</td>
<td>-.640</td>
<td>.648</td>
<td>-.492</td>
<td>.361</td>
</tr>
<tr>
<td>Confrontational</td>
<td>.761</td>
<td>.229</td>
<td>-.514</td>
<td>-.215</td>
<td>-.455</td>
<td>.309</td>
<td>-.143</td>
<td>.259</td>
</tr>
<tr>
<td>Nasty</td>
<td>.758</td>
<td>-.147</td>
<td>-.632</td>
<td>-.305</td>
<td>-.636</td>
<td>.646</td>
<td>-.520</td>
<td>.331</td>
</tr>
<tr>
<td>Mean</td>
<td>.752</td>
<td>-.199</td>
<td>-.533</td>
<td>-.317</td>
<td>-.716</td>
<td>.632</td>
<td>-.623</td>
<td>.343</td>
</tr>
<tr>
<td>Accusatory</td>
<td>.747</td>
<td>-.046</td>
<td>-.480</td>
<td>-.335</td>
<td>-.466</td>
<td>.281</td>
<td>-.160</td>
<td>.183</td>
</tr>
<tr>
<td>Threatening</td>
<td>.736</td>
<td>-.201</td>
<td>-.372</td>
<td>-.268</td>
<td>-.599</td>
<td>.592</td>
<td>-.552</td>
<td>.175</td>
</tr>
<tr>
<td>Hateful</td>
<td>.730</td>
<td>-.151</td>
<td>-.547</td>
<td>-.242</td>
<td>-.722</td>
<td>.681</td>
<td>-.555</td>
<td>.387</td>
</tr>
<tr>
<td>Bold</td>
<td>.717</td>
<td>.251</td>
<td>-.452</td>
<td>-.339</td>
<td>-.420</td>
<td>.243</td>
<td>-.176</td>
<td>.114</td>
</tr>
<tr>
<td>Crude</td>
<td>.711</td>
<td>-.249</td>
<td>-.494</td>
<td>-.268</td>
<td>-.706</td>
<td>.680</td>
<td>-.550</td>
<td>.316</td>
</tr>
<tr>
<td>Comative</td>
<td>.706</td>
<td>.057</td>
<td>-.471</td>
<td>-.206</td>
<td>-.634</td>
<td>.376</td>
<td>-.284</td>
<td>.268</td>
</tr>
<tr>
<td>Destructive</td>
<td>.700</td>
<td>-.177</td>
<td>-.425</td>
<td>-.338</td>
<td>-.577</td>
<td>.661</td>
<td>-.496</td>
<td>.343</td>
</tr>
<tr>
<td>Argumentative</td>
<td>.679</td>
<td>.134</td>
<td>-.437</td>
<td>-.266</td>
<td>-.409</td>
<td>.288</td>
<td>-.149</td>
<td>.176</td>
</tr>
<tr>
<td>Disagreeable</td>
<td>.614</td>
<td>-.242</td>
<td>-.455</td>
<td>-.270</td>
<td>-.435</td>
<td>.526</td>
<td>-.505</td>
<td>.154</td>
</tr>
<tr>
<td>Divided</td>
<td>.496</td>
<td>-.237</td>
<td>-.245</td>
<td>-.252</td>
<td>-.336</td>
<td>.436</td>
<td>-.288</td>
<td>.035</td>
</tr>
<tr>
<td>Assertive</td>
<td>.433</td>
<td>.402</td>
<td>-.360</td>
<td>-.116</td>
<td>-.323</td>
<td>.025</td>
<td>-.119</td>
<td>.127</td>
</tr>
<tr>
<td>Evidence-based</td>
<td>-.252</td>
<td>.694</td>
<td>.349</td>
<td>.125</td>
<td>.469</td>
<td>-.688</td>
<td>.409</td>
<td>-.321</td>
</tr>
<tr>
<td>Assured</td>
<td>.022</td>
<td>.626</td>
<td>.116</td>
<td>.061</td>
<td>.147</td>
<td>-.164</td>
<td>.040</td>
<td>-.148</td>
</tr>
<tr>
<td>Confident</td>
<td>.132</td>
<td>.576</td>
<td>-.033</td>
<td>.065</td>
<td>.053</td>
<td>-.245</td>
<td>.171</td>
<td>-.156</td>
</tr>
<tr>
<td>Genuine</td>
<td>-.215</td>
<td>.570</td>
<td>.371</td>
<td>-.051</td>
<td>.537</td>
<td>-.498</td>
<td>.323</td>
<td>-.197</td>
</tr>
<tr>
<td>Matter-of-fact</td>
<td>-.005</td>
<td>.559</td>
<td>.124</td>
<td>-.037</td>
<td>.125</td>
<td>-.405</td>
<td>.110</td>
<td>-.245</td>
</tr>
<tr>
<td>Dignified</td>
<td>-.246</td>
<td>.543</td>
<td>.268</td>
<td>.112</td>
<td>.352</td>
<td>-.515</td>
<td>.150</td>
<td>-.288</td>
</tr>
<tr>
<td>Nice</td>
<td>-.593</td>
<td>.099</td>
<td>.876</td>
<td>.055</td>
<td>.624</td>
<td>-.383</td>
<td>.156</td>
<td>-.474</td>
</tr>
<tr>
<td>Gentle</td>
<td>-.452</td>
<td>-.028</td>
<td>.809</td>
<td>-.104</td>
<td>.505</td>
<td>-.307</td>
<td>.034</td>
<td>-.461</td>
</tr>
<tr>
<td>Pleasant</td>
<td>-.582</td>
<td>.174</td>
<td>.805</td>
<td>.110</td>
<td>.559</td>
<td>-.404</td>
<td>.125</td>
<td>-.444</td>
</tr>
<tr>
<td>Careful</td>
<td>-.415</td>
<td>.156</td>
<td>.786</td>
<td>.045</td>
<td>.576</td>
<td>-.463</td>
<td>.146</td>
<td>-.490</td>
</tr>
<tr>
<td>Positive</td>
<td>-.494</td>
<td>.318</td>
<td>.725</td>
<td>.096</td>
<td>.567</td>
<td>-.542</td>
<td>.263</td>
<td>-.514</td>
</tr>
<tr>
<td>Composed</td>
<td>-.599</td>
<td>.398</td>
<td>.711</td>
<td>.265</td>
<td>.624</td>
<td>-.648</td>
<td>.283</td>
<td>-.584</td>
</tr>
<tr>
<td>Levelheaded</td>
<td>-.572</td>
<td>.369</td>
<td>.699</td>
<td>.259</td>
<td>.602</td>
<td>-.541</td>
<td>.176</td>
<td>-.572</td>
</tr>
</tbody>
</table>
Table 3 Continued.

