2019

When Empathy Only Goes So Far: Development of a Trait Parochial Empathy Scale

Anna Maria C. Behler
Virginia Commonwealth University

Follow this and additional works at: https://scholarscompass.vcu.edu/etd
Part of the Quantitative Psychology Commons, and the Social Psychology Commons

© Anna Maria C. Behler

Downloaded from
https://scholarscompass.vcu.edu/etd/6005

This Dissertation is brought to you for free and open access by the Graduate School at VCU Scholars Compass. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of VCU Scholars Compass. For more information, please contact libcompass@vcu.edu.
WHEN EMPATHY ONLY GOES SO FAR: DEVELOPMENT OF A TRAIT PAROCHIAL EMPATHY SCALE

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Psychology at Virginia Commonwealth University

By: ANNA MARIA C. BEHLER
M.S., Virginia Commonwealth University, 2017
M.A., CUNY Queens College, 2014
B.S., Fordham University – Rose Hill, 2011

Co-Directors: Jeffrey D. Green, Ph.D. & Jennifer Joy-Gaba, Ph.D.
Associate Professors of Psychology
Department of Psychology

Virginia Commonwealth University
Richmond, Virginia
July, 2019
Acknowledgements

There are several people to whom I would like to say thank you: First, to my advisors, Dr. Jeffrey Green & Dr. Jennifer Joy-Gaba, for their kindness, encouragement, support, and unparalleled mentorship over the past four years. To my committee members, Drs. Daniel Berry, Terri Sullivan, and Sharon Zumbrunn, for their guidance on this project. To my parents, brother, and sister, who push me to be the best version of myself every day. To Hanna, who is always by my side reminding me that it’s the little things that make life so special. And last but not least, to Don, for always believing in me, and for making my dreams come true.
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>vi</td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Empathy and Group Identity</td>
<td>1</td>
</tr>
<tr>
<td>Empathy: A Multi-Faceted Construct</td>
<td>3</td>
</tr>
<tr>
<td>TPES Item Generation</td>
<td>6</td>
</tr>
<tr>
<td>Study 1</td>
<td>9</td>
</tr>
<tr>
<td>TPES Scale Description</td>
<td>9</td>
</tr>
<tr>
<td>Method</td>
<td>9</td>
</tr>
<tr>
<td>Results</td>
<td>13</td>
</tr>
<tr>
<td>Discussion</td>
<td>24</td>
</tr>
<tr>
<td>Study 2</td>
<td>26</td>
</tr>
<tr>
<td>Method</td>
<td>26</td>
</tr>
<tr>
<td>Results</td>
<td>26</td>
</tr>
<tr>
<td>Discussion</td>
<td>35</td>
</tr>
<tr>
<td>Study 3</td>
<td>38</td>
</tr>
<tr>
<td>Method</td>
<td>39</td>
</tr>
<tr>
<td>Results</td>
<td>41</td>
</tr>
<tr>
<td>Discussion</td>
<td>42</td>
</tr>
<tr>
<td>Study 4</td>
<td>44</td>
</tr>
<tr>
<td>Method</td>
<td>44</td>
</tr>
<tr>
<td>Results</td>
<td>45</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
</tr>
<tr>
<td>-------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Discussion</td>
<td>50</td>
</tr>
<tr>
<td>Study 5</td>
<td>52</td>
</tr>
<tr>
<td>Method</td>
<td>53</td>
</tr>
<tr>
<td>Results</td>
<td>57</td>
</tr>
<tr>
<td>Discussion</td>
<td>58</td>
</tr>
<tr>
<td>General Discussion</td>
<td>60</td>
</tr>
<tr>
<td>References</td>
<td>64</td>
</tr>
<tr>
<td>Appendices</td>
<td>72</td>
</tr>
<tr>
<td>Appendix A</td>
<td>72</td>
</tr>
<tr>
<td>Appendix B</td>
<td>74</td>
</tr>
<tr>
<td>Appendix C</td>
<td>75</td>
</tr>
<tr>
<td>Appendix D</td>
<td>76</td>
</tr>
<tr>
<td>Appendix E</td>
<td>77</td>
</tr>
<tr>
<td>Appendix F</td>
<td>79</td>
</tr>
<tr>
<td>Appendix G</td>
<td>80</td>
</tr>
<tr>
<td>Appendix H</td>
<td>81</td>
</tr>
<tr>
<td>Appendix I</td>
<td>82</td>
</tr>
<tr>
<td>Appendix J</td>
<td>83</td>
</tr>
<tr>
<td>Appendix K</td>
<td>85</td>
</tr>
</tbody>
</table>
List of Tables

Table 1. Factor Loading from Exploratory Factor Analysis – Ingroup Subscale .................. 15
Table 2. Factor Correlation Matrix – TPES Ingroup Subscale ........................................ 18
Table 3. Factor Loading from Exploratory Factor Analysis – Outgroup Subscale ............... 19
Table 4. Factor Correlation Matrix – TPES Outgroup Subscale ....................................... 22
Table 5. Regression Weights – Study 1 ............................................................................. 23
Table 6. Regression Weights – Study 2 ............................................................................. 34
Table 7. Regression Weights – Study 4 ............................................................................. 49
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.</td>
<td>Scree Plot – Ingroup Subscale of the TPES</td>
<td>15</td>
</tr>
<tr>
<td>Figure 2.</td>
<td>Scree Plot – Outgroup Subscale of the TPES</td>
<td>18</td>
</tr>
<tr>
<td>Figure 3.</td>
<td>Confirmatory Factor Analysis – Ingroup Subscale (Initial Model)</td>
<td>29</td>
</tr>
<tr>
<td>Figure 4.</td>
<td>Confirmatory Factor Analysis – Ingroup Subscale (Respecified Model)</td>
<td>31</td>
</tr>
<tr>
<td>Figure 5.</td>
<td>Confirmatory Factor Analysis – Outgroup Subscale (Initial Model)</td>
<td>32</td>
</tr>
<tr>
<td>Figure 6.</td>
<td>Confirmatory Factor Analysis – Outgroup Subscale (Respecified Model)</td>
<td>33</td>
</tr>
<tr>
<td>Figure 7.</td>
<td>Confirmatory Factor Analysis – Ingroup Subscale (Final Model)</td>
<td>47</td>
</tr>
<tr>
<td>Figure 8.</td>
<td>Confirmatory Factor Analysis – Outgroup Subscale (Final Model)</td>
<td>48</td>
</tr>
</tbody>
</table>
Abstract

Empathy, the ability to feel and/or understand another’s emotional state, plays a significant role in interpersonal interactions, mitigating hostility and enhancing affiliation and helping. However, empathy also biases interpersonal reactions. For example, at the group level empathy can become amplified towards members of their ingroup and blunted towards individuals in outgroups, a term called parochial empathy. Currently, no validated measures of parochial empathy at the dispositional level exist, and development of such a scale would be important to understanding the role of group-based emotions in prejudice and discrimination against outgroups. I conducted five studies to develop and validate a self-report Trait Parochial Empathy Scale (TPES) that could measure tendencies to respond with parochial empathy across any set of group membership categories. Study 1 assessed the factor structure of the TPES through exploratory and confirmatory factor analyses while Study 2 attempted to replicate the Study 1 factor structure and assess concurrent and divergent validity of the TPES using attitudinal measures. Study 3 assessed the temporal consistency of the TPES. Study 4 examined whether the TPES could be flexibly used across a variety of groups by assessing its relation to various outcomes across different ingroup and outgroup combinations. Finally, Study 5 assessed the ability of the TPES to predict in vivo behavior.
Empathy, the ability to vicariously feel and understand the emotions of others, is an essential component of maintaining positive relationships and enhancing cooperation among individuals. Most notably, empathy can serve as a catalyst for helping others and has been positively linked to prosocial and altruistic behaviors (Eisenberg & Miller, 1987; Batson et al., 2016). Empathic individuals are often also better able to perspective take and consider situational factors, thus making them more adept at understanding the others’ decisions, even if they might not agree or be able to relate to them (Vescio, Sechrist, & Paolucci, 2003; Simon, Moss, & O’Brien, 2017).

At first glance, empathy may seem to be the principal answer to resolving intergroup conflict. Indeed, it has often been looked to as such with many interventions focusing on empathy to reduce intergroup tensions and hostilities (Zaki & Cikara, 2015). However, empathy has come under closer scrutiny in recent years, with some researchers noting that empathic responses to the plights of others are neither automatic nor universal.

**Empathy and Group Identity**

Empathy can motivate hostility towards people dissimilar to a target of empathy, even in the absence of provocation (Buffone & Poulin, 2014). These dissimilarities can take many forms including, but not limited to, race, religion, gender, political affiliation, and even more superficial categorizations such as school affiliation or team membership. Although empathizing
with stigmatized or dissimilar others can increase compassionate responding to that group (Batson et al., 1997), feeling empathy for one group decreases empathy and prosociality toward other groups (Batson & Ahmad, 2009). Further, the benefits of empathy do not necessarily transfer when considering how individuals exhibit empathic concern for members of ingroups versus outgroups (Bruneau, Cikara, & Saxe, 2015; 2017). There is even evidence to suggest that when controlling for trait empathy, there are notable unique and opposite effects on attitudes towards ingroup and outgroup members. This phenomenon is known as parochial empathy. The word parochial means limited in range or scope. In the context of empathy, this refers to the human tendency to exhibit an intergroup empathy bias, allocating empathy unevenly across groups and ultimately showing greater empathic concern for members of our ingroup. This manifests itself in several ways, most notably via holding prejudicial attitudes about outgroups and exhibiting negative behavioral responses toward them.

Several researchers have acknowledged the importance in understanding these intergroup variations in empathy (Redford & Ratliff, 2017; Weisz & Zaki, 2018), but there remains no validated trait measure to assess this construct. The research thus far has assessed situation-specific empathic responses. For example, some of the most prolific work in this area has used narrative descriptions to assess attitudes towards members of ingroups and outgroups as an assessment of parochial empathy (Bruneau, Cikara, & Saxe, 2015; 2017; Bruneau & Kteily, 2017). Because individuals differ in the extent to which they may have personal experience within the different hypothetical scenarios described in narrative descriptions, narrative-based scales may elicit more measurement error compared to scales asking about general experiences. For example, one such scenario in these narratives involves an individual being left at the altar on the day of their wedding. This may elicit different levels of empathy not due to
ingroup/outgroup identification, but rather due to how relevant the concept of weddings is to the reader in that moment, or due to the reader having experienced the same event either themselves or through a close other.

Therefore, the purpose of the proposed studies is to develop and validate a scale that a) measures both cognitive and affective components of trait parochial empathy, b) is independent of participants’ personal experiences with specific scenarios, c) measures empathy for ingroup and outgroup members across various types of group memberships, and d) can predict behavior towards outgroup members. I address a limitation of previous work that has examined this construct by not linking the scale items to specific scenarios: the scale is intended to be used flexibly with different ingroup/outgroup combinations. In doing so, this removes potential confounds such as a participant potentially having had similar (or no) personal experiences in a specific context. Perhaps most importantly, the creation of such a measure will allow researchers to control for trait levels of parochial empathy.

The Trait Parochial Empathy scale (TPES) will help address a gap in empathy measurement, allowing researchers to quickly assess participants’ efficacy and willingness to empathize with people as a function of their group membership in a variety of domains (e.g., political affiliation, race, religion, sexuality, gender identity, etc.). In contrast to prior measures of parochial empathy, this scale will be generalizable across situations and groups. This flexibility will allow researchers to better assess moderators of empathy failures such as group membership, stereotypes, and perceived ability to empathize.

**Empathy: A Multi-Faceted Construct**

Early research on empathy conceptualized it as a construct containing four distinct factors (Davis, 1983). The Interpersonal Reactivity Index (IRI) is a commonly used and well-validated
measure of empathy (Chrysikou & Thompson, 2016), and contains unique subscales that map on to each construct. The first is perspective-taking, or an individual’s tendency to put themselves in someone else’s shoes and consider other viewpoints. The second is fantasy, or an individual’s propensity to insert themselves into fictional or imaginative scenarios. This is best characterized by how deeply affected one is by the emotions of fictional characters in movies, television, or literature. The next factor is empathic concern, which assesses a person’s feelings of sympathy and concern when considering the plight of others. Last is personal distress, which is defined as feelings of anxiety and uneasiness in interpersonal settings where tensions are high.

When developing potential items for the TPES, I considered all four subscales to be applicable to the goals of the scale development. However, items contained in the fantasy subscale did not seem particularly applicable to the goals of the proposed scale as they were currently written, as mentally inserting oneself into fictional worlds was not relevant to the types of questions contained in the TPES measure. This subscale includes items such as “After seeing a play or movie, I have felt as though I were one of the characters” and “When I am reading an interesting story or novel, I imagine how I would feel if the events in the story were happening to me”. While these items speak to an individual’s ability to put themselves in another’s situation, the construct of parochial empathy is centered around how we allocate empathy differently towards ingroup and outgroup members in various domains. These discrepancies in empathy have important real-world implications. As such, I felt that the items needed to be grounded more in reality, such as assessing how likely a person is to try and consider the perspective of others when reading news stories about an outgroup.

A second well-validated measure known as the Basic Empathy Scale (BES, Joliffe & Farrington, 2006) conceptualizes empathy as having two distinct factors—affective and
cognitive. The affective subscale assesses an individual’s ability to share the emotions of another, whereas the cognitive subscale examines how much they are able to understand the emotions and perspectives of others. Like the affective component of the BES, the Toronto Empathy Questionnaire (Spreng, McKinnon, Mar, & Levine, 2009) conceptualizes empathy as a primarily emotional construct, and thus was also drawn upon for some of the items in the initial pool. Finally, the Questionnaire for Cognitive and Affective Empathy (QCAE; Reniers, Corcoran, Drake, Shrayne, & Volim, 2011) also focused on the affective components of empathy, but included items intended to assess one’s ability and willingness to actively perspective take.

