Spring 5-15-2020

The Effects of Monitoring Expression and Outgroup Familiarity on Judgments of Other-Race Interaction Partners

Katlin Bentley
Washington University in St. Louis

WASHINGTON UNIVERSITY IN ST. LOUIS

Department of Psychological & Brain Sciences

Dissertation Examination Committee:
Tammy English, Chair
Hillary Anger Elfenbein
Calvin Lai
Randy Larsen
Clara Wilkins

The Effects of Monitoring Expression and Outgroup Familiarity
on Judgments of Other-Race Interaction Partners
by
Katlin Bentley

A dissertation presented to
The Graduate School
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

May 2020
St. Louis, Missouri

Table of Contents

List of Tables

List of Figures

List of Appendices

Acknowledgements

I would like to express my sincere gratitude to my advisor and committee chair, Dr. Tammy English, for her dedicated support, enthusiasm, patience, and guidance throughout this research project and my graduate studies. I would also like to thank the other members of my core dissertation committee, Drs. Clara Wilkins and Randy Larsen, for providing valuable feedback and insight for this project. Thanks also go out to my other committee members, Drs. Hillary Anger Elfenbein and Calvin Lai, for sharing their expertise and time while serving on my committee. I would also like to acknowledge the Department of Psychological and Brain Sciences for providing me with academic and financial support to complete this project and m.

Isidro, Tabea, Claire, Tess, Kathleen, and Molly – Thank you for putting up with all of my research- and non-research-related questions and rants. Please know that you all have played a significant role in helping me stay sane through this process, and I greatly appreciate your academic and emotional support. To the other members of the Emotion and Relationships Lab – thank you for helping with data collection and analyses and for providing feedback on my ideas.

Thank you to my family and friends for their tireless support and encouragement throughout my time at WashU. Finally I'd like to thank my bestie and husband Everett. Over the last five years, you've perfected the art of predicting when I'll come home hangry, when I ought to take a break from staring intensely at my computer screen, or when I just need a hug. Thank you for always being there for me and for helping me celebrate my victories and conquer the hurdles of graduate school.

<div align="right">Katie Bentley</div>

Washington University in St. Louis

May 2020

ABSTRACT OF THE DISSERTATION

The Effects of Monitoring Expression and Outgroup Familiarity

on Judgments of Other-Race Interaction Partners

by

Katlin Bentley

Doctor of Philosophy in Psychological and Brain Sciences

Washington University in St. Louis, 2020

Professor Tammy English, Chair

Research on emotion communication demonstrates that people are more accurate at recognizing

emotion when evaluating members of one's racial ingroup compared to assessments made for

outgroups. It is unclear what leads us to make erroneous outgroup judgments. Two factors may

play a central role in this process: judges' attentiveness to and knowledge about partners' group-

specific expressive behaviors. In this project, I tested moderators of people's ability to accurately

detect emotions during an in-person interaction when paired with a same- or other-race partner.

Findings indicate that when playing a cooperative game, people are surprisingly adept at

accurately judging outgroup emotions, but these assessments were susceptible to bias. Judges

were generally highly attentive to outgroup partners' expressivity, and this monitoring did not

benefit judgment accuracy. Outgroup exposure predicted somewhat improved assessments, but

the effects were inconsistent. Finally, people liked partners who they were able to judge more

accurately, regardless of the racial makeup of the dyad. This research builds on prior work

documenting miscommunication between racial groups by demonstrating that accurately judging

others' emotional experiences can be a complex process and being less attentive to or familiar

with outgroup expressivity cues does not necessarily interfere with judges' ability to form

accurate assessments.