| Delicate  | Tolerant | .658 | Agreeable | .583 | Inquisitive | Humorous | Funny | Sarcastic | Civil | Respectful | Mature | Unkind | Thoughtful | Antagonistic | Understanding | Compassionate | Ineffective | Unproductive | Childish | Immature | Effective | Productive | Rational | Malicious | Mindful | Obnoxious | Educational | Sensible | Pragmatic | Juvenile | Harsh | Amiable | Collected | Constructive | Impassive |
|-----------|---------|------|-----------|------|------------|----------|------|----------|------|-----------|-------|--------|-----------|-------------|-------------|--------------|------------|-----------|----------|---------|----------|-----------|----------|----------|--------|--------|---------|---------|----------|----------|----------|
| -0.298    | -0.365  | 0.070| -0.361    | -0.096| -0.193     | 0.182    | 0.331| -0.365   | -0.632| -0.628    | -0.564| 0.661   | 0.472     | 0.693       | -0.426      | -0.245       | 0.555      | 0.631     | 0.553    | 0.603    | -0.366   | -0.431   | -0.476   | 0.676    | -0.552  | 0.630    | -0.366   | 0.332    | 0.200    |
| -0.076    | 0.028   | 0.374| -0.130    | -0.208| -0.887     | 0.122    | -0.047| -0.467   | -0.297| -0.204    | 0.608 | -0.490  | 0.580     | -0.069      | 0.231       | -0.036       | -0.268     | -0.275    | -0.284   | -0.286   | -0.505   | 0.501    | 0.520    | -0.477   | -0.463   | 0.463    | 0.532    | 0.535    | 0.528    |
| -0.374    | 0.399   | -0.236| -0.181    | 0.021 | -0.021    | 0.030    | -0.070| -0.117   | -0.297| -0.122    | -0.775| 0.768   | 0.768     | -0.451      | -0.383      | 0.612        | -0.569     | -0.345    | -0.204   | -0.337   | 0.034    | 0.103    | 0.202    | -0.633   | -0.137   | -0.647   | 0.361    | 0.361    | 0.428    |
| -0.236    | -0.273  | -0.364| -0.009    | 0.020 | 0.020     | -0.032   | -0.070| -0.117   | -0.536| -0.319    | -0.654| 0.663   | 0.663     | -0.414      | -0.284      | -0.375       | -0.869     | -0.584    | -0.345   | -0.654   | 0.565    | 0.579    | 0.202    | -0.633   | -0.661   | -0.681   | 0.509    | 0.509    | 0.166    |
| -0.006    | 0.115   | -0.555| 0.328     | -0.006| -0.207    | -0.555   | 0.328| -0.414   | -0.284| -0.204    | -0.613| 0.663   | 0.663     | 0.315       | 0.232       | -0.272       | -0.547     | -0.566    | -0.294   | -0.538   | -0.789   | -0.762   | -0.337   | -0.734   | -0.681   | -0.681   | 0.330    | 0.330    | 0.546    |
| -0.364    | -0.555  | 0.006| -0.308    | 0.006 | -0.194    | 0.043    | 0.028| -0.232   | -0.248| 0.248     | 0.546| 0.546   | 0.546     | -0.043      | -0.462      | -0.716       | -0.490     | -0.343    | -0.264   | -0.353   | -0.297   | -0.272   | -0.590   | -0.734   | 0.504    | 0.504    | 0.349    | 0.349    | 0.113    |