Ultimately I developed items that took into consideration an individual’s propensity to feel and accurately discern the emotions of ingroup and outgroup members, as well as their motivation to attend to the needs of those in their ingroup and outgroup, to develop comparable items that encompassed elements found within all four of the aforementioned measures. Thus, I predict that the final parochial empathy measure will contain elements of empathy that are in alignment with these well-validated measures (Davis, 1983; Jolliffe & Farrington, 2006; Spreng, McKinnon, Mar, & Levine, 2009; Reniers, Corcoran, Drake, Shrayne, & Volim, 2011). However, while some of these elements will map onto both the ingroup and outgroup components of the measure, I also believe there will be distinct differences between these two subscales, further underscoring my proposition that parochial empathy is a unique form of empathy. Further, while there will be some overlap on these two subscales, they are intended to be used separately, as I am hypothesizing that they will serve as unique predictors of other self-report measures and behaviors. In addition to their utility as individual scales, I believe that using
them in combination might offer greater predictive power of attitudes and behavior than currently existing empathy scales.

**TPES Item Generation**

I developed an item pool of 70 questions that reflect domains derived from the previously described elements of validated empathy scales (Davis, 1983; Jolliffe & Farrington, 2006; Spreng, McKinnon, Mar, Levine, 2009; Reniers, Corcoran, Drake, Shrayne, & Volim, 2011). These domains are: affective matching (feeling what another feels), empathic concern (feeling “for” another), perspective-taking (“putting oneself in another’s shoes”), and empathic accuracy (accurately inferring another’s emotional state). A panel of psychology graduate students from various fields of study (clinical, counseling, developmental, and health disciplines) rated the adequacy of each item as an assessment tool. All panelists were familiar with the construct of empathy as well as the principles of scale development. Thus, they rated items based upon their relevancy to the construct being measured, clarity of wording, and conciseness of each question.

Items that were rated highly and consistently across raters (Cronbach’s alpha ≥ .70) were retained. As shown in Appendix A, the final scale includes 28 items that were pilot tested by a sample of 300 undergraduate students at Virginia Commonwealth University. An exploratory factor analysis identified either a four or six factor solution. However, this pilot study was limited in that there were few items used to assess convergent and discriminant validity, and there were no screening requirements to ensure that participants considered themselves as members of the political categories of interest. Thus, I hope to reassess the factor structure using a more reliable sample and then confirm in a more generalizable population.

Given that all of the items were created with various elements of empathy in mind, I predict the scale will show internal consistency, and will demonstrate construct validity. I also
predict that the TPES will exhibit convergent and discriminant validity with several personality traits as well as general attitudes towards outgroups (see Measures in Study 1).

Across five studies, I will assess the following seven hypotheses:

H1: The Trait Parochial Empathy Scale (TPES) will capture distinct affective and cognitive factors of empathy across two unique subscales—ingroup empathy and outgroup empathy (Studies 1 and 2).

H2: TPES scores on both subscales will be positively related to trait agreeableness (see Melchers et al., 2016), as well as trait empathy (see Davis, 1983), indicating convergent validity. A more in depth explanation of how these constructs will be assessed is provided in the measures section of Study 1.

H3: Scores on the outgroup subscale will be negatively related to social dominance orientation (see Ho et al., 2015), trait hostility (see Bruneau et al., 2017), outgroup dehumanization (see Kteily, Bruneau, Waytz, & Cotterill, 2015) and explicit attitudes toward an ingroup (Studies 1, 2, 3, and 4). They will also be positively related to malicious envy (see Lange & Crucius, 2015). These relationships will also serve as indicators of convergent validity, and further information about how these constructs will be assessed can be found in the measures section of Study 1.

H4: TPES scores will not be significantly correlated with conscientiousness, neuroticism, openness to experience, and extraversion, indicating discriminant validity (see Melchers et al., 2016; Studies 1, 2, 3, and 4). As with the previous trait items, additional information on these constructs can be found in the measures section of Study 1.

H5: The TPES will show strong temporal stability (Study 3).
H6: The relation between the TPES and convergent constructs will replicate across different categories of group membership (Study 4).

H7: TPES scores will predict greater helping for ingroups versus outgroups (Study 5) over and above trait empathy.

In today’s sociopolitical climate, individuals may believe that a lack of empathy for others, particularly outgroup members, may be at the root of many of our current dilemmas (Zaki & Cikara, 2015). The goals of this project are to increase our understanding of the underlying causes and consequences of social issues. Indeed, the goal of this project is to show that the TPES will increase people’s understanding of parochial empathy’s function as an underlying mechanism of prejudice and discrimination at the individual level. It is important to note, however, that this scale should not be seen as a substitute for prejudice measures, but is intended to work in conjunction with them. I contend that parochial empathy is a latent contributor to prejudice, and it may help predict prejudicial attitudes towards outgroups, particularly in individuals for whom the empathy gap is largest. Further, an individual may exhibit an empathic bias for their ingroup while still being high in overall empathy for everyone, and they may not consciously hold prejudicial attitudes. Therefore, this series of studies will serve as the foundation for future research exploring how we can increase empathic responses while reducing negative attitudes towards outgroups.
Study 1

TPES Scale Description

The TPES is designed to measure tendencies to empathize differently with ingroup and outgroup members by having participants report their empathic tendencies about people from each type of stated/given group. The scale is flexible to use with different group membership categories, which participants would select from a pre-specified list of group categories. As an example, participants might be asked to identify which group pair they identify with the most and least, thus specifying their “ingroup” and “outgroup” (e.g., “Republican” and “Democrat”). The purpose of this selection is to ensure that participants are choosing their appropriate ingroup and an outgroup with which they least identify, as simply saying “all other parties” or “another political party” may be too vague. Further, prejudicial attitudes are typically targeted at others based on their specific group membership. Therefore, it is important that participants select a specific outgroup to think about when answering the relevant items on the TPES.

In this study, participants were first report their tendencies to empathize with members of their ingroup, and were then asked to complete other questionnaires. Lastly, participants reported their tendencies to empathize with members of their outgroup. The TPES was scored by calculating means for the ingroup and outgroup questionnaires as two separate subscales.

Method

Participants

I recruited 366 U.S. college undergraduates from a large, urban research university. Participants were U.S. citizens aged 18 and over ($M_{\text{age}} = 19.51$, $SD = 3.15$). Of these, seven participants were removed from the final analyses for failing attention check items, leaving 359 participants in the final sample. The sample was mostly women (74.4%), but was reasonably
racially and ethnically diverse (49.9% White, 18.7% Black/African-American, 15.9% Asian, 10.9% Hispanic/Latino, 6.6% multiracial or other).

Aside from United States residency and age requirements, there were no other exclusion criteria. As this scale is likely to be used in studies that utilize college undergraduate populations, the first study used this as the population of interest. However, later studies will seek to replicate the findings of Study 1 in more generalizable samples.

**Procedure**

Participants completed the TPES using political group categories (see Appendix A), as this is a group affiliation that often overlaps several other identities (race, religion, gender, etc.). In addition, recent findings that suggest political decision-making and empathy may be linked (Sirin, Valentino, & Villalobos, 2017). Between completing the ingroup and outgroup questionnaires of the TPES, participants completed the measures listed below (Appendices B-J), and lastly reported demographics (Appendix K).

**Convergent/Discriminant Measures**

**International Personality Item Pool (Mini-IPIP; Donnellan et al., 2006).** The Mini-IPIP measures traits of extraversion ($\alpha = .79$), conscientiousness ($\alpha = .71$), openness ($\alpha = .69$), positive emotionality (i.e., neuroticism; $\alpha = .72$), and agreeableness ($\alpha = .71$). All items were rated on a scale from 1 (Disagree strongly) to 5 (Agree strongly). I expected that because extraversion and conscientiousness are conceptually unrelated to empathy or outgroup attitudes, both extraversion and conscientiousness would be unrelated to either subscale, showing discriminant validity. I also expected that agreeableness and positive emotionality would be positively correlated with both subscales of the TPES in accordance with previous research.
findings (Joliffe & Farrington, 2006; Graziano, Bruce, Sheese, & Tobin, 2007), providing evidence of convergent validity (see Appendix B).

**Social Dominance Orientation Scale (SDO; Ho et al., 2015; \( \alpha = .84 \)).** This 8-item scale reflects preferences for social group inequality. Items were rated from 1 (Very negative) to 7 (Very positive), with higher scores indicating more preference for social hierarchies. Because empathy involves greater concern for others and the ability to consider someone else’s circumstances, I expected that this measure would negatively correlate with scores on the TPES, providing evidence of convergent validity (see Appendix C).

**Ascent Measure (Kteily et al., 2015).** This scale measures perceptions of dehumanization of other groups. Participants were shown the “Ascent of Man” diagram showing human physiological evolution, moving from early apelike ancestors to more culturally modern human ancestors. Participants used a slider along the bottom of the diagram to show the “evolvedness” of the group listed (see Appendix D). By its very definition empathy involves concern for the plight of others, making it highly unlikely that participants exhibiting higher empathy for others would view them as less human. Thus, I expected that this measure would negatively correlate with scores on the outgroup subscale, providing further evidence of convergent validity.

**Brief Aggression Questionnaire (BAQ; Webster et al., 2013; \( \alpha = .78 \)).** The BAQ measures self-reported tendencies to be violent, feel anger, and make hostile attributions. Items are rated from 1 (Extremely uncharacteristic of me) to 5 (Extremely characteristic of me) (see Appendix E). Given the antisocial nature of these items, I expected that this measure would negatively correlate with scores on the ingroup and outgroup subscales, providing additional evidence of convergent validity.
**Feelings Thermometer.** This scale measures explicit attitudes toward different groups. Participants reported how cold \((0 = \text{unfavorable})\) or warm \((10 = \text{favorable})\) they felt towards that group (see Appendix F). This measure is being included for exploratory purposes.

**Marlowe-Crowne Social Desirability Scale-Short Form (MCSD; Reynolds, 1982; \(\alpha = .65\)).** This 13-item scale measures tendencies to respond in socially desirable ways. Each item is marked True or False \((\text{True} = 1; \text{False} = 0)\). Higher scores indicate greater social desirability bias. Scores > 9 were used to filter out participants whose social desirability tendencies may have biased their responses (see Appendix G).

**Benign and Malicious Envy Scale (BeMaS; Lange & Crucius, 2015).** The 10-item scale consists of ten self-report items that assess two dimensions of envy at the trait level. The benign subscale \((\alpha = .82)\) assesses envy as a source of motivation, whereas the malicious subscale \((\alpha = .84)\) assesses feelings of bitterness and one’s desire for harm to befall an envied target. The 5-item malicious subscale was of interest for the purposes of this dissertation. All items are rated on a 6-point Likert-type scale, ranging from 1 \((\text{strongly disagree})\) to 6 \((\text{strongly agree})\). Given the negative emotional valence associated with envy, I expected that ingroup empathy subscale scores would be negatively correlated with this measure, providing evidence of convergent validity (see Appendix H).

**Single Item Trait Empathy Scale (SITES; Konrath, Meier, & Bushman, 2018).** This measure assesses trait empathy, and participants responded to the statement “I am a very empathic person” on a 1 \((\text{Not very true of me})\) to 5 \((\text{Very true of me})\) scale (see Appendix I). This item is being included for exploratory purposes to see if this newly validated measure will map on to both subscales of the TPES.
Interpersonal Reactivity Index (IRI; Davis, 1983; \( \alpha = .83 \)). This measure also assesses trait empathy but includes the four unique subscales discussed in the introduction of this document (i.e., Perspective Taking, Fantasy, Empathic Concern, and Personal Distress). All items are assessed on a 5-point Likert-type scale ranging from 1 (\textit{Does not describe me well}) to 5 (\textit{Describes me very well}). I included this measure in order to assess how the ingroup and outgroup subscales of the TPES map onto these well-validated empathy items (see Appendix J). I expected that trait empathy would relate positively to both subscales of the TPES.

**Results**

Exploratory Factor Analysis

Factor analysis is a multivariate statistical procedure that allows a large pool of items to be collapsed into a smaller number of distinguishable subscales, otherwise known as factors. Further, it allows researchers to examine the relationship between subscales in order to determine the underlying factors of a theoretical construct (Williams, Brown, & Osman, 2010). The ultimate goal of a factor analysis is to develop a parsimonious means by which to analyze the various elements of a particular construct.

Exploratory factor analysis (EFA) is one subtype for this category of analyses. As the name suggests, there are no \textit{a priori} hypotheses regarding the specific factor structure that will result from the pool of items. This is often leveraged as a weakness of EFA, with critics suggesting that the results are based on the researcher’s subjective interpretation of the data rather than a theoretical formulation (Tabachnick & Fidell, 2007). Thus, it is important to note that the initial TPES items were informed by previously validated empathy measures. Further, my interpretation of the resulting factors was also informed by previous empathy research, as well as research on ingroup and outgroup dynamics.
In order to determine the factor structure of the Trait Parochial Empathy Scale, an EFA assuming no \textit{a priori} factor structure was performed using principal axis factoring and a Promax rotation including all 28 items on data from a sample of 344 participants. This was conducted separately for both the in-group and outgroup subscales. The EFA was conducted in SPSS using an alpha level of .05. I specified a promax rotation to identify eigenvalues greater than one, a standard cutoff point in EFA. More specifically, factors that fall below an eigenvalue of 1 do not add significant amount of explained variance to the overall factor solution. The promax rotation also assumes that there will be significant correlations of at least .30 between factors, which is expected among subscales of the same measure, making this a statistically appropriate choice. Additionally, the pattern matrix was analyzed to identify how individual items load onto specific factors. Items should load onto a primary factor with a weight of +/- .40 or higher, and should not load onto a secondary factor with a weight inside the range of +/- .40. Lastly, factor correlations were assessed to ensure that there was not a violation of the collinearity assumption (i.e., factors should not be correlated above .70).