Chapter 1: Introduction

We rapidly form impressions about others. From pinpointing someone's immediate emotional state to determining more enduring information like personality traits, being able to accurately judge the characteristics of our interaction partners is a crucial component of interpersonal communication. Although there is a wealth of literature on interpersonal judgments, there is limited application of these findings to the field of intergroup relations. Research demonstrates that achieving harmonious cross-group communication can be difficult, in part due to our tendency to misinterpret or misattribute outgroup characteristics (e.g., Cortes, Demoulin, Rodriguez, Rodriguez, & Leyens, 2005; Elfenbein & Ambady, 2002; Freeman, Schiller, Rule, & Ambady, 2010; Judd, Ryan & Park, 1991). It remains unclear which step in the judgment process most often results in failure. This issue might stem from two sources of error. People may fail to notice relevant behavioral cues when judging characteristics about outgroup interaction partners, particularly if being in an intergroup interaction leads people to behave differently than they would if they were with an ingroup partner. Alternatively, they may lack knowledge about these cues, making it hard to know which behaviors judges should attend to when interacting with an outgroup partner and, if the right ones are noticed, what these cues represent. Research has not yet conclusively determined the extent to which either factor may contribute to low judgment accuracy of outgroup members. This project aims to advance our understanding of how people make judgments about racial outgroup members and uncover where misperceptions can arise by examining how readily people can judge the characteristics of an other-race interaction partner and testing whether attending to their partner's expressions or being familiar with their partner's racial group (operationalized as outgroup exposure) moderates these outcomes.

1.1 Communicating Our Emotions

Nonverbal behavior allows us to convey our thoughts and feelings to those around us and to ascertain the internal states of others. We use a variety of *behavioral cues* from visual (e.g., face and body) and vocal (e.g., tone of voice) channels to share social signals (Scherer, Clark-Polner, & Mortillaro, 2011; Vinciarelli, Pantic, & Bourlard, 2008). Displaying and perceiving nonverbal behavior frequently occur at a subconscious level that does not require full attention or reasoning capabilities (de Gelder, 2006). Rapid processing of behavioral cues is particularly useful in situations where we need to distinguish whether someone has harmful or helpful behavioral intentions (Adams, Ambady, Macrae, & Kleck, 2006).

Research on emotion communication primarily focuses on facial expressions. This trend is somewhat unsurprising, as the face is one of the most readily accessible and specialized tools for conveying and perceiving emotion (Jack, 2013; Nusseck, Cunningham, Wallraven, & Bülthoff, 2008). Accordingly, there is less work examining the role of emotional body language and vocal prosody. Some researchers suggest that this dearth originates from the long-held assumption that non-facial information only serves to amplify facial expressions (Harrigan, 2005). This theory is no longer *in vogue* (Dael, Mortillaro, & Scherer, 2011), and there have been several recent studies demonstrating that body movement (e.g., Atkinson, Tunstall, & Dittrich, 2007; de Gelder & Huis in 't Veld, 2016; Watson & de Gelder, 2017) and vocal cues (e.g., Jiang, Paulmann, Robin, & Pell, 2015; Johnstone & Scherer, 2000; Zhu, 2013) alone can convey emotional information. Despite the literature asymmetry, no one cue appears to be the best measure of affective state. There are a few possible reasons why: emotional processes are not always accompanied by perceptible behavioral cues, there are individual differences in cue use in response to emotion elicitors, and there may be external factors such as social or cultural

norms that adjust expression (Barrett, 2006; Cacioppo, Berntson, Larsen, Poehlmann, & Ito, 2000).

While it is possible to transmit emotional content using one type of behavioral cue, using multiple nonverbal channels increases our ability to accurately send and recognize social signals (Jack, 2013; Nusseck et al., 2008). Bodily and vocal cues do not map easily onto discrete emotion categories; however, they can be indicative of the arousal level (i.e., low or high emotion intensity) and valence (i.e., positive or negative emotion) of one's affective state (Barrett, 2006). Displaying a combination of nonverbal behaviors simultaneously or within quick succession of one another can help resolve cues that are more ambiguous. For example, leaning back in a chair during a social interaction could have a variety of meanings depending on the context. One might be performing this movement as an expression of relaxation, dominance, or a desire to avoid the situation. Alternatively, there may be no intended emotional content (e.g., stretching). When presented in conjunction with other nonverbal behaviors such as gaze, facial expression, head movement, and arm position, we can determine how something as minimal as torso movement links to the person's internal state. Our perceptions of emotional facial expressions are strongly influenced by the presence of other behavioral cues (Wieser & Brosch, 2012). Congruency between multiple cues is the most effective way to communicate emotion (de Gelder, 2006; Meeren, van Heijnsbergen, & de Gelder, 2005); however, when there is apparent discordance between facial, bodily, and vocal expressions, certain cues are processed preferentially over others. Bodily and vocal cues presented in conjunction with facial expressions are often weighed more heavily over the facial information; they can also lead to faster and more accurate emotion recognition than viewing facial expressions alone (Meeren et al., 2005; Van den Stock, Righart, & de Gelder, 2007).