76
## Appendix E: Inter-Item Correlations Between HATE and CARE Scales

### Table 6

*Inter-item Correlations Between HATE and CARE Scales*

<table>
<thead>
<tr>
<th></th>
<th>Civil</th>
<th>Respectful</th>
<th>Rational</th>
<th>Genuine</th>
<th>Collected</th>
<th>Thoughtful</th>
<th>Dignified</th>
<th>Compassionate</th>
<th>Careful</th>
<th>Understanding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attacking</td>
<td>-.76</td>
<td>-.75</td>
<td>-.58</td>
<td>-.35</td>
<td>-.63</td>
<td>-.67</td>
<td>-.33</td>
<td>-.43</td>
<td>-.59</td>
<td>-.58</td>
</tr>
<tr>
<td>Threatening</td>
<td>-.58</td>
<td>-.60</td>
<td>-.54</td>
<td>-.40</td>
<td>-.48</td>
<td>-.60</td>
<td>-.29</td>
<td>-.33</td>
<td>-.42</td>
<td>-.44</td>
</tr>
<tr>
<td>Extreme</td>
<td>-.72</td>
<td>-.69</td>
<td>-.64</td>
<td>-.41</td>
<td>-.63</td>
<td>-.61</td>
<td>-.36</td>
<td>-.33</td>
<td>-.46</td>
<td>-.48</td>
</tr>
<tr>
<td>Hostile</td>
<td>-.79</td>
<td>-.78</td>
<td>-.62</td>
<td>-.34</td>
<td>-.69</td>
<td>-.68</td>
<td>-.33</td>
<td>-.42</td>
<td>-.58</td>
<td>-.57</td>
</tr>
<tr>
<td>Accusatory</td>
<td>-.59</td>
<td>-.53</td>
<td>-.42</td>
<td>-.14*</td>
<td>-.48</td>
<td>-.37</td>
<td>-.18</td>
<td>-.26</td>
<td>-.40</td>
<td>-.46</td>
</tr>
<tr>
<td>Abrasive</td>
<td>-.62</td>
<td>-.60</td>
<td>-.49</td>
<td>-.17*</td>
<td>-.51</td>
<td>-.49</td>
<td>-.23</td>
<td>-.31</td>
<td>-.46</td>
<td>-.43</td>
</tr>
<tr>
<td>Forceful</td>
<td>-.48</td>
<td>-.54</td>
<td>-.37</td>
<td>-.19</td>
<td>-.36</td>
<td>-.36</td>
<td>-.17*</td>
<td>-.24</td>
<td>-.37</td>
<td>-.38</td>
</tr>
<tr>
<td>Targeting</td>
<td>-.65</td>
<td>-.65</td>
<td>-.53</td>
<td>-.28</td>
<td>-.55</td>
<td>-.57</td>
<td>-.25</td>
<td>-.37</td>
<td>-.51</td>
<td>-.50</td>
</tr>
</tbody>
</table>