**Ingroup Subscale**

For the in-group subscale, a scree plot (Cattell, 1966; Figure 1) revealed an inflection point between the third and fourth-highest eigenvalues. However, the eigenvalue cutoff of one pointed to a possible 6 factor solution, though there was only a small difference between the amount of cumulative variance in the items explained by the first five factors (54.28%) compared to the first six (58.04%).

These small differences between the fifth and sixth factors in amounts of variance explained, in combination with the inflection point between the third and fourth eigenvalue,
suggested that further examination of the item loadings was necessary to make a decision about which factors to retain.

Figure 1. Scree plot for the ingroup subscale of the TPES showing an inflection point between the third and fourth eigenvalues.

The item loadings for the first six factors in this EFA appear in Table 1. I used .40 as a general cutoff for determining whether an item loads meaningfully onto a factor, and ensured that item cross-loadings were at least .15 less than the item’s highest factor loading (Worthington and Wittaker, 2006) to determine if an item achieved simple structure.

Table 1.

<table>
<thead>
<tr>
<th>Factor Name &amp; Items</th>
<th>Factor Loading</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Factor 1: Empathic Accuracy</strong></td>
<td></td>
</tr>
<tr>
<td>#3: I can tell when a person who is a ______ is upset even when they don’t say anything.</td>
<td>.52* .03 .14 .07 .01 .13</td>
</tr>
<tr>
<td>#16: I can often tell when someone who is a ______ is hiding their true emotions.</td>
<td>.76* -.09 -.06 .09 -.01 .09</td>
</tr>
<tr>
<td>#17: I am quick to spot when someone who is a ______ is feeling awkward or uncomfortable.</td>
<td>.78* .09 .09 -.02 -.04 .04</td>
</tr>
<tr>
<td>#18: It’s easy for me to tell when a ______ is interested in what I’m saying.</td>
<td>.63* .14 .12 -.07 .08 -.08</td>
</tr>
<tr>
<td>#19: ______ people tell me I’m good at understanding what they are thinking and feeling.</td>
<td>.54* .10 .21 -.07 .13 -.22</td>
</tr>
</tbody>
</table>
#20: I can usually figure out how _______ people are feeling before they tell me.  

<table>
<thead>
<tr>
<th>Factor 2: Motivated Empathy</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>#22: When I think I’m about to criticize someone who is a ______, I first try to imagine how I would feel if I was in their place.</td>
<td>.11</td>
</tr>
<tr>
<td>#23: I can usually appreciate a ______’s viewpoints even if I don’t agree with them.</td>
<td>.09</td>
</tr>
<tr>
<td>#25: When people who are ______ tell me they are taken advantage of, I try to understand why they feel that way.</td>
<td>-0.08</td>
</tr>
<tr>
<td>#27: When I read stories about people who are ______ in the news, I try to imagine what I would feel like in their position.</td>
<td>-0.03</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Factor 3: Non-Empathic Responses</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>#2: I find it annoying when ______ people get excited.</td>
<td>-.17</td>
</tr>
<tr>
<td>#4: I have a hard time feeling “in tune” with the feelings of people who are ______.</td>
<td>.04</td>
</tr>
<tr>
<td>#10: When people who are ______ start to talk about their problems, I try to change the topic to something else.</td>
<td>-.07</td>
</tr>
<tr>
<td>#11: I’m not really interested in how ______ feel.</td>
<td>-.11</td>
</tr>
<tr>
<td>#12: People who are ______ don’t deserve my compassion.</td>
<td>-.07</td>
</tr>
<tr>
<td>#21: I’m not usually aware of the feelings of people who are ______.</td>
<td>.20</td>
</tr>
<tr>
<td>#24: Even when I try to consider the perspective of a ______, I find I just can’t understand them.</td>
<td>.14</td>
</tr>
<tr>
<td>#28: I don’t believe people who are ______ when they say they are disadvantaged.</td>
<td>-.06</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Factor 4: Affective Sharing</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>#1: When I see a ______ person get excited, I get excited too.</td>
<td>.02</td>
</tr>
<tr>
<td>#6: When I’m talking with someone who is a ______, I tend to feel the same emotions they are feeling.</td>
<td>.26</td>
</tr>
<tr>
<td>#7: It makes me happy to see people who are ______ happy.</td>
<td>.17</td>
</tr>
<tr>
<td>#8: It upsets me to see a person who is a ______ being treated disrespectfully.</td>
<td>.02</td>
</tr>
<tr>
<td>#9: I enjoy making people who are ______s feel better.</td>
<td>-.04</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Factor 5</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>#13: I often feel badly about the problems experienced by people who are ______.</td>
<td>-.02</td>
</tr>
<tr>
<td>#14: I feel sorry for the way that ______s get taken advantage of.</td>
<td>.06</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Factor 6</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>#26: It seems to me that people who are ______ are treated fairly in society.</td>
<td>.08</td>
</tr>
</tbody>
</table>
The five items that did not achieve simple structure were considered not to be a meaningful part of the factor solution and appear at the bottom of Table 1. When examining this pattern of item loadings, I found that items 5 and 15 did not load onto any factor, suggesting they should be removed. Additionally, only items 13 and 14 loaded onto the fifth factor. A factor should have at least three items, making this two-item fifth factor problematic. Finally, only a single item (item 26) loaded with simple structure on the sixth factor, suggesting the factor be removed. The first four factors accounted for approximately 50% of the cumulative item variance and for these reasons were retained.

Twenty-three of the original 28 items loaded with simple structure onto four factors (See Table 1). Six items loaded onto Factor 1, representing Empathic Accuracy, or how well a person believes they are able to recognize the emotions and perspectives of others. Four items loaded onto Factor 2, which I have termed Motivated Empathy, or the tendency to engage in perspective taking and consider situational factors that might influence another’s thoughts, feelings, and behaviors. The third factor included eight items, all of which represent non-empathic, and potentially even antagonistic, attitudes and behaviors (note: all of these items are reverse-coded in the scale). This factor will be referred to as Non-Empathic Responding. Finally, the fourth factor contained five items representing Affective Sharing, or the tendency to recognize and feel similar emotions when interacting with others. (See Table 2 for factor correlations.)
Table 2.

*Factor Correlation Matrix – TPES Ingroup Subscale*

<table>
<thead>
<tr>
<th>Factor</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Empathic Accuracy</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>Motivated Empathy</td>
<td>.33</td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>Non-Emathic Responding</td>
<td>-.07</td>
<td>-.37</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>Affective Sharing</td>
<td>.46</td>
<td>.50</td>
<td>-.44</td>
<td></td>
</tr>
</tbody>
</table>

**Outgroup Subscale**

The same criteria were used to determine the factor structure of the outgroup subscale of the TPES. A scree plot (Cattell, 1966; Figure 2) once again revealed an inflection point between the third and fourth-highest eigenvalues. However, the eigenvalue cutoff of one pointed to a 4 or 5 factor solution. There was once again a small difference between the amount of cumulative variance in the items explained by the first four factors (55.78%) compared to the first five (60.00%). Taken together, this once again suggested that further examination of the item loadings was necessary to determine the factor structure of this subscale. The item loadings for the first 5 factors in this EFA appear in Table 3.
Figure 2. Scree plot for the outgroup subscale of the TPES showing an inflection point between the third and fourth eigenvalues.

The five items that did not achieve simple structure were considered not to be a meaningful part of the factor solution and appear at the bottom of Table 3. When examining this pattern of item loadings, I found that item 7 did not load onto any factor, suggesting removal of that item. Additionally, items 12 and 24 loaded onto the multiple factors, suggesting that these items are not reliable. However, this presented an additional issue, as removing item 24 left the fifth factor with only 2 items. As previously stated, a factor should have at least three items, making this two-item fifth factor problematic. Furthermore, when examining the factor correlation matrix, I noted that factor five exhibited low correlations (.08-.30) with all other factors. Taken together, this suggested the removal of the fifth factor. The first four factors accounted for approximately 56% of the cumulative item variance and for these reasons were retained.

Table 3.

*Factor Loading from Exploratory Factor Analysis – Outgroup Subscale*

<table>
<thead>
<tr>
<th>Factor Name &amp; Items</th>
<th>Factor Loading</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Factor 1: Motivated Empathy</strong></td>
<td></td>
</tr>
<tr>
<td>#8: It upsets me to see a person who is a _____ being treated disrespectfully.</td>
<td>.79* - .05 -.02 .06 - .12</td>
</tr>
<tr>
<td>#9: I enjoy making people who are _____s feel better.</td>
<td>.50* .00 .31 .15 -.15</td>
</tr>
<tr>
<td>#13: I often feel about the problems experienced by people who are ____.</td>
<td>.62* .00 .01 -.21 .30</td>
</tr>
<tr>
<td>#14: I feel sorry for the way that _____s get taken advantage of.</td>
<td>.61* -.02 .15 -.20 .35</td>
</tr>
<tr>
<td>#22: When I think I’m about to criticize someone who is a _____, I first try to imagine how I would feel if I was in their place.</td>
<td>.77* .03 -.07 .05 -.10</td>
</tr>
<tr>
<td>#23: I can usually appreciate a _____’s viewpoints even if I don’t agree with them.</td>
<td>.83* .08 -.21 .02 -.08</td>
</tr>
<tr>
<td>#25: When people who are ____ tell me they are taken advantage of, I try to understand why they feel that way.</td>
<td>.84* .01 -.09 -.03 -.04</td>
</tr>
</tbody>
</table>
When I read stories about people who are ______ in the news, I try to imagine what I would feel like in their position.

Factor 2: Empathic Accuracy

| #3: I can tell when a person who is a ______ is upset even when they don’t say anything. | -.04 | .67* | -.05 | .02 | -.07 |
| #16: I can often tell when someone who is a ______ is hiding their true emotions. | -.01 | .74* | .13 | -.08 | .03 |
| #17: I am quick to spot when someone who is a ______ is feeling awkward or uncomfortable. | .05 | .81* | -.04 | .09 | .02 |
| #18: It’s easy for me to tell when a ______ is interested in what I’m saying. | .19 | .59* | .09 | -.02 | -.04 |
| #20: I can usually figure out how ______ people are feeling before they tell me. | -.03 | .82* | -.04 | .06 | -.01 |

Factor 3: Affective Sharing

| #1: When I see a ______ person get excited, I get excited too. | -.05 | .01 | .75* | .06 | -.07 |
| #5: When I see people who are ______ seem anxious, I get anxious as well. | -.04 | .03 | .67* | -.17 | .01 |
| #6: When I’m talking with someone who is a ______, I tend to feel the same emotions they are feeling. | -.03 | -.04 | .82* | .02 | -.02 |
| #15: I take part in activities to try and help give a voice to people who are ______. | -.11 | .05 | .62* | -.17 | .22 |
| #19: ______ people tell me I’m good at understanding what they are thinking and feeling. | .11 | .32 | .43* | .04 | .10 |

Factor 4: Non-Empathic Responses

| #2: I find it annoying when ______ people get excited. | .00 | -.13 | .13 | .44* | .01 |
| #4: I have a hard time feeling “in tune” with the feelings of people who are ______. | -.05 | -.19 | .22 | .40* | .26 |
| #10: When people who are ______ start to talk about their problems, I try to change the topic to something else. | -.04 | .02 | -.08 | .61* | .05 |
| #11: I’m not really interested in how ______s feel. | .17 | .02 | -.02 | .66* | .07 |
| #21: I’m not usually aware of the feelings of people who are ______. | -.21 | .26 | -.28 | .55* | .15 |

Factor 5

| #24: Even when I try to consider the perspective of a ______, I find I just can’t understand them. | -.05 | .02 | .09 | .45 | .47* |
| #26: It seems to me that people who are ______ are treated fairly in society. | -.26 | .04 | .01 | .12 | .56* |
| #28: I don’t believe people who are ______ when they say they are disadvantaged. | .11 | -.13 | .02 | .25 | .60* |

No Factor Loading

| #7: It makes me happy to see people who are ______ happy. | .37 | -.01 | .38 | .27 | -.24 |
| #12: People who are ______ don’t deserve my compassion. | .40 | -.04 | -.20 | .56* | .05 |
Similar to the ingroup subscale, 23 of the original 28 items loaded with simple structure onto four factors (See Table 3). For the outgroup subscale, Factor 1 represented Motivated Empathy, and contained eight items. Factor 2 represented Empathic Accuracy and contained five items. Factor 3 also had a five-item loading and represented Affective Sharing. Finally, the last five items loaded onto the Non-Empathic Responding factor.

Item 7 did not load onto any of the factors and was deleted. Additionally, items 12 and 24 loaded onto multiple factors with correlations greater than or equal to .40, meaning that they could not be reliably distinguished as belonging to a distinct factor. This meant that items 26 and 28 were the only two that grouped into a fifth factor, and thus were also removed (see Table 4 for factor correlations).

Although there was a similar pattern to the findings regarding the qualities of the factors, there were some notable differences. Specifically, these differences occurred in the order in which the factors loaded, as well as the number of items that loaded onto each subscale. The Empathic Accuracy factor was nearly identical for both subscales, except for Item 19, “_____ people tell me I’m good at understanding what they are thinking and feeling”, which did not load onto this factor for the outgroup measurement. Notably, this item instead loaded onto the Affective Sharing factor of the outgroup subscale. For the outgroup, Affective Sharing also included Items 5 (“When I see people who are _____ seem anxious, I get anxious as well”) and 15 (“I take part in activities to try and help give a voice to people who are _____”). The Motivated Empathy factor contained the same four items (22, 23, 25, 27) for both the ingroup and outgroup. However, for the outgroup, this factor also included items 8 and 9, which were categorized as Affective Sharing for the ingroup subscale. Finally, the Non-Empathic Responding factor contained all the same items for the ingroup and outgroup subscales, except
for items 12, 24, and 28, which were all deleted from the outgroup measure.