In addition to processing multiple behavioral cues, we also glean contextual information from the demographics of the sender that shape how we identify and interact with others (Jack & Schyns, 2015; Wieser & Brosch, 2012). This supplementary information can help us determine whether the person we are interacting with is a member of our ingroup or not. Although there seems to be some universality in the encoding and decoding of certain behavioral cues, there is also evidence that nonverbal behavior differs notably across groups (e.g., Elfenbein & Ambady, 2002).

1.2 Group Differences in Emotion Communication

While there is still disagreement on the exact level to which emotional nonverbal behavior is universal or culturally specific, there is growing consensus that innate factors interact with culture-specific learning to produce variations in expression. Certain cues seem to be more ubiquitous than others, particularly in the case of facial expressions, although there is still contention about the extent of universality in this domain (e.g., Gendron, Roberson, van der Vyver, & Barrett, 2014; Jack, Sun, Delis, Garrod, & Schyns, 2016). Bodily and vocal expressions of emotion appear to be more susceptible to cultural variations; very few of these behavioral cues have consistent meaning across different groups (Harrigan, 2005). For example, mutual gaze and physical touch during interpersonal interactions (Schofield, Parke, Castañeda, & Coltrane, 2008; Sorokowska et al., 2017), symbolic gestures (Matsumoto & Hwang, 2013), and smiling and brow furrowing during webcam use (McDuff, Girard, & el Kaliouby, 2017) have markedly different meanings between cultures. It is important to note that groups can and do display most, if not all, of these behaviors, but the contexts in which they occur, associated meaning, and frequency of use are often dependent on cultural background. Additionally, cultures differ in the utilization of situational factors during emotion perception, sometimes

relying on context over behavioral cues displays when making emotion judgments (Stanley, Zhang, Fung, & Isaacowitz, 2012).

Recent work on these differences suggest the existence of nonverbal expressive *dialects* originating from cultural variations on innate encoding and decoding processes (Elfenbein & Ambady, 2002, 2003; Elfenbein, Beaupré, Lévesque, & Hess, 2007). Nonverbal divergence between groups may occur randomly over time as distance, both social and geographical, increase or might result from purposeful creation of a distinct social identity (Elfenbein, 2013). As nonverbal behavior becomes more specialized and reinforced across groups, certain cues may end up becoming more diagnostic for group-specific emotion decoding than others. There is some evidence that individuals from different cultures focus on different portions of the face when judging emotion expressions. For example, East Asians tend to focus on the eyes when perceiving emotion while Western Caucasians focus on the mouth (Kelly et al., 2011; Yuki, Maddux, & Masuda, 2007), and Eastern perceivers' internal representations of facial features converge highly when presented with eye-related cues while Westerners respond more uniformly to mouth and eyebrow cues (Jack, 2013). These differences likely arise from cultural preferences to emphasize certain sections of the face when expressing emotion. While dialects in body language and vocal prosody are relatively understudied, one would assume that the theory still applies to these cues. Cultural groups should vary in their use of bodily and vocal behaviors when expressing emotion or making emotion judgments.

People can identify someone's cultural background by observing expressions containing behavioral dialects (Marsh, Elfenbein, & Ambady, 2003). These nuances in nonverbal behavior may reinforce perceived familiarity and similarity to self during social interactions, resulting in more favorable attitudes about individuals with similar expressive dialects. Having the ability to

detect small aberrations from behavioral cues shared by one's cultural group could be highly

adaptive, as it would allow for rapid distinguishing between possible enemies and allies.

1.3 Judging the Characteristics of Others

Our ability to judge emotional and social characteristics about others influences the way

we navigate interpersonal situations (Ambady, Bernieri, & Richeson, 2000; Letzring, Wells, &

Funder, 2006; Todorov, Olivola, Dotsch, & Mende-Siedlecki, 2015). Judgements typically

involve at least two people: someone making a judgment, or a *judge*, and someone being judged,

or a *target* (Funder, 1995; Ickes & Hodges, 2013). Judgments can vary in their *accuracy*, or

correctness, and *bias*, or direction, such as overestimating or underestimating levels of a given

trait (Gagné, & Lydon, 2004; West & Kenny, 2011). Research on empathic accuracy and

personality judgment has shown that judging others accurately depends on numerous factors