*Note:* *Indicates the correlation is significant at the p < .05 level (2-tailed). All other correlations are significant at the p < .01 level (2-tailed).*
Appendix F: Study 3 Demographic Information

Table 12

*Study 3 Demographic Information*

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N = 96</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>19.03 (1.8)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Min-Max</td>
<td>18-33</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>89.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black or African American</td>
<td>2.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>American Indian or Alaska Native</td>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multiracial</td>
<td>4.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did Not Specify</td>
<td>2.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethnicity (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic or LatinX</td>
<td>5.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did Not Specify</td>
<td>2.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix G: B-SAT Affirmation Condition

Instructions: Use the following scale to rate your agreement with the statements below.

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Strongly Disagree</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Strongly Agree</td>
<td></td>
</tr>
</tbody>
</table>

1. I love to learn new things.
2. My friends can trust me.
3. I always try to keep my word.
4. I am always curious about the world.
5. There are people in my life who care as much about my feelings and well-being as they do about their own.
6. I value my ability to think critically.
7. My friends value my good judgement.
8. I can express love to someone else.
9. I treat all people equally, regardless of who they might be.
10. Despite challenges, I always remain hopeful about the future.
Appendix H: B-SAT No Affirmation Condition

Instructions: Think about the celebrity David Beckham. Use the scale to rate your agreement with the statements.

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly Disagree</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>Strongly Agree</td>
</tr>
</tbody>
</table>

1. He loves to learn new things.
2. His friends can trust him.
3. He always tries to keep his word.
4. He is always curious about the world.
5. There are people in his life who care as much about his feelings and well-being as they do about their own.
6. He values his ability to think critically.
7. His friends value his good judgement.
8. He can express love to someone else.
9. He treats all people equally, regardless of who they might be.
10. Despite challenges, he always remains hopeful about the future.
### Appendix I: PANAS-X Form

This scale consists of a number of words and phrases that describe different feelings and emotions. Read each item and then mark the appropriate answer in the space next to that word. Indicate to what extent you feel this way **right now, in this moment**. Use the following scale to record your answers:

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very slightly or not at all</td>
<td>A little</td>
<td>Moderately</td>
<td>Quite a bit</td>
<td>Extremely</td>
</tr>
</tbody>
</table>

| ____ cheerful | ____ shaky | ____ lively |
| ____ disgusted | ____ happy | ____ ashamed |
| ____ attentive | ____ timid | ____ at ease |
| ____ bashful | ____ alone | ____ scared |
| ____ sluggish | ____ alert | ____ drowsy |
| ____ daring | ____ upset | ____ angry at self |
| ____ surprised | ____ angry | ____ enthusiastic |
| ____ strong | ____ bold | ____ downhearted |
| ____ scornful | ____ blue | ____ sheepish |
| ____ relaxed | ____ shy | ____ distressed |
| ____ irritable | ____ active | ____ blameworthy |
| ____ delighted | ____ guilty | ____ determined |
| ____ inspired | ____ joyful | ____ frightened |
| ____ fearless | ____ nervous | ____ astonished |
| ____ disgusted with self | ____ lonely | ____ interested |
| ____ sad | ____ sleepy | ____ loathing |
| ____ calm | ____ excited | ____ confident |
| ____ afraid | ____ hostile | ____ energetic |
| ____ tired | ____ proud | ____ concentrating |
| ____ amazed | ____ jittery | ____ dissatisfied with self |
BIOGRAPHY OF THE AUTHOR

Kathryn L. Roderick was born in Fredericksburg, Virginia on October 7, 1996. She was raised in Fredericksburg, Virginia and graduated from Chancellor High School and the Commonwealth Governor’s School in 2015. She attended the University of Lynchburg and graduated summa cum laude with a Bachelor of Science in Psychology in 2019. She moved to Maine and entered the Psychology graduate program at the University of Maine in the fall of 2020. After receiving her degree, Kathryn will continue at the University of Maine to obtain her doctoral degree. Kathryn is a candidate for the Master of Arts degree in Psychology from the University of Maine in August 2022.