Table 4.

*Factor Correlation Matrix – TPES Outgroup Subscale*

<table>
<thead>
<tr>
<th>Factor</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Empathic Accuracy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Motivated Empathy</td>
<td>.25</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Empathic Responding</td>
<td>-.65</td>
<td>-.22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affective Sharing</td>
<td>.45</td>
<td>.23</td>
<td>-.36</td>
<td></td>
</tr>
</tbody>
</table>

**Convergent/Discriminant Validity**

Hierarchical linear regression was used to assess the relationship between the TPES and the pre-test scale measures while controlling for levels of social desirability (see Table 5 for full results of this analysis). All assumptions were verified and met prior to running this analysis, including univariate and multivariate normality, an absence of multicollinearity, and homoscedasticity.

I expected that the following traits would significantly predict scores on the TPES, indicating convergent validity: agreeableness, openness, malicious envy, social dominance orientation, hostility, dehumanization, and the IRI empathy measure. More specifically, I hypothesized that agreeableness, openness, and the IRI would be positively related to both subscales, whereas malicious envy, social dominance orientation, hostility, and dehumanization would be negatively related to each subscale. I also expected that extraversion and
conscientiousness would be traits that would not significantly predict scores on either subscale of the TPES, indicating discriminant validity of the measure. These predictions were mostly supported, with the personality traits of agreeableness, openness, extraversion, and conscientiousness converging and diverging with the TPES scores exactly as expected.

These were a few inconsistencies my initial predictions. First, malicious envy did not predict scores on either subscale of the TPES. Interestingly, however, was that whereas both trait empathy and social dominance orientation were significant positive predictors of ingroup empathy, neither predicted outgroup empathy. Furthermore, both aggression and dehumanization were significant positive predictors of ingroup empathy but were even stronger negative predictors of outgroup empathy. This pattern of results underscores the nuances of how empathy is considered differently towards ingroups and outgroups.

Table 5.

*Regression Weights – Study 1*

<table>
<thead>
<tr>
<th>Variable (N = 359)</th>
<th>M</th>
<th>SD</th>
<th>β, p Ingroup</th>
<th>β, p Outgroup</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social Desirability</td>
<td>6.59</td>
<td>2.64</td>
<td>-.03, .54</td>
<td>.15, .006*</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>4.12</td>
<td>.69</td>
<td>.18, .002*</td>
<td>.20, &lt; .001*</td>
</tr>
<tr>
<td>Openness</td>
<td>3.94</td>
<td>.70</td>
<td>.16, .003*</td>
<td>.16, .002*</td>
</tr>
<tr>
<td>Extraversion</td>
<td>3.13</td>
<td>.98</td>
<td>-.05, .27</td>
<td>-.02, .75</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>3.53</td>
<td>.83</td>
<td>.06, .26</td>
<td>-.02, .74</td>
</tr>
<tr>
<td>Empathy</td>
<td>3.47</td>
<td>.48</td>
<td>.18, .002*</td>
<td>.10, .07</td>
</tr>
<tr>
<td>Malicious Envy</td>
<td>2.29</td>
<td>.95</td>
<td>.02, .71</td>
<td>.03, .59</td>
</tr>
<tr>
<td>Aggression/Hostility</td>
<td>2.45</td>
<td>.65</td>
<td>.09, .10</td>
<td>-.13, .02*</td>
</tr>
</tbody>
</table>
An exploratory factor analysis provided unique four factor solutions for each subscale of the TPES: Empathic Accuracy, Motivated Empathy, Affective Sharing, and Non-Emathic Responding. However, while both the ingroup and outgroup subscales of the TPES produced similar four factor solutions, the order in which these factors were extracted as well as the items contained within each factor differed in a slight, albeit notable, manner. These variations underscore the notion that there are differences in the ways we allocate empathy towards ingroups versus outgroups. Empathic accuracy was the factor that loaded most prominently for the ingroup subscale, which aligns with the typical definition of empathy—an ability to understand the perspectives and feelings of others (Davis, 1983). This was the factor that was most similar between the two subscales, suggesting that what shapes our beliefs about how well we are able to empathize with others is similar across group memberships, regardless of our actual ability to do so. Likewise, the Non-Emathic Responding factor was also relatively consistent across the ingroup and outgroup subscales, indicating that there are similarities in the thoughts we hold and behaviors we exhibit towards others when we lack empathy for them.

Interestingly, the primary element of the outgroup subscale was Motivated Empathy, which potentially suggests that maintaining empathy for outgroups is a more active cognitive and emotional process requiring internal and/or external motivation. This aligns with previous work suggesting that while empathy often appears to be an automatic neurological process, this emotion is also influenced by other factors such as contextual appraisal or one’s relationship with
the target (Singer & Lamm, 2009; Esmenio et al., 2019). Notably, there were also some items that shifted between the Motivated Empathy and Affective Sharing factors when considering the ingroup and outgroup subscales, suggesting that perhaps what we view simply as sharing in someone else’s experiences requires greater effort when that person belongs to a different group.

There were also differences in the relationships between the TPES ingroup and outgroup subscales in regard to their relationships with other traits. Both subscales related positively to personality traits like agreeableness and openness, but there were some differences regarding their relationships with aggression, social dominance, and even trait empathy. This pattern suggested that while higher ingroup empathy is associated with positive traits such as agreeableness and openness, it is also related to traits that are associated with higher rates of hostility, prejudice, and discrimination. These findings align with the notion that empathy is not an entirely positive force, and tends to be allocated in accordance with our ability to relate to a specific person or group (Bloom, 2017).

The outgroup subscale of the TPES showed no relationship to trait empathy which may seem surprising at first, but may make sense when considering the parochial nature of empathy that I am trying to highlight. That is to say that my results underscore the importance of the TPES as an assessment tool when considering attitudes towards different groups, as traditional empathy measures such as the IRI may only be capturing how individuals allocate empathy toward those in our ingroup. It was also interesting to note that outgroup empathy was positively associated with social desirability but not trait empathy, suggesting that a strong empathic preference towards outgroup members may be somewhat driven by a desire to be viewed favorably by others rather than true feelings of empathic concern.
While there were several strengths of Study 1, it was limited in that it used a college undergraduate sample. Therefore, a follow-up study will confirm the Study 1 factor structure and convergent/discriminant validity findings with a more generalizable population.
Study 2

Study 2 attempted to confirm the factor structure of the TPES obtained in Study 1 with a new, more generalizable sample. Participants were recruited from Amazon’s Mechanical Turk website.

Method

Participants

I recruited 406 participants via Amazon’s Mechanical Turk website. A sample size of ≥ 300 is considered appropriate for conducting well-powered confirmatory factor analyses (Meyers, Gamst, & Guarino, 2016). The sample was mostly White (78.3%), but was reasonably balanced in terms of gender (53.3% female; $M_{age} = 35.75, SD = 12.52$). Based on the political identifications requested at the beginning of the survey, 61.6% of the respondents identified as Republican or Conservative, and the remaining 38.4% identified as Democrat or Liberal.

Procedure

Because this was a confirmatory factor analysis, the procedure and measures were identical to Study 1.

Results

Confirmatory Factor Analysis

Confirmatory factor analyses allow researchers to verify preconceived ideas about their data. I used a confirmatory factor analysis (CFA) with maximum-likelihood (ML) estimation to assess the fit indices suggested by results of the EFA in Study 1. ML estimation is the primary estimation procedure used in confirmatory factor analyses (Flora & Curran, 2004). A CFA requires a specific number of factors and their inter-correlations to be specified in advance, with the overall purpose of this analysis being to examine the fit indices of a model in order to attain a
model that is “identified”, (i.e., offers the best fit). Therefore, I expected that the same number of factors found in the EFA for Study 1 would be identified in this analysis, and that the retained items of the TPES would load onto the factors in the same way. I used IBM SPSS AMOS to conduct the CFA. This is a useful structural equation modeling software that also allows for the creation of visual representations of models, including how individual items map onto each latent factor, as well as how the factors correlate with one another. As in Study 1, I considered the ingroup and outgroup subscales of the TPES separately and conducted a unique CFA for each. The manifest variables for each the CFA were the 23 retained items from the TPES, and the 4 latent constructs were the four factors of the TPES that had been revealed by the previously conducted EFA. Assumptions of skewness and kurtosis were checked and met for each of the manifest variables for both analyses. However, the ingroup analysis produced 11 multivariate outliers, leaving a final sample size of 395 participants, and the outgroup analysis produced 15 multivariate outliers, for a final sample of 391 participants.

**Ingroup Confirmatory Factor Analysis (see Figure 3).** There are a number of statistics that should be attained in order to assess model fit. As such, I used the following criteria: First, individual items should load onto factors with correlations above $r = .40$. All of the scale items loaded significantly onto their respective first-order factors (all $p’s < .001$). Additionally, the individual factors were not correlated above $r = .70$, as this would indicate the presence of multicollinearity among the subscales of the measure. It should be noted that the Empathic Accuracy and Non-Empathic Responding factors were the only two that did not share a significant relationship, and that this pattern of correlations between the factors replicated the findings from Study 1.

In addition to the correlations between items and factors, there were a number of indices
that needed to be examined to determine model fit (McDonald & Ho, 2002; Meyers, Gamst, & Guarino, 2016). The first are the absolute fit indices: The $\chi^2$ goodness-of-fit test was significant, $\chi^2 (224) = 753.15, p < .001$. The chi-squared test should be non-significant as an indicator of good fit, however this statistic is problematic, as it tends to be significant in large samples due to the sample size rather than as the result of an accurately detected significant effect (Gatignon, 2010). Thus, there are other, more accurate absolute indices that were used to confirm model fit.

The Goodness of Fit Index (GFI) assesses the fit between the predicted model and resulting covariance matrix, and should be $\geq .95$. In this case, the GFI value was only .86, indicating that the model fit was less than ideal. Additionally, the root mean square error of approximation (RMSEA) should fall below .10. Lower values of this statistic indicate better model fit, such that a value of $\leq .10$ is considered adequate and $\leq .08$ is considered good, though a value $\leq .05$ is ideal. The RMSEA of the initial model was .076, which indicated a good fit.

Finally, there is a relative fit measure known as the Comparative Fit Index (CFI), that accounts for model improvement in a comparison between a null model (i.e., one in which all items are uncorrelated) and the proposed model. Its value can range from 0-1, with obtained values closer to 1 representing a better model fit. It should be $\geq .95$ to be considered good, or $\geq .90$ for an adequate fit. In this case, the initial model produced a CFI of .86, indicating a less than adequate fit.

In order to determine the model fit of the TPES, it was important to take all of the aforementioned statistics into account. The RMSEA accounts for discrepancies due to sample size, making it one of the most popularly reported indices of model fit (Jackson, Gillaspy, & Purc-Stephenson, 2009), and leading some to argue that priority should be given to this value when making the final interpretation of model fit. However, Kenny, Kaniskan, & McCoach
(2016) warn that no one index of fit is the single gold standard to assess fit, as each can be influenced by a number of factors, including the sample size and model degrees of freedom. Thus, although the RMSEA should be weighted more heavily, I wanted to be conservative and determine the quality of a model by examining all of these statistics together, as no one on its own can be used to accurately assess the model’s overall identification. I concluded that the model was marginally adequate, given the good RMSEA value and while still accounting for the remaining indices which were less than ideal, though nearing acceptable levels.

Figure 3. Confirmatory factor analysis for the ingroup subscale of the TPES.

I examined the modification indices to determine whether adjustments could be made to the model in order for its fit to be improved. Upon examination of these indices, I noted that there were several error terms that could be correlated to improve the model fit. This should only be done if there is a theoretical rationale for why the error terms of two separate items should be
correlated, and so I only did so for items that could be justifiably linked (Kenny, 2011). It should be noted that this is not indicative of the presence of multicollinearity, but rather an indication that the items may be measuring similar constructs even within their unique factors.

On the Affective Sharing factor, items 1 “When I see a _____ person get excited, I get excited too” and 6 “When I’m talking with someone who is a _____, I tend to feel the same emotions they are feeling” were correlated ($r = .37$). On the Empathic Accuracy factor, items 2 “I find it annoying when _____ people get excited” and 4 “I have a hard time feeling ‘in tune’ with the feelings of people who are _____” shared a significant amount of error variance ($r = .20$). Regarding Motivated Empathy, items 22 “When I think I’m about to criticize someone who is a _____, I first try to imagine how I would feel if I was in their place” and 23 “I can usually appreciate a _____’s viewpoints even if I don’t agree with them” shared related error terms ($r = .37$). Finally, items 3 “I can tell when a person who is a _____ is upset even when they don’t say anything”, 16 “I can often tell when someone who is a _____ is hiding their true emotions”, 17 “I am quick to spot when someone who is a _____ is feeling awkward or uncomfortable”, 18 “It’s easy for me to tell when a _____ is interested in what I’m saying” and 19 “People who are _____ tell me I’m good at understanding what they are thinking and feeling” shared differing amounts of error variance within the Empathic Accuracy factor (see Figure 4 for full depiction of these relationships).