(e.g., Connelly & Ones, 2010; Human & Biesanz, 2013). One framework that concisely

describes the process of making a judgment and the factors that influence these assessments is

the Realistic Accuracy Model (RAM; Funder, 1995). This model defines a four-step process to

making an accurate judgment: relevance (i.e., the target must behave in a way indicative of the

trait of interest), availability (i.e., the judge must have access to this information), detection (i.e.,

the judge must notice this information), and utilization (i.e., the judge must interpret this

information correctly). In the context of the RAM, Funder proposed four moderators of

accuracy: a good trait (i.e., accuracy increases for characteristics or states that produce easily

observed behaviors), a good target (i.e., targets whose behaviors are consistently in line with

expectations about the characteristic or state are more accurately judged), a good judge (i.e.,

judges who are more accurate are better at cue detection and utilization), and good information

(i.e., judges have access to information relevant to the characteristic or state being judged, and

having more high quality information tends to lead to better judgments). This project primarily focuses on the second half of the RAM, specifically the extent to which judges attend to relevant behavioral cues, whether they use these cues to appropriately shape their judgments, and if this process of cue detection and utilization changes depending on the racial makeup of the dyad. Judges would need to know which behavioral cues differ based on the group membership of their interaction partner, pay attention to these group-specific cues, and apply the information gleaned from their observations in a valid way to be able to accurately form assessments about outgroup members. However, attentiveness alone will not ensure accurate judgments. Judges' ability to make effective assessments is contingent on targets' expression of observable behaviors that are relevant to the characteristic being judged.

In the context of intergroup interactions, there is a substantial amount of literature demonstrating group differences in emotion communication and recognition accuracy. One notable finding amongst this data is the trend that people tend to be better at assessing emotion expressions displayed by members of their own racial or ethnic background (i.e., ingroup) in comparison to those produced by people of dissimilar ancestry (i.e., outgroup). This phenomenon is referred to as the *ingroup advantage* of emotion recognition and is thought to stem from the dialect theory of cultural variations on innate encoding and decoding processes (Elfenbein & Ambady, 2002, 2003; Elfenbein et al., 2007). Members of a specific group are expected to have increased knowledge about the group's nonverbal accents and decoding and display rules, and thus be more effective at rapidly and accurately decoding the emotions of fellow ingroup members compared to assessments for outgroups. This tendency is thought to pervasively shape the way we recognize and process emotional faces. Judges' use of ingroup-favoring perceptual strategies for emotion decoding has been shown to inconsistently vary as a function of the

target's race, indicating either unconscious persistence or haphazard application of these biased mechanisms during face processing (Blais, Jack, Scheepers, Fiset, & Caldara, 2008; Brielmann, Bülthoff, & Armann, 2014; Fu, Hu, Wang, Quinn, & Lee, 2012). While research on this topic primarily focuses on judging emotional facial expressions, the effect also manifests beyond facial movements (e.g., body language: Kleinsmith, De Silva, & Bianchi-Berthouze, 2006; vocal prosody: Thompson & Balkwill, 2006), although there is limited research examining other channels of nonverbal expression. According to *social identity theory* (Tajfel & Turner, 2004), we tend to use perceptual shortcuts during social categorization that lead us to exaggerate the distinctness of the ingroup from other groups and believe that members of an outgroup are very similar to one another. These assumptions may be working in tandem with ingroup-favoring face perception strategies, such that people tend to be more motivated to fully attend to those they classify as ingroup members and more successful at decoding their expressions compared to situations featuring outgroup members. Alongside our facial processing and social categorization biases, our interactions with and perceptions of outgroup members are often shaped by other factors, namely our knowledge (or lack thereof) of outgroup norms and the stereotypical beliefs we ascribe to outgroup members.

1.4 Interacting with an Other-Race Partner

Intergroup conflict and misunderstandings often stem from decreased exposure to people of other races (Allport, 1954). This idea makes sense; it is easier to maintain prejudiced or ignorant beliefs about racial outgroups if you rarely encounter them during your daily routine or if members of one's ingroup advocate similar stereotypes. Exposure to and interactions with racial outgroups (i.e., intergroup contact) can be highly beneficial in dismantling misguided beliefs about outgroups. Intergroup contact increases familiarity with outgroup norms, decreases

anxiety related to intergroup interactions, and increases people's ability to empathize with outgroups (Al Ramiah & Hewstone, 2013; Pettigrew & Tropp, 2008). Why does this effect occur? Encounters with other races allow people to reshape the way they perceive racial outgroups, drawing from specific knowledge gained from these interactions rather than solely relying on stereotypes (Pettigrew, 1998). Over time, people may feel less threat and anxiety in response to intergroup interactions while becoming more knowledgeable about outgroup characteristics and behavioral nuances.

While intergroup contact generally tends to be beneficial for intergroup relations (Pettigrew & Tropp, 2006), it is important to note that direct (i.e., in-person) interactions with outgroups are not always as successful. Interracial interactions can be cognitively and emotionally taxing due to the balance between juggling monitoring self-expression and regulation alongside concerns about prejudice (Richeson & Shelton, 2007). People may feel stressed or anxious in anticipation of or during interactions with outgroup members (Hyers & Swim, 1998; Littleford, Wright, & Sayoc-Parial, 2005; Shelton, 2003; Stephan & Stephan, 1989; Vorauer, Hunter, Main, & Roy, 2000). Engaging in these interactions can lead to increased biases toward and avoidance of outgroups (Paolini, Hewstone, Voci, Harwood, & Cairns, 2006; Shelton, Dovidio, Hebl, & Richeson, 2009; Trawalter, Richeson, & Shelton, 2009), concerns about being perceived as a confirmation of group stereotypes (Shelton, Richeson, & Vorauer, 2006), and activation of stress-related physiological responses (Blascovich, Mendes, Hunter, Lickel, & Kowai-Bell, 2001; Mendes, Blascovich, Lickel, & Hunter, 2002; Page-Gould, Mendoza-Denton, & Tropp, 2008). While these outcomes may be amplified due to the context in which the interactions occur, there is also a concern that interacting and improving relations with

certain outgroups will not lead to similar attitude changes for other outgroups, even if interactions occur under optimal conditions (Amir, 1976; Forbes, 1997).

1.4.1 Concerns about group-specific stereotypes

Research on intergroup interactions demonstrates that our motivations and goals for accurate self- and partner perception shift depending on the racial group membership of the judge and the target. Frameworks such as the Stereotype Content Model (Fiske, Cuddy, Glick, & Xu, 2002; Fiske, Xu, Cuddy, & Glick, 1999) and the Behaviors from Intergroup Affect and Stereotypes Map (Cuddy, Fiske, & Glick, 2007) align group stereotypes along two dimensions, competence and warmth, depending on a group's perceived majority or minority status in society. Although these constructs vary independently (Wojciszke, 2005), perceptions of groups' warmth and competence tend to be negatively correlated (Yzerbyt, Kervyn, & Judd, 2008), such that stereotyped groups are often perceived as either warm but incompetent (i.e., liked but disrespected) or competent but cold (i.e., respected but disliked). People generally try to avoid confirming group-specific stereotypes during interracial interactions as these stereotypes determine which emotions others broadly feel towards that group (e.g., admiration, pity) and how they act around members of the group (Bergsieker, Shelton, & Richeson, 2010; Cuddy et al., 2007; Fiske et al., 1999, 2002). This aversion to confirming group stereotypes, also known as *stereotype threat*, can influence stress levels, attentiveness, working memory, and self-regulation and impair performance on social and cognitive tasks depending on the saliency of one's group membership (Johns, Inzlicht, & Schmader, 2008; Steele & Aronson, 1995; Schmader, Johns, & Forbes, 2008). Simply being in the same room as an other-race person is sufficient to induce stereotype threat (Inzlicht & Ben-Zeev, 2000; Shih, Ambady, Richeson, Fujita, & Gray, 2002; Steele & Aronson, 1995).