After these respecifications, the GFI was .89, which was nearer to approaching a good fit, though still slightly less than ideal. However, the CFI was .91 and the RMSEA improved from .076 to .063, both indicating a good fit. Overall, the respecified model indicated a much-improved fit from the original (see Figure 4).
Outgroup Confirmatory Factor Analysis (see Figure 5). The chi-square for the outgroup subscale was once again significant, $\chi^2 (224) = 932.12, p < .001$. The GFI, CFI and RMSEA did not approach acceptable levels, at .81, .85, and .09, respectively. Overall, the model fit was less than adequate and I again decided to examine the modification indices.
There were indications that Item 19 was problematic, as the modification indices recommended that it also be allowed to cross-load onto the Empathic Accuracy scale. There were also a number of items that shared error variance, including items 2 and 4, similar to the ingroup subscale \((r = .28)\). Regarding Motivated Empathy, items 8 “It upsets me to see a person who is a ____ being treated disrespectfully”, 9 “I enjoy making people who are _____s feel better”, 13 “I often feel about the problems experienced by people who are _____”, 14 “I feel sorry for the way that _____s get taken advantage of”, 22 “When I think I’m about to criticize someone who is a _____, I first try to imagine how I would feel if I was in their place” and 23 “I can usually appreciate a _____’s viewpoints even if I don’t agree with them” all shared varying amounts of error variance (see Figure 6 for full depiction of these relationships).
After these respecifications, the GFI was .87, which was approaching a good fit, though still slightly less than ideal. However, the CFI was .90 and the RMSEA improved from .090 to .075, both indicating a marginally good fit. Overall, the respecified model indicated a much-improved fit from the original (see Figure 6).

![Figure 6. Respecified model for the outgroup subscale of the TPES.](image)

**Convergent/Discriminant Validity**

Once again, hierarchical linear regression was used to assess the relationship between the TPES and the pre-test scale measures while controlling for levels of social desirability (see Table 6 for full results of this analysis). All assumptions were verified and met prior to running this analysis, including univariate and multivariate normality, an absence of multicollinearity, and homoscedasticity.
The pattern of results was nearly identical to those found in Study 1, with a few minor exceptions. As in Study 1, social desirability did not predict TPES ingroup scores, but it did positively predict scores on the outgroup subscale. Dehumanization positively predicted TPES scores on the ingroup subscale and negatively predicted them on the outgroup subscale. Once again, malicious envy did not predict scores on either subscale of the TPES.

In Study 1, both trait empathy and social dominance orientation were significant positive predictors of ingroup empathy, while neither predicted outgroup empathy. However, in Study 2, trait empathy significantly predicted both ingroup and outgroup scores. Furthermore, social dominance orientation negatively predicted ingroup empathy while positively predicting outgroup scores. There were also some differences regarding personality traits and the TPES, with agreeableness only predicting ingroup scores, and openness not predicting either subscale. Finally, aggression did not predict either subscale in this study. This pattern of results, though not completely identical to those of Study 1, still underscores the nuances of how empathy is considered differently towards ingroups and outgroups.

Table 6.

_Regression Weights – Study 2_

<table>
<thead>
<tr>
<th>Variable (N = 406)</th>
<th>M</th>
<th>SD</th>
<th>$\beta, p$ Ingroup</th>
<th>$\beta, p$ Outgroup</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social Desirability</td>
<td>6.32</td>
<td>3.12</td>
<td>-.05, .29</td>
<td>.12, .02*</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>3.88</td>
<td>.78</td>
<td>.24, &lt;.001*</td>
<td>.06, .34</td>
</tr>
<tr>
<td>Openness</td>
<td>3.44</td>
<td>.50</td>
<td>.07, .17</td>
<td>-.05, .31</td>
</tr>
<tr>
<td>Extraversion</td>
<td>2.80</td>
<td>1.00</td>
<td>.02, .66</td>
<td>.10, .03*</td>
</tr>
</tbody>
</table>
The results of Study 2 replicated the EFA from Study 1 almost perfectly, and in a more generalizable sample of adults across the United States. Using the factor structure found in Study 1, I was able to achieve models that were approaching a good fit but were less than ideal and thus required respecification. However, for both subscales, the adjustments that were needed were not to the item loadings themselves, but rather to the error terms between some of the items within factors. The main issue was that Item 19 needed to be dropped from the outgroup subscale of the TPES, but no other items gave an indication of being problematic.

For the ingroup subscale, a majority of the related error terms were on the Empathic Accuracy factor, likely due to the moderate to large correlations between all of the items that loaded onto that factor ($r$’s between .32-.54). This is also an indication that those items are tapping into a similar construct, and thus are reasonably likely to share some error variance. For the outgroup subscale, nearly all of the related error terms were on the Motivated Empathy factor. As previously discussed in Study 1, it may be that feelings of empathy towards outgroup members may require greater effort and motivation. This factor loaded most strongly for the outgroup subscale in Study 1, and several of these items shared large correlations ($r$’s between
.59-.82) in Study 2. As in the ingroup subscale, this is an indication that they are likely tapping into a similar construct and share error variance.

It was interesting to note that the Empathic Accuracy and Non-Empathic Responding factors were not correlated for the ingroup subscale, but were negatively related for the outgroup subscale. This may be due to the fact that responding empathically to our ingroups may be a more automatic process, and we may believe we are more accurate in discerning the emotions of those who are more similar to us regardless of whether we choose to respond empathically or not. However, showing empathy to outgroup members may require us to first believe that they are in need of empathy, once again suggests that having empathy for outgroup members may be a more effortful process. These possibilities await future empirical examination.

The convergent and discriminant validity findings were also similar, with a few notable exceptions. Trait empathy was significantly related to both subscales, where it had previously only been related to outgroup empathy. However, the relationship between scores on the IRI and the ingroup subscale were noticeably weaker. The personality trait items did not map on exactly as before, though this may be due to the fact that Study 1 was conducted in a college sample rather than an MTurk sample, and there is wide variability among college samples specifically with regard to the Big Five personality items (Corker et al., 2017). There is also a possibility that the results may have differed due to the fact that the political demographics of this sample were markedly different. While not a central focus of this study, it is possible that the relationships of personality and empathy may differ between liberals and conservatives. Future work should examine these relationships again in different samples to determine whether the convergent and discriminant validity findings of Study 1 or Study 2 can be replicated, and whether the political affiliation of the sample plays a role. (This will be somewhat addressed in Study 4, where the
convergent and discriminant measures were utilized in a new MTurk sample, though not using political affiliation as a categorizing variable.)

Now that the factor structure of the TPES has been established as well as the convergent and discriminant validity of the construct, the next step is to determine whether the measure exhibits temporal stability. The focus of Study 3 will be to establish whether scores on the TPES are consistent across multiple timepoints.
Study 3

The aim of study 3 was to establish the temporal stability of the Trait Parochial Empathy Scale (TPES) once its latent factor structure was explored in Study 1 and confirmed in Study 2. Temporal stability refers to the how stable a measure will be across timepoints, and test-retest reliability typically is the method used to assess whether a scale meets this criterion. This is an especially important form of reliability, as it indicates that a scale is a true and meaningful reflection of a construct and can be trusted to produce similar results over time.

According to Devellis (2017), temporal instability may result from a number of factors including:

1) An actual change in the construct being measured (e.g., mean levels of confidence in the economy may increase or decrease amongst a sampled group over time).

2) Naturally occurring variations in the construct being measured (e.g., student anxiety scores may change cyclically, in relation to the occurrence of midterm and final examinations).

3) Changes that are the result of a difference in the subjects themselves rather than a change in the construct being measured (e.g., a participant is distracted or tired and does not attend to questions as carefully, resulting in less agreement and lower reliability of their scores across time).

4) Unreliability of the measure itself. Of the four sources of temporal instability, this item is clearly problematic, as it indicates the measure is not appropriate to assess the construct of interest.

In the case of the TPES, temporal stability is a crucial component of the measure, especially in light of how empathy towards outgroups may be shaped by current events. If the TPES is to be considered a reliable measure, participants should exhibit similar scores across
time points. Moreover, as I proposed a measure to assess parochial empathy at the trait level, I expected that it should be fairly stable, even when taking into account current events, as personality traits tend to be stable across one’s lifetime (Soldz & Vaillant, 1999). However, given that there may be large-scale events that people encounter that may influence their attitudes towards outgroup members, I assessed whether any such event occurred during the duration of data collection.

**Method**

**Participants**

According to Shoukri, Asyali, & Donner (2004), 101 participants provided sufficient power to detect the minimum acceptable effect of ICC = .70 with two time-point measurements (α = .05, power ≥ .95). However, I overrecruited in order to account for unusable data, participant carelessness, and participants who fail to complete the measures at both timepoints, collecting data from a total of 200 participants via Amazon’s Mechanical Turk website. A total of 101 participants completed responses at both time points, but 3 had to be excluded because they did not complete one of the two subscales of the TPES, leaving a final sample of 98 participants. Although this was slightly less than the desired sample of 101, a second power analysis revealed that the power was still acceptable at just under .95, well above the minimum threshold of .80.

The sample was mostly White (84.7%), and made up of women (59.2%; M\(_{age}\) = 31.96, SD = 10.66). Based on the political identifications requested at the beginning of the survey, 69.4% of the respondents identified as Democrat or Liberal, and the remaining 29.6% identified as Republican or Conservative.
Procedure

Participants completed the political groups TPES and several of the measures described in Study 1 (Social Dominance Orientation, Dehumanization, Social Desirability, Trait Empathy) before reporting demographics. The ingroup and outgroup subscales were presented in counterbalanced order at each timepoint. Fourteen days later, participants were contacted to complete the measures a second time, and were given 48 hours to respond from the time the second survey was initially posted to the MTurk website. The 48-hour window was provided to allow ample time for respondents to view the message, as not all MTurk users may log in daily. However, nearly all participants ( > 90%) completed the survey in less than 24 hours.

Test-retest reliability has been assessed in the literature through timespans ranging from days (Egger et al., 2006; Miller et al., 2002) to years (Schatz, 2010) dependent upon the construct being examined. I chose a 14-day period between timepoints 1 and 2 to allow for enough time to lapse so that participants might not fully recall the items in the measure. This is similar to other timeframes used in research on other similar personality measures such as the Portrait Values Questionnaire (Sandy, Gosling, Schwartz, & Koelkebeck, 2017).

Current Events Items. In addition to the measures above, participants also responded to the following four items asking if they had recently come across any information in the news that has changed their opinion toward their selected outgroup. All items were rated on a 1 (Strongly Disagree) to 5 (Strongly Agree).

Item 1: I have recently read/watched a news item containing information that has made me feel more positively towards _______ .

---

1 This data was collected for purposes beyond the scope of this dissertation, and the results will not be presented here.
Item 2: I have recently read/watched a news item containing information that has made me feel more negatively towards _______.

Item 3: I have recently had a positive encounter with a person who is _______.

Item 4: I have recently had a negative encounter with a person who is _______.

**Results**

In Study 3, one-way random intraclass correlations between time 1 and time 2 scores on the TPES were used to measure temporal consistency of the measure. For the purposes of this analysis, an intraclass correlation of .70 was considered the minimum acceptable threshold for ascertaining adequate stability of the measure over time. ICC values of .70-.79 represent moderate reliability, while .80-.89 are classified as good, and those greater than .90 are labeled excellent (Koo & Li, 2016). The reason for using the ICC statistic rather than a simple Pearson correlation is because the ICC is intended to assess reliability, stability and consistency of items, whereas the Pearson statistic simply captures whether the data is related in a linear fashion.

The measure showed good reliability for both subscales across time. For the ingroup subscale, the ICC (97, 97) = .88, 95% CI = .82 to .92, and for the outgroup subscale, the ICC (97, 97) = .86, 95% CI = .79 to .90. From these analyses, we can conclude that the TPES exhibits strong temporal stability.

I then used a repeated measures ANOVA to assess whether there was a significant difference between subscales while controlling for any positive or negative experiences participants may have had in between timepoints. There were no significant differences between participants scores on the TPES across time for either the ingroup subscale, $F(1, 93) = 1.30, p = .26$, or the outgroup subscale $F(1, 93) = 1.10, p = .30$. 
Discussion

Test-retest reliability is a crucial component of a personality measure, as personality traits, by definition, tend to be stable over time. Thus, in order to make the claim that the TPES is a meaningful measure of parochial empathy as a trait, it was important to establish that the scale shows consistency across time. The scale showed strong reliability across two timepoints. Importantly, levels of parochial empathy were not affected by current events, as there was no significant difference between the timepoints even when controlling for positive and negative encounters with outgroup members. This suggests that although parochial empathy differs from trait empathy, it can still be assessed at the dispositional level, and that the TPES has the capability to do so.

One limitation of Study 3 is the use of political affiliation as a means of establishing ingroups and outgroups. It is important to note that at the time of data collection², there were no large-scale events (e.g., a national election) that may have brought about differences in people’s levels of parochial empathy towards political groups. If an event such as this had taken place, people’s political identities may temporarily have been more salient, and there may have been an unexpected fluctuation in the data that could not be accurately captured with only two timepoints. However, an event like an election potentially is a situational factor that might produce only a temporary fluctuation in one’s state level of parochial empathy, and may not actually be indicative of a change at the trait level. This is why it was important to control for any recent events that individual participants may have experienced in between timepoints. It may be beneficial for future research to assess the temporal stability of the TPES using additional timepoints spread across a longer period to establish whether there are any significant

² All data was collected between May 29th-June 17th, 2019.
increases or decreases in parochial empathy as a function of current events. However, the chosen time period of two weeks was consistent with personality and related research.

With the completion of studies 1, 2, and 3, the factor structure of the TPES has been confirmed, and the scale has been established as a reliable measure. However, in all three studies, political affiliation was used to determine ingroup and outgroup identity. In the next study, I hope to establish that the factor structure of the TPES can once again be replicated when examining different types of ingroup and outgroup identities.
Study 4

Study 4 assessed the generalizability of the TPES to non-political group identities. Generalizability is an important component of any measure, as it assesses the extent to which the scale can be reliably used across a variety of different conditions and among samples from varying populations. Moreover, the intent of these scales is to versatility apply to most ingroup-outgroup pairs. A new version of the scale was developed and tested, this time assessing race as the ingroup/outgroup identity categorization (see Appendix A). If a similar factor structure is found amongst this new group categorization, then it can be said that the TPES can be used reliably among populations from different ingroup/outgroup designations (Gregorich, 2006).

Method

Participants

Data was collected from 414 U.S.-residing, English-speaking adults recruited from Amazon M-Turk. Participants who completed Studies 2 or 3 were not eligible. The sample was mostly women (57.0%; \( M_{\text{age}} = 31.82, SD = 10.25 \)). Based on the racial identifications requested at the beginning of the survey, 70.0% of the respondents identified as White, 12.6% as Black/African-American, 8.2% as Asian, 7.9% as Hispanic/Latino, and < 1% identified as Middle Eastern or Native American. Middle Easterners were identified as the most common outgroup (31.6%), followed by Black/African-American (24.2%), Asian (18.4%), White (9.7%), Native American (9.7%), and Hispanic/Latino (6.5%).

Procedure

A new version of the TPES was posted to Amazon MTurk reflecting racial identity categorizations. The items were identical except that the ingroup/outgroup options reflected racial groups rather than political affiliations (e.g., White, Black, Asian, etc.). Participants
completed the refined version of the TPES and the same convergent and discriminant validity measures as in Studies 1 and 2.

**Results**

**Confirmatory Factor Analysis**

I used a confirmatory factor analysis (CFA) with mean-likelihood (ML) estimation, with the intent of replicating the factor structure found in Study 2. Assumptions of skewness and kurtosis were checked and met for each of the manifest variables for both analyses. However, the ingroup analysis produced 8 multivariate outliers, leaving a final sample size of 406 participants, and the outgroup analysis produced 12 multivariate outliers, for a final sample of 402 participants.

**Ingroup Confirmatory Factor Analysis.** I used the same criteria as in Study 2 to assess model fit: first, all of the scale items loaded significantly onto their respective first-order factors (all r’s > .40, p’s < .001), Second, the individual factors were not correlated with one another above r = .70, as this could indicate the presence of multicollinearity among the subscales of the measure. It appeared that the Motivated Empathy and Affective Sharing factors were highly correlated, r = .86, p < .001, but all others were significantly correlated within a normal range, as in Study 2. Additionally, although Empathic Accuracy and Non-Empathic Responding factors shared a significant relationship (r = .25, p < .001), the correlation was significantly lower than the ones shared between all other factors (all r’s ranging from .51-.86, p’s < .001). Because these two factors were not significantly related in Studies 1 and 2, it was interesting to note that they still shared the smallest relationship of all the factors here.

In addition to the correlations between items and factors, I once again examined the same set of fit indices as previously. The first were the absolute fit indices: The $\chi^2$ goodness-of-fit test...
was significant, $\chi^2 (224) = 622.93, p < .001$. However, as discussed in Study 2, this statistic tends to be significant in large samples due to the sample size, and so there are more accurate absolute indices that should be examined. The GFI value was .88, indicating that the model fit was less than ideal. The Comparative Fit Index (CFI) of the initial model was .89, indicating a slightly less than adequate fit. However, the root mean square error of approximation (RMSEA) was .066, indicating a good fit, although a value $\leq .05$ is ideal. The value of .066 was notably lower (i.e., improved) than the initial CFA performed in Study 2.

Given the good RMSEA value and the remaining indices that were less than ideal, though approaching acceptable levels, I concluded that the model was adequate. However, I still wanted to use the same respecifications (correlating related error variances within the same factors) that were used in Study 2 as this was intended to be a replication. After these respecifications, the GFI improved very slightly to .89, which is still less than ideal. However, the CFI rose to .91 and the RMSEA improved from .066 to .063, both indicating a good fit. Overall, the respecified model indicated a slightly improved fit from the original (see Figure 7 for final model).
Outgroup Confirmatory Factor Analysis. As in Study 2, the chi-square for the outgroup subscale was once again significant and much higher than for the ingroup subscale, $\chi^2(224) = 811.24, p < .001$. The GFI and CFI did not approach acceptable levels, at .85 and .84, respectively. However, the RMSEA was exactly .08, indicating an adequate model fit. This initial model provided a similar, though slightly better, fit than in Study 2.

After the respecifications, the GFI was .89, which was still slightly less than ideal. However, the CFI rose to .90 and the RMSEA improved from .080 to .065, both indicating a marginally good fit. Overall, with taking all of the aforementioned statistics into consideration, but giving priority to the RMSEA, it appears that the respecified model indicated a much-improved fit from the original (see Figure 8 for final model).
Convergent/Discriminant Validity

As in Studies 1 and 2, hierarchical linear regression was used to assess the relationship between the TPES and the pre-test scale measures while controlling for levels of social desirability (see Table 7 for full results of this analysis). All assumptions were verified and met prior to running this analysis, including univariate and multivariate normality, an absence of multicollinearity, and homoscedasticity.

The pattern of results was similar to those in Studies 1 and 2, with a few exceptions. As in the previous studies, social desirability did not predict TPES ingroup scores, but it did positively predict scores on the outgroup subscale. Dehumanization once again positively
predicted TPES scores on the ingroup subscale and negatively predicted them on the outgroup subscale, providing further evidence that empathy is not allocated equally amongst ingroup and outgroup members, and that there is the potential for extremely negative outcomes as a result of this difference. Aggression was unrelated to both subscales once again. This time, however, malicious envy negatively predicted scores on the ingroup subscale of the TPES, but was unrelated to outgroup empathy. This relationship between envy and empathy did not appear in any of the previous studies, and may be a function of the types of groups being compared (political in Studies 1 and 2, racial in Study 4).

As in Study 1, trait empathy only significantly predicted outgroup scores in this sample. Unlike in the previous studies, social dominance orientation negatively predicted both ingroup and outgroup scores. As in Study 2, openness did not predict scores on either subscale. However, agreeableness was once again significant positive predictor of both ingroup and outgroup empathy scores. This pattern of results, though not completely identical to the studies that preceded them, is largely consistent with them and underscores the idea that there are important differences in the ways empathy is considered towards ingroups versus outgroups.

Table 7.

Regression Weights – Study 4

<table>
<thead>
<tr>
<th>Variable (N = 414)</th>
<th>M</th>
<th>SD</th>
<th>β, p</th>
<th>β, p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Ingroup</td>
<td>Outgroup</td>
</tr>
<tr>
<td>Social Desirability</td>
<td>6.55</td>
<td>2.73</td>
<td>-.05, .32</td>
<td>.14, .005*</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>3.80</td>
<td>.78</td>
<td>.32, &lt;.001*</td>
<td>.28, &lt;.001*</td>
</tr>
<tr>
<td>Openness</td>
<td>3.46</td>
<td>.51</td>
<td>.04, .41</td>
<td>.06, .19</td>
</tr>
<tr>
<td>Extraversion</td>
<td>2.69</td>
<td>.94</td>
<td>-.06, .19</td>
<td>.05, .31</td>
</tr>
</tbody>
</table>
### Discussion

Using new comparison categories, the results of Study 4 replicated the CFA from Study 2 very closely. Using the exact factor structure confirmed in Study 2, I was able to achieve models that indicated marginally good fits for both the ingroup and outgroup subscales. The results were also similar to Study 2 in that the initial and final models of the ingroup subscale exhibited overall better fit indices relative to outgroup subscale.

However, it seemed that the ingroup subscale seemed to exhibit a stronger fit than in Study 2, even prior to respecification. This may have been due to the identity categories that were used, as perhaps racial identity is a more central tenet of the self than political affiliation. That is to say that our race is unchangeable and typically a discrete category, whereas political beliefs can fluctuate and some people may hold beliefs that overlap with both major political parties. This is also in line with recent work which suggests that other racial groups specifically may be seen as rivals and sources of economic and social competition, therefore resulting in feelings of envy (Richins et al., 2019). However, the research also suggests that these feelings are moderated by whether the target group is seen as admired or advantaged in some way. Future work might examine whether levels of parochial empathy are influenced by feeling

<table>
<thead>
<tr>
<th></th>
<th>3.58</th>
<th>0.84</th>
<th>0.05, 0.34</th>
<th>-0.04, 0.45</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conscientiousness</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Empathy</td>
<td>2.98</td>
<td>0.45</td>
<td>0.09, 0.09</td>
<td>0.17, &lt;.001*</td>
</tr>
<tr>
<td>Malicious Envy</td>
<td>2.46</td>
<td>1.08</td>
<td>-0.17, 0.002*</td>
<td>0.02, 0.66</td>
</tr>
<tr>
<td>Aggression/Hostility</td>
<td>2.45</td>
<td>0.67</td>
<td>0.09, 0.13</td>
<td>0.01, 0.90</td>
</tr>
<tr>
<td>Social Dominance</td>
<td>2.52</td>
<td>1.17</td>
<td>-0.12, 0.03*</td>
<td>-0.23, &lt;.001*</td>
</tr>
<tr>
<td>Dehumanization</td>
<td>0.27</td>
<td>0.80</td>
<td>0.14, 0.005*</td>
<td>-0.16, &lt;.001*</td>
</tr>
</tbody>
</table>
disadvantaged relative to another outgroup. Additionally, it would be interesting to examine how parochial empathy functions when considering individuals belonging to multiple racial groups (or other outgroup categorization of interest). There is evidence to suggest that holding multiple outgroup categorizations humanizes a target, making them more complex and less homogenous, and thus reduces prejudicial attitudes against them (Brewer, 2000; Levy et al., 2019). In this vein, it is possible that empathy may also be allocated more evenly towards individuals who identify as bi- or multi-racial relative to others.

Now that the factor structure and convergent/discriminant validity of the TPES has been mostly replicated using a new sample and different identity demographics, the next step is to determine whether the measure has predictive validity. The focus of Study 5 will be to establish whether scores on the TPES can predict behavior towards ingroup and outgroup members.
Study 5

The goal of Study 5 was to determine the utility of the Trait Parochial Empathy Scale in predicting behavior over and above a standard empathy measure such as the Interpersonal Reactivity Index (Davis, 1983). Predictive validity is the extent to which a measure can be relied upon to predict an outcome. This is considered one of the most important types of validity, especially when considering personality traits, as it can allow us to ascertain how an individual may perform on a specific related criterion measure or how they may behave under certain conditions (Roberts et al., 2007; Rauthmann, Sherman, Nave, & Funder, 2015). Further, if a scale can map onto observable and theoretically relevant behaviors, then this provides converging evidence that the items are measuring the construct as intended.

In the current study, TPES scores were used to predict the likelihood of students donating to different campus groups as a function of their ingroup and outgroup designation. Although the students all belonged to the shared ingroup of “VCU Student,” the manipulation attempted to appeal to their other, perhaps more salient ingroup identities. Identity salience is an individual’s awareness of a particular identity and how important it is to their self-concept. On a larger scale, this is somewhat comparable to how individuals in the United States consider themselves “Americans,” but are often much more aware of their political, racial, religious, or gender identities, and thus find them much more personally important.

This study employed a manipulation inspired by the Minimal Groups Paradigm (Tajfel, 1970), which has been used in previous research to showcase that invoking social categories even around trivial or superficial aspects of one’s identity (e.g., preference for a particular artist) is enough to elicit preferential bias for one’s ingroup relative to an outgroup (Diehl, 1990; Gaertner & Insko, 2000). This manipulation also requires allocation of some sort of concrete...
resource between groups, with the understanding that the participants will not be directly benefiting from those resources. In the context of Study 5, students were asked to allocate money between two groups – republicans and democrats.

I expected that participants who felt more empathy for their ingroup would exhibit this via their decisions to make more charitable donations to support a group they identify with relative to an outgroup. In a real-world context, this is similar to how people may be more likely to attend to a cause that has personally affected them (e.g., an individual who has lost a loved one to Alzheimer’s may be more likely to donate to an Alzheimer’s-related charity versus another type of organization). However, whereas in the previous example a charitable donation was still made, and thus the outcome can still be considered a positive one, parochial empathy in a different context can have much more negative consequences. For example, a juror who is less likely to feel empathy for a defendant of a different race may judge them much more harshly. If parochial empathy exists at the trait level, it can exert a negative influence upon interactions between people of different social groups, some of which can have devastating real-world implications. Thus, being able to more accurately predict how one might behave when engaging with such individuals can be invaluable, especially if the outgroup member is in a position that should elicit empathy.

Method

Participants

An *a priori* power analysis conducted using G*Power software (Faul, Erdfelder, Buchner, & Land, 2009) and assuming a small effect size as a conservative estimate ($r^2 = .09$), revealed that 100 participants would be sufficient to detect an effect (power $\geq 0.95$, alpha $\leq 0.05$). Due to time constraints, 87 participants were sampled. However, 3 participants’ data had
to be removed due to carelessness (n = 2) and one (n = 1) due to a computer not recording responses correctly, leaving a final sample of 84 participants.

All participants were college students recruited using VCU’s undergraduate research pool. The sample was mostly women (69.0%; $M_{age} = 19.22$, $SD = 1.24$), but was reasonably diverse in terms of racial identification. Approximately 40.5% of participants identified as White, 20.2% as Black/African-American, 19.0% as South Asian, 13.1% as more than one race, and the remaining 7.2% as Other. Based on the political identifications requested at the beginning of the survey, 80.9% of the respondents identified as Democrat or Liberal, and the remaining 19.1% identified as Republican or Conservative.

**Procedure**

Participants were told they were completing a study to determine which on campus groups put forth the most persuasive funding campaigns. After consenting, participants were informed that they would be viewing videotaped appeals from students belonging to various groups on campus and would then select the appeal they found the most convincing, and thus, worthy of funding. Participants were told that the funding came from an alumni grant, with one of the stipulations of the grant being that students get to decide how the funding gets allocated across groups.