10

In the current project, I focus on interactions involving East Asians and Whites. Based on the high-status ascribed to Whites in the U.S., they are often stereotyped as competent but also more likely to act in a prejudiced way during interracial interactions (Cuddy et al., 2007; Vorauer, Main, & O'Connell, 1998). Whites are aware of these stereotypes (Frantz, Cuddy, Burnett, Ray, & Hart, 2004) and typically desire to be seen as non-prejudiced and likeable by other-race interaction partners (Bergsieker et al., 2010). Most studies examining interracial interactions focus on interactions between Black and White populations. Research on stereotype threat and interpersonal judgment accuracy among East Asians, particularly when interacting with Whites, is limited in comparison. Although Asians' status relative to other underrepresented minority groups in the U.S. varies (e.g., Bergsieker et al., 2010; Craig & Richeson, 2012; O'Brien & Major, 2005), Asians are still subject to group specific biases that differ from those experienced by other races (Landrine, Klonoff, Corral, Fernandez, & Roesch, 2006; Sue, Bucceri, Lin, Nadal, & Torino, 2007). Recent work examining hierarchies among racial groups in the U.S. suggests that the status ascribed to non-White minorities depends on two distinct dimensions: inferiority (i.e., perceptions of the group's abilities and achievement) and cultural foreignness (i.e., the extent to which the group is considered prototypically American); according to this Racial Position Model (Zou & Cheryan, 2017), Asian are deemed similar to Whites in terms of superiority but perceived as more culturally foreign (i.e., less American) than Whites. Due to the relatively high status associated with Asians, they are likely to be stereotyped as intellectually and economically competent, diligent, and high-achieving (Lin, Kwan, Cheung, & Fiske, 2005; Siy & Cheryan, 2013); however, Asians are also perceived as less American, regardless of their orientation towards American culture (Cheryan & Monin, 2005; Shih et al., 2002; Sue et al., 2007). Based on the stereotypes associated with this racial group, one would

11

expect that Asians interacting with an other-race partner may prioritize being seen as likeable and more American regardless of how competent they are assumed to be.

1.4.2 Getting to know outgroup members

Beyond the scope of live interactions, people tend to have improved recognition accuracy for images of same-race emotional faces rather than other-race ones (Anzures, Quinn, Slater, Tanaka, & Lee, 2013; Meissner & Brigham, 2001). The *perceptual expertise hypothesis* (Meissner & Brigham, 2001) proposes that people typically have more contact with members of their own racial group, leading them to have extensive experience with variation within the group and thus be faster and more accurate at recognizing same-race faces over other-race ones. Subsequently, low contact with other-race faces is expected to result in low expertise at decoding outgroup emotion on dimensions that actually do vary across groups (McLin & Malpass, 2001). Repetitive exposure to other-race faces tends to diminish this cross-race effect in short- and long-term settings (Elliott, Wills, & Goldstein, 1973; Goldstein & Chance, 1985; Malpass, Lavigueur, & Weldon, 1973; Sangrigoli, Pallier, Argenti, Ventureyra, & de Schonen, 2005); however, the effects are not always consistent (Malpass & Kravitz, 1969; Ng & Lindsay, 1994). One explanation may be that exposure alone is not necessarily enough to increase perceptual expertise. Instead, motivated individuation when processing other-race stimuli seems to be a more effective way to improve perception of outgroup faces (Hugenberg, Miller, & Claypool, 2007). Doing so requires somewhat attentive and effortful processing of outgroup members in a more nuanced, personalized way to lessen reliance on features of the overall outgroup when assessing outgroup targets (Walker & Hewstone, 2006).

One method of frequent exposure that is expected to induce outgroup individuation is intergroup contact. Increased contact with outgroups should provide opportunities to witness

group-specific variations in nonverbal behavior and increase familiarity with the expression and meaning of cues across different situations. Elfenbein and Ambady's (2002) meta-analysis on the ingroup advantage of emotion recognition found support for this idea: this advantage tends to be smaller for groups that are physically close to one another (i.e., share a regional border or reside within the same country). Developing and maintaining relationships with outgroup members is also associated with improved familiarity with outgroup expressive behavior. Individualized contact with members of different racial groups leads to improved performance in outgroup emotion recognition (Bukach, Cottle, Ubiwa, & Miller, 2012; Hancock & Rhodes, 2008; Walker, Silvert, Hewstone, & Nobre, 2008; Yankouskaya, Humphreys, & Rotshtein, 2014). Living in a region with a diverse population also leads to decreases in the ingroup advantage (Soto & Levenson, 2009), even if the person has only recently moved to the area (e.g., international college students; Elfenbein & Ambady, 2003; Yan, Andrews, Jenkins, & Young, 2016).

There are a few caveats limiting the potential benefits of developing outgroup familiarity through repeated exposure. First, observing outgroup members' emotional behavior is likely contingent on the type and quality of relationship one has with an outgroup member and whether people feel comfortable correcting misjudgments as they occur. One would expect that being in a close relationship with an outgroup member would lead to more experiences with outgroup behavior across many contexts and greater exchange of feedback compared to less close relationships, such as those with an acquaintance. An additional concern is that recurring contact with an outgroup should only boost judgment accuracy for targets from that particular group. So having more social connections with people from a specific racial or ethnic category is expected to associate with improved decoding of behavioral cues for members of that particular group but not necessarily for other outgroups. Finally, the dispersion of different racial groups within a

population tends to be uneven, and unsurprisingly minorities are much more likely to regularly

encounter non-minorities compared to the rate that non-minorities cross paths with minorities

(Omi & Winant, 2014). As a result, there may be limited opportunities for certain groups to

develop expertise through repeated interactions with people of other races or ethnicities.

1.5 The Present Research

Successfully navigating a social situation can be a complex process, particularly if we are

worried about being perceived accurately while simultaneously trying to accurately perceive the

characteristics and experiences of others. Past research on intergroup interactions has

demonstrated that we tend to be less accurate when judging outgroup social partners; however,

little is known about how these erroneous judgments occur. This project addresses gaps in the

literature by examining emotion-related judgments in the context of an in-person intergroup

interaction and testing the factors that may influence judgment accuracy. Study 1 involved

randomly assigning East Asian and White participants to interact with a same- or other-race

stranger for ten minutes, then provide self-other ratings of emotional experience, attention to

partner expressivity, partner likeability, and outgroup exposure. Study 2 expanded on this design

by incorporating an experimental component where expressivity attention was manipulated to

test causal links between expression attention and judgment accuracy alongside testing for

effects of outgroup exposure.

Recruitment was limited to these specific racial groups for a few reasons. Dyads featuring

Asian and White participants are not commonly studied in intergroup research, let alone in the

context of an in-person interaction, so focusing on these groups allowed me to examine a novel

type of cross-race interaction. These groups also differ notably in the way they experience and

express emotion (Tsai, 2007; Gross, Richards, & John, 2006; Matsumoto, 1990) and the

stereotypes they anticipate encountering during interracial interactions (Bergsieker et al., 2010; Cuddy et al., 2007; Sue et al., 2007). The paradigm used in both studies also addresses several limitations from prior research on interpersonal judgments and the ingroup advantage of emotion recognition. Participants in these studies are often asked to assess pictures, written paragraphs, or films featuring outgroup members (e.g., Beaupré & Hess, 2006; Kang & Lau, 2013; Kunstman, Tuscherer, Trawalter, & Lloyd, 2016; Ma-Kellams & Blascovich, 2012; Wickline, Bailey, & Nowicki, 2009). Judgments do not necessarily follow a live interaction. These designs may not have given people a fair test of judgment accuracy, as directly interacting with someone can provide judges with more useful contextual information about targets than just remotely observing them (Letzring et al., 2006).

This research aims to inform our understanding of how judgments about outgroup targets go wrong. While there has been substantial research on emotion and personality judgments and on intergroup relations, there has been little overlap between these areas of work. This project builds on studies documenting miscommunication between racial groups by considering the mechanisms that interfere with accurate judgment of outgroup interaction partners and the factors that may boost judgment accuracy.

The following research questions are addressed in this project:

1. Are people's judgments of outgroup members' emotions less accurate compared to those targeting ingroup members in the context of an in-person interaction?

2. Do people monitor their partner's expressivity more when interacting with an outgroup member compared to those interacting with an ingroup member, and does this monitoring interfere with judgment accuracy?

3. Is greater representation of an outgroup within one's social network associated with improved judgment accuracy when interacting with an unknown member from this group?

4. Do judges who are more effective at assessing their partner's emotions like them more than inaccurate judges? Do partners like accurate judges more than inaccurate ones?

1.5.1 Hypothesis 1

I expected that people would be less accurate judging the emotions of other-race interaction partners compared to judgments made for same-race ones. Across Studies 1 and 2, I focused on judgments about emotional experience to compare how readily people could assess target characteristics that differ in their expressivity. Given what we know about group-specific dialects in emotion communication and the factors that influence interpersonal judgments, accuracy was expected to vary as a function of the racial makeup of the dyad.