Participants were told that due to the high number of groups participating, they would only view two videos of groups from within the same category and would be asked to answer questions specifically relating to the groups in their assigned category. They were then given a randomly assigned number via an online random number generator and told that they had been assigned to the view video appeals from the political group’s category. In reality all participants viewed the same groups (e.g., College Republicans and Young Democrats, which are two active
groups at Virginia Commonwealth University). To increase believability, the experimenter checked a list of various groups to verify which assignment the participant would be receiving. This cover story was also intended to increase face validity of the study, to provide a rationale for why participants were answering the items on the political groups TPES in relation to this persuasive messaging competition.

**Trait Measures.** Participants completed the personality, trait empathy, and social desirability measures from Study 1, as well as the political groups TPES. The Dehumanization, Envy, and Aggression/Hostility measures were not included, as they were deemed inappropriate given the cover story. Similar to Studies 1-4, the ingroup and outgroup subscales of the TPES were presented in counterbalanced order at the beginning and end of the set of measures, with all other scales placed in between.

**Video Appeal.** Participants viewed two pre-taped video appeals from student members of various organizations (in reality the members were research assistants). In each video, the student explained the purpose of their organization, and explained why funding would be valuable to them. Each of the two research assistants taped two versions of the video—one for the College Republicans and another for the Young Democrats. There were two different research assistants representing the two organizations, and the videos were counterbalanced to ensure that both the research assistance and the organization they represented was randomized, for a total of four possible combinations.

Additionally, in order to invoke empathy, the experimenter informed participants that these student groups have decided to take part in this fundraising campaign due to a shortage of funds, and not simply as part of a competition. Further, they would be told that any student group
who was unable to obtain a minimum amount of funding will be unable to run their scheduled events in the following semester.

Participants were further told that these student-led groups are in need of funding and were given the opportunity to allocate $15 across groups, again with the understanding that if a group did not receive adequate funds, it would be unable to run events the following semester. Participants would be required to allocate the entirety of the money, and the web-based survey was set up to ensure that the total amount of money was assigned among the two groups. The names of the two organizations were also counterbalanced at this point in the survey, to avoid any order effects when participants were making their donations. Lastly, participants responded to questions designed to assess their motivations for the allocation. They answered an open-ended question about why they made their allocation choices, followed by specific questions (e.g., “I wanted to give the highest amount of money to the group I support the most,” “I wanted to help all of the groups meet their funding goals equally,” and “I gave less money to the group I support the least.”) and demographics (see Appendix K for a full list of questions).

Additional items were included to assess which video appeal participants found most persuasive in an effort to increase the face validity of the manipulation by alluding back to the initial cover story. For example, participants responded to items such as “Which of the videos did you find most persuasive?”, along with related follow-up questions including, “You indicated that you found the video appeal from the _____ most persuasive. Please elaborate on what parts of their message appealed to you most”, “I might consider joining the _____”, and “I would be interested in attending a future event held by the ______”.

Suspicion Check. Due to the use of a cover story as a part of the manipulation, it was crucial to probe the participants to assess whether they had any suspicions about the true purpose
of the study. After participants completed the online questionnaire items, the experimenters asked them to describe what they thought the study was about in their own words. At this time, the experimenter noted whether any participants expressed suspicions about the purpose of the study or had guessed the hypotheses. No participants did so, and therefore all 84 of the participants who passed the attention check measures were included in the final analyses. Participants were then debriefed and the study session was concluded.

Results

Manipulation Check

Participants selected the video appeal from the Young Democrats of VCU as the most persuasive video 82% of the time, which was almost exactly as many identified as Democratic or Liberal. This did not differ based on the order in which the videos were presented, $F(3, 81) = 2.00, p = .12$, suggesting that participants were likely convinced based on the presumed political affiliation of the student in the video rather than the message itself.

Additionally, the more strongly a participant agreed with the statement “I wanted to give the highest amount of money to the group I support the most,” the greater amount of money they gave to the school organization that represented their ingroup relative to the one that represented an outgroup, $t(83) = 5.56, p < .001$. Conversely, if participants reported a stronger desire to help all the groups equally, there was a significantly lower difference between the amounts they donated $t(83) = -4.08, p < .001$. Participants’ reported intentions accurately reflected the differences in their donations amounts, suggesting that the manipulation was, in fact, successful.

Principal Dependent Measures

Separate ingroup donation and outgroup donation scores were calculated as outcome measures. More specifically, if participants identified as Democrat or Liberal, the amount of
money they donated to the Young Democrats of VCU was considered their “ingroup donation” and any money donated to the College Republicans of VCU was considered their “outgroup donation”. The scores were calculated in the opposite direction for those who identified as Republican or Conservative.

I then used a hierarchical regression analyses to determine whether participant scores on the ingroup and outgroup scales of the TPES uniquely predicted their funding allocations over and above trait empathy and social desirability. I also decided to include social desirability in the first block as this had been used as a control variable in the previous studies. Both the ingroup and outgroup TPES scores were entered as predictors in Block 2, with the ingroup donation amount as the dependent variable. Block 1 was not significant, $R^2 = .01, F(2, 83) = .59, p = .55$, with neither empathy ($\beta = .02, t(83) = .15, p = .88, 95\% CI [-2.06, 2.39]$ nor social desirability ($\beta = -.12, t(83) = -1.08, p = .28, 95\% CI [-.53, .16]$ predicting the amount of money donated to participants’ ingroups. Further, when TPES ingroup and outgroup scores were added into the model, they did not add significant predictive value, $\Delta R^2 = .02, \Delta F(2, 79) = .83, p = .44$. When examined individually, neither scores on the ingroup subscale, $\beta = .13, t(83) = 1.07, p = .29, 95\% CI [-.89, 2.96]$, nor outgroup subscale $\beta = -.11, t(83) = -.92, p = .36, 95\% CI [-.52, .93]$ predicted the amount of money donated. The overall model effect size was very small, $r^2 = .04$, with TPES scores explaining only 4% of the variance in ingroup donation amounts.

Next, I repeated the analysis with outgroup donation amount as the dependent variable. Once again, Block 1 containing trait empathy and social desirability was not significant, $F(2, 83) = .65, p = .52$, and Block 2 containing the TPES subscales did not add significant predictive value, $\Delta R^2 = .02, \Delta F(2, 79) = .92, p = .40$. Also, as with the ingroup donations, neither scores on the ingroup subscale, $\beta = -.13, t(83) = -1.11, p = .27, 95\% CI [-2.99, .85]$, nor outgroup subscale
β = .12, \( t(83) = 1.00, p = .32, 95\% \text{ CI } [ - .86, 2.58] \) predicting the amount of money donated. The model effect size was once again very small, \( r^2 = .04 \), with TPES scores explaining only 4% of the variance in outgroup donation amounts.

Although the TPES did not predict behavior on its own, this may have been due to the underpowered sample. It was interesting to note that overall people donated significantly more to the Young Democrats (\( M = 10.61, SD = 3.76 \)) than to the College Republicans (\( M = 4.37, SD = 3.76 \)), \( t(83) = 7.61, p < .001, 95\% \text{ CI } [4.61, 7.87] \), which aligned with the strongly democratic/liberal skew of the sample. This makes it especially likely that the lack of power contributed to the null results, given the skew in the donations made, and also the fact that neither trait empathy nor social desirability were able to predict donation amounts on their own.

**Discussion**

The results of this study did not provide enough evidence to suggest that the TPES was a significant predictor of behavior. However, it should be noted that the sample was underpowered, and the obtained effect size was smaller than anticipated. While there was a clear imbalance in the donations that aligned with students’ stated political preferences, there was not enough data collected at this time to determine whether the TPES was able to accurately predict behavior. Thus, additional data collection will be crucial.

Apart from the sample size, another potential limitation of the study could be the way in which the outcome was assessed. While using money as an outcome variable is a close simulation of what would happen in a real-world scenario, the experiment was limited by the fact that it was in a lab-setting and thus not the participants’ own money they were donating. Thus, participants may not have been as personally invested in the outcome since the money was not their own, and since they themselves do not actually belong to these organizations in the way
that a voter might belong to a political party. Therefore, we might see a stronger effect if it were the participants’ own money and if the personal stakes were higher. Also, the students may have seen their individual $15 amount as being relatively low relative to the total amount of money that would be donated across all students, leading them to believe their donation would not make that much of a difference and was therefore inconsequential. For future data collection, it might be helpful to increase the amount of money students believe they will be donating, or perhaps to include an item assessing how helpful participants feel their donation might be to their chosen organization. Additionally, in order to improve the external validity of this study, future work might look at assessing a person’s level of parochial empathy as well as their actual donations made to political candidates (or other groups of interest).

Importantly, despite the small sample size and somewhat limited external validity, the data was still trending in expected directions. More specifically, the ingroup subscale of the TPES was positively related to ingroup donations and negatively related to the monetary amounts allocated towards the outgroups, and this pattern was mirrored with the outgroup subscale. If the pattern of results continues along this path when additional data is collected, then I will be able to establish the predictive validity of the TPES.

General Discussion

Through five studies, I refined and validated a novel measure of dispositional parochial empathy by assessing the properties of the TPES across various samples and group membership categories. In Study 1, I established that parochial empathy seems to be divided into four distinct factors: Empathic Accuracy, Motivated Empathy, Non-Empathic Responding, and Affective Sharing. These four factors were consistent when examining how this trait is exhibited towards both ingroup and outgroup members, though there were some differences in how items loaded
onto each factor that reflected the differences in how we allocate empathy towards different individuals. In Study 2, I confirmed the four-factor structure of the TPES, though the model required some respecification. In Study 3, I established that the TPES can reliably measure levels of parochial empathy across time (i.e., 14 days). In Study 4, I once again confirmed the factor structure of the TPES using different ingroup/outgroup categorizations to show that the scale is generalizable across different demographics. The previous studies had all used political affiliation for the ingroup and outgroup categorizations, whereas Study 4 used racial groups. This was important, as it provides evidence that the TPES can be utilized to assess empathy towards a variety of ingroup and outgroup pairs across demographic variables. Finally, in Study 5, I assessed the predictive validity, but was unable to demonstrate that the scale offers predictive validity. However, this was likely due to a smaller than anticipated sample size, and all of the results were still trending in the predicted directions.

Cognitive and affective empathy can be a powerful predictor of intergroup cooperation (Batson & Ahmad, 2009), but conversely, empathy is often most difficult when empathy targets are perceived to be dissimilar (Stumer, Snyder, Kropp, & Siem, 2006). The TPES potentially can be used to facilitate the measurement of dispositional-level empathy for ingroup and outgroup members, which is a necessary step in identifying and understanding factors which influence intergroup communication, cooperation, and hostility. These studies have allowed me to form a solid understanding of the scale’s properties and limitations, which I will build upon in future research on intergroup empathy and social identities. I anticipate that use of this scale will help us understand factors that facilitate or hinder intergroup empathy, and ultimately inform interventions to reduce negative prejudicial attitudes towards outgroup members and improve intergroup relations.
An additional strength of the studies comes from the design of the scale, which lends itself to potentially being able to categorize participants into one of four possible subgroups: Overall High Empathy (high ingroup/high outgroup), Antipathy (low ingroup/low outgroup), Altruistic Empathy (low ingroup/high outgroup), and Parochial Empathy (high ingroup/low outgroup). While this was not a main goal of the present series of studies, it would be interesting to examine whether there are discernable patterns between participants who fall into these four group categorizations. For example, given the relationship between outgroup empathy and dehumanization, would people in the “Parochial Empathy” category be more likely to actively choose to engage in harmful behaviors against outgroup members? Similarly, would individuals who can be described as “Altruistic Empathy” exhibit similarly negative behavior towards people in their own groups? Further, what would be the motivations for holding negative attitudes and/or exhibiting such negative behaviors towards one’s own group?

It is important to note that these studies pose several limitations. First, because participants were recruited consistently from the populations of college students and MTurk workers from the United States, the participant samples may not be generalizable to global populations. This is due to the nature of some of the categories selected, which are based on American group designations (e.g., U.S. political parties and racial designations). I hope to address this in future studies by assessing the TPES in samples that reflect group categories relevant to different cultures and nations. Additionally, for the studies involving political affiliation categories, it might be helpful to screen for participants who do not follow politics but still chose to participate. Although I included an item assessing how important a person’s political affiliation was to them, it would be helpful to capture whether participants are truly not engaged in the political process, or just generally do not feel it is an important part of their
identity. This would also be important for future work examining other categories such as religious affiliation, where a person might have been raised as a particular religion, but may not actively participate in any of its traditions nor hold any of its beliefs.

Another limitation is that this scale only deals with one facet of a person’s identity at a time. The TPES is designed to assess parochial empathy at the trait level, and so while I expected that this would be a relatively stable predictor of other attitudinal and behavioral outcomes, it is also important to consider the implications of holding multiple identities. While the scale and its questions were designed to have participants consider just one of their many identities, there is a theoretical framework that speaks directly to the importance of considering multiple aspects of identity simultaneously. Intersectionality Theory posits that a person’s various identities interact with and are reinforced by one another (Thomas & Crenshaw, 2004). Further, by its very definition, intersectionality tells us that we cannot disentangle an individual’s various identities from one another. Consider an individual who identifies as White, heterosexual, and a woman. To whom might their empathy be allocated more strongly—those who are of the same racial group, those who share the same sexual identity, or might it be those who are of the same gender? In the scale’s current form, the demographic questions will include an assessment of the importance of the specific identity being assessed in the TPES, and this item can be used to control for identity centrality. Although examining the interaction of identities is beyond the scope of this dissertation, this is a promising avenue for future exploration on this topic.
References


APPENDIX A
Parochial Empathy Measure (TPES) – developed by author

First, identify which of the following groups you identify with the MOST, and then identify which group is LEAST like the group you belong to.

<table>
<thead>
<tr>
<th>Political Affiliation (S1-3, 5)</th>
<th>Race/Ethnicity (S4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liberal</td>
<td>White/Caucasian</td>
</tr>
<tr>
<td>Conservative</td>
<td>Black/African-American</td>
</tr>
<tr>
<td>Democrat</td>
<td>Hispanic/Latino</td>
</tr>
<tr>
<td>Republican</td>
<td>Middle Eastern</td>
</tr>
<tr>
<td></td>
<td>Asian</td>
</tr>
<tr>
<td></td>
<td>Native American</td>
</tr>
</tbody>
</table>

Participants report which of several groups they identify with MOST, and which group they identify with LEAST. They are shown the following scale twice (see next page), first with their ingroup labels in the blanks, and second with the outgroup label in the blanks.

Not at all like me
Neither like me nor dislike me
Very much like me

1  2  3  4  5

Affective Sharing/Emotional Contagion
1. When I see a _________ person get excited, I get excited too.
2. I find it annoying when _______ people get excited.
3. I can tell when a person who is a _________ is upset even when they don’t say anything.
4. I have a hard time feeling “in tune” with the feelings of people who are _______.
5. When I see people who are __________seem anxious, I get anxious as well.
6. When I’m talking with someone who is a _____________, I tend to feel the same emotions they are feeling.
7. It makes me happy to see people who are _______ happy.

Empathy in action/empathic concern
1. It upsets me to see a person who is a ________ being treated disrespectfully.
2. I enjoy making people who are __________ feel better.
3. When people who are __________ start to talk about their problems, I try to change the topic to something else.
4. I’m not really interested in how __________ feel.
5. People who are __________ don’t deserve my compassion.
6. I often feel bad about the problems experienced by people who are ________.
7. I feel sorry for the way that __________ get taken advantage of.
8. I take part in activities to try and help give a voice to people who are __________.
Empathic accuracy
1. I can often tell when someone who is a _______ is hiding their true emotions.
2. I am quick to spot when someone who is a ___________ is feeling awkward or uncomfortable.
3. It’s easy for me to tell when a _____________ is interested in what I’m saying.
4. _____________ people tell me I’m good at understanding what they are thinking and feeling.
5. I can usually figure out how ______________ people are feeling before they tell me.
6. I’m not usually aware of the feelings of people who are ______ .

Perspective-taking
1. When I think I’m about to criticize someone who is a ____________, I first try to imagine how I would feel if I was in their place.
2. I can usually appreciate a _____________’s viewpoints, even if I don’t agree with them.
3. Even when I try to consider the perspective of a ____________, I find that I just can’t understand them.
4. When people who are ______________ tell me they are taken advantage of, I try to understand why they feel that way.
5. It seems to me that people who are _______ are treated fairly in society.
6. When I read stories about people who are ___________ in the news, I try to imagine what I would feel like in their position.
7. I don’t believe people who are _______ when they say they are disadvantaged.
APPENDIX B
Mini-IPIP

Here are a number of characteristics that may or may not apply to you. For example, do you agree that you are someone who likes to spend time with others? Please choose the number next to each statement to indicate the extent to which you agree or disagree with that statement.

<table>
<thead>
<tr>
<th></th>
<th>Disagree Strongly</th>
<th>Disagree a little</th>
<th>Neither agree nor disagree</th>
<th>Agree a little</th>
<th>Agree strongly</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Am the life of the party.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Sympathize with others’ feelings</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Get chores done right away.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Have frequent mood swings.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Have a vivid imagination.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Don’t talk a lot. (R)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Am not interested in other people’s problems. (R)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Often forget to put things back in their proper place. (R)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Am relaxed most of the time. (R)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Am not interested in abstract ideas. (R)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Talk to a lot of different people at parties.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Feel others’ emotions.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Like order.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Get upset easily.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Have difficulty understanding abstract ideas. (R)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Keep in the background. (R)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Am not really interested in others. (R)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>Make a mess of things. (R)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Seldom feel blue. (R)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>Do not have a good imagination. (R)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

APPENDIX C
Social Dominance Orientation Scale – 8 item

Instructions: Show how much you favor or oppose each idea below by selecting a number from 1 to 7 on the scale below. You can work quickly; your first feeling is generally best.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Strongly Oppose</td>
<td>Somewhat Oppose</td>
<td>Slightly Oppose</td>
<td>Neutral</td>
<td>Slightly Favor</td>
<td>Somewhat Favor</td>
<td>Strongly Favor</td>
</tr>
</tbody>
</table>

Protrait dominance:
1. An ideal society requires some groups to be on top and others to be on the bottom.
2. Some groups of people are simply inferior to other groups.

Contrait dominance:
3. No one group should dominate in society.
4. Groups at the bottom are just as deserving as groups at the top.

Protrait anti-egalitarianism:
5. Group equality should not be our primary goal.
6. It is unjust to try to make groups equal.

Contrait anti-egalitarianism:
7. We should do what we can to equalize conditions for different groups.
8. We should work to give all groups an equal chance to succeed.

Note: The con-trait items should be reverse-scored before computing a composite scale mean.

Citation

APPENDIX D
Ascent Measure of Dehumanization

People can vary in how human-like they seem. Some people seem highly evolved whereas others seem no different than lower animals. Using the image below, indicate using the sliders how evolved you consider the average member of each group to be:

Note: The group labels associated with these sliders will be based on the ingroup and outgroup categories that participants select at the beginning of the study.

APPENDIX E

Brief Aggression Questionnaire (Webster et al., 2015)

Instructions: Indicate the extent to which each of these statements describes you.

<table>
<thead>
<tr>
<th>Statement</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Given enough provocation, I may hit another person.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>(Does not describe me at all)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>If I have to resort to violence to protect my rights, I will.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>(Does not describe me at all)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>There are people who pushed me so far that we came to blows.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>(Does not describe me at all)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I feel angry for a person when his or her feelings have been hurt by someone else.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>(Does not describe me at all)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I tell my friends openly when I disagree with them.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>(Does not describe me at all)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>When people annoy me, I may tell them what I think of them.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>(Does not describe me at all)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>My friends say that I’m somewhat argumentative.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>(Does not describe me at all)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I am an even-tempered person.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>(Does not describe me at all)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sometimes I fly off the handle for no good reason.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>(Does not describe me at all)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
I have trouble controlling my temper.  
1 (Does not describe me at all) 2 3 4 5 (Describes me very well)

Other people always seem to get the breaks.  
1 (Does not describe me at all) 2 3 4 5 (Describes me very well)

I sometimes feel that people are laughing at me behind my back.  
1 (Does not describe me at all) 2 3 4 5 (Describes me very well)

When people are especially nice, I wonder what they want.  
1 (Does not describe me at all) 2 3 4 5 (Describes me very well)

APPENDIX F
Feelings Thermometer Scale

How do you feel toward people who are [outgroup label] in general?

10 = Extremely warm/favorable
5 = Neither favorable nor unfavorable
0 = Extremely cold/unfavorable

How do you feel toward people who are [ingroup label] in general?

10 = Extremely warm/favorable
5 = Neither favorable nor unfavorable
0 = Extremely cold/unfavorable
APPENDIX G
Marlowe-Crowne Social Desirability Scale – 13 item

INSTRUCTIONS: For each item below, please indicate whether the statement is true of you or false of you by circling “T” or “F” respectively. If neither seems to apply exactly to you, then circle the answer which is closest to how you truly feel.

1. T F It is sometimes hard for me to go on with my work if I am not encouraged.
2. T F I sometimes feel resentful when I don’t get my way.
3. T F On a few occasions, I have given up doing something because I thought too little of my ability.
4. T F There have been times when I felt like rebelling against people in authority even though I knew they were right.
5. T F No matter who I’m talking to, I’m always a good listener.
6. T F There have been occasions when I took advantage of someone.
7. T F I’m always willing to admit it when I make a mistake.
8. T F I sometimes try to get even rather than forgive and forget.
9. T F I am always courteous, even to people who are disagreeable.
10. T F I have never been irked when people expressed ideas very different from my own.
11. T F There have been times when I was quite jealous of the good fortune of others.
12. T F I am sometimes irritated by people who ask favors of me.
13. T F I have never deliberately said something that hurt someone’s feelings.

APPENDIX H
The Benign and Malicious Envy Scale (BeMaS)

**English version**

*Below, you will find statements related to situations when you lack another's superior quality, achievement, or possession and you either desire it or wish that the other lacks it. Please indicate for every statement how much you agree or disagree with it. There are no right or wrong answers. Don’t hesitate to indicate the first answer that comes to your mind.*

**benign1**  
When I envy others, I focus on how I can become equally successful in the future.

**malicious1**  
I wish that superior people lose their advantage.

**benign2**  
If I notice that another person is better than me, I try to improve myself.

**benign3**  
Envyng others motivates me to accomplish my goals.

**malicious2**  
If other people have something that I want for myself, I wish to take it away from them.

**malicious3**  
I feel ill will towards people I envy.

**benign4**  
I strive to reach other people’s superior achievements.

**malicious4**  
Envious feelings cause me to dislike the other person.

**benign5**  
If someone has superior qualities, achievements, or possessions, I try to attain them for myself.

**malicious5**  
Seeing other people’s achievements makes me resent them.

APPENDIX I
Single-Item Trait Empathy Scale

To what extent does the following statement describe you?

“I am an empathetic person.”

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Not very true of me</td>
<td>Slightly true of me</td>
<td>Somewhat true of me</td>
<td>Moderately true of me</td>
<td>Very true of me</td>
</tr>
</tbody>
</table>

*Note:* An empathetic person understands others’ feelings, and experiences care and concern for them.

Citation: Konrath, S., Meier, B. P., & Bushman, B. J. (2018). Development and validation of the single item trait empathy scale (SITES). *Journal of Research in Personality, 73*, 111-122.
Appendix J
INTERPERSONAL REACTIVITY INDEX

The following statements inquire about your thoughts and feelings in a variety of situations. For each item, indicate how well it describes you by choosing the appropriate letter on the scale at the top of the page: A, B, C, D, or E. When you have decided on your answer, fill in the letter next to the item number. READ EACH ITEM CAREFULLY BEFORE RESPONDING. Answer as honestly as you can.

ANSWER SCALE

1  2  3  4  5
Does not describe me      Describes me very well

1. I daydream and fantasize, with some regularity, about things that might happen to me. (FS)
2. I often have tender, concerned feelings for people less fortunate than me. (EC)
3. I sometimes find it difficult to see things from the "other guy's" point of view. (PT) (-)
4. Sometimes I don't feel very sorry for other people when they are having problems. (EC) (-)
5. I really get involved with the feelings of the characters in a novel. (FS)
6. In emergency situations, I feel apprehensive and ill-at-ease. (PD)
7. I am usually objective when I watch a movie or play, and I don't often get completely caught up in it. (FS) (-)
8. I try to look at everybody's side of a disagreement before I make a decision. (PT)
9. When I see someone being taken advantage of, I feel kind of protective towards them. (EC)
10. I sometimes feel helpless when I am in the middle of a very emotional situation. (PD)
11. I sometimes try to understand my friends better by imagining how things look from their perspective. (PT)
12. Becoming extremely involved in a good book or movie is somewhat rare for me. (FS) (-)
13. When I see someone get hurt, I tend to remain calm. (PD) (-)
14. Other people's misfortunes do not usually disturb me a great deal. (EC) (-)
15. If I'm sure I'm right about something, I don't waste much time listening to other people's arguments. (PT) (-)
16. After seeing a play or movie, I have felt as though I were one of the characters. (FS)
17. Being in a tense emotional situation scares me. (PD)
18. When I see someone being treated unfairly, I sometimes don't feel very much pity for them.
19. I am usually pretty effective in dealing with emergencies. (PD)

20. I am often quite touched by things that I see happen. (EC)

21. I believe that there are two sides to every question and try to look at them both. (PT)

22. I would describe myself as a pretty soft-hearted person. (EC)

23. When I watch a good movie, I can very easily put myself in the place of a leading character. (FS)

24. I tend to lose control during emergencies. (PD)

25. When I'm upset at someone, I usually try to "put myself in his shoes" for a while. (PT)

26. When I am reading an interesting story or novel, I imagine how I would feel if the events in the story were happening to me. (FS)

27. When I see someone who badly needs help in an emergency, I go to pieces. (PD)

28. Before criticizing somebody, I try to imagine how I would feel if I were in their place. (PT)

Appendix K
General Information – Demographics

1) What is your gender?

___ Male
___ Female
___ Trans-male
___ Trans-female
___ Non-binary
___ Other/Prefer not to answer

2) What is your age? ______ 

3) What is your ethnicity?

___ Hispanic or Latino
___ Not Hispanic or Latino
___ Unknown

4) What is your race?

___ American Indian/Alaska Native
___ East Asian
___ South Asian
___ Native Hawaiian/Pacific Islander
___ Black or African American
___ White
___ More than one race – Black and White
___ More than one race – Other
___ Other or Unknown

5) Is English your native language? ______ Yes ______ No

If no, please list your native language: _______________________________________

6) What country were you born in? ___________________________________________
7) What school clubs, organizations, or teams have you been a part of during your time at VCU? (Please list all that apply, past and present. If none, write none.)

______________________________________________________________________________

8a) You indicated that you MOST identify as _____. How important is this identity to you?*

Not at all important

1 2 3 4 5 6 7

Extremely important

8b) How would you rate your political affiliation?**

Very Liberal

1 2 3 4 5 6 7

Very Conservative

9) How long have you lived in the United States?

____ My whole life

____ Less than my whole life (specify number of years: ___________)

10a) In your own words, what do you think the current study was about?

______________________________________________________________________________

______________________________________________________________________________

______________________________________________________________________________

10b) Did anything seem strange or out of the ordinary when you were participating?****

______________________________________________________________________________

______________________________________________________________________________

______________________________________________________________________________

* This item will be used in all studies. The blank will be auto-populated with the ingroup the participant selected (i.e., the group MOST identified with).

** This item will be used in Studies 1, 2, and 3 only.

*** This item will be used in Study 4 only.

**** This item will be used in Study 5 only.