1.5.2 Hypothesis 2

I predicted that expressivity attention during the interaction would moderate judgment accuracy. It is unclear how much people normally pay attention to their partner's expressions during social interactions. Failure to notice outgroup behavioral cues due to lower attentiveness to partner's expressivity may explain why people tend to be less accurate when judging outgroup targets. I predicted that judges who were more focused on their partner's expressions during the interaction would be better able to judge their partners accurately compared to judges who were less attentive to their partner's expressivity.

1.5.3 Hypothesis 3

An alternative explanation for low accuracy of outgroup judgments might be a lack of familiarity with outgroup cues in social settings. So even if judges are attending to their partner's expressions, they may not know which cues to look for, identify invalid cues, or apply valid cues incorrectly. I predicted that judges who were more familiar with outgroups would assess members of these specific groups more accurately than judges with low familiarity. In this project, social network diversity was used as a proxy for outgroup exposure, as having more other-race people within one's social network was expected to indicate greater levels of meaningful contact with members of specific racial groups and subsequently greater competency for judgments of other members of these groups compared to someone with low outgroup exposure (Elfenbein & Ambady, 2003).

For those in mixed-race dyads, the magnitude of the outgroup exposure effect was expected to differ by judge race. Whites are currently one of the largest racial groups in the U.S., particularly in comparison to recent population estimates for Asians (U.S. Census Bureau, 2019). Given these disproportionate demographics, one would expect that Whites may have limited opportunities to regularly encounter and experience meaningful contact with Asians. White participants with more social network diversity may have presumably been exposed to more exemplars of the characteristics and behaviors associated with people of other races. Subsequently, White judges with more Asians in their network may be more effective at assessing the emotions of Asian interaction partners compared to White judges with few Asians in their network. This effect may not occur for Asian participants. Given their status as a minority in the U.S., Asians are anticipated to have frequent contact with Whites in a variety of situations due to the asymmetric number of White people relative to Asians in a given area. So

17

the number of Whites in one's social network may not be as beneficial for Asian judges' given

their opportunity to inadvertently develop expertise through repeated exposure to Whites.

1.5.4 Hypothesis 4

I expected that judges who were more accurate would like their partners more than would

inaccurate judges and that partners who were more accurately judged would like these judges

more than less accurate ones. People tend to ascribe more favorable characteristics to people who

are easily judged over targets who are more difficult to characterize (Colvin, 1993). I anticipated

that dyads in which judges are able to accurately assess their partners will end up having more

cohesive interactions and better rapport than ones where judges are less accurate as accurate

judges may be more effective at responding to their partner's experience and their partners may

in turn feel more understood. I also conduct exploratory analyses examining partner liking as a

potential moderator of judgment accuracy, to test whether judges' liking of targets affected their

ability to accurately perceive targets' experiences.

Chapter 2: Study 1 Introduction

Study 1 examined the factors that may naturally shape judgments of outgroup members following an in-person interaction. East Asian and White participants were randomly paired with a same- or other-race partner and played a game together for ten minutes. After the interaction, they completed assessments about their own and their partner's emotional experience. Participants also reported how much they attended to their partner's expressions, how much they liked their partner, and described their social network. Targets' self-rated emotional experience was compared against judges' perceived ratings to determine judgment accuracy. Dyad type and judges' attention to targets' expressivity, outgroup exposure, and liking of targets were tested as moderators of accuracy. Accuracy within dyads was examined as a predictor of judges' ratings of target liking.

Chapter 3: Study 1 Methods

3.1 Sample

Participants ($N = 320$) were recruited via Washington University's Psychology Department Subject Pool and flyering within the St. Louis community.[1] The sample was 63.4% female, 35.9% male, and .6% nonbinary. Age ranged from 18 to 36 years old ($M = 19.90$, $SD = 1.67$). Participants were evenly split across two racial groups: 49.1% East Asian or East Asian American and 50.9% White, Caucasian, or European American. The majority of White participants (97.5%) identified as U.S. citizens, while 57.7% of Asian participants identified as U.S. citizens. In terms of generational status, 42.3% of Asians were first generation Americans (i.e., they and their parents are foreign born), 46% of Asians were second generation Americans (i.e., at least one parent was born outside of the U.S), and 11.7% of Asians were third-and-higher generation (i.e., those with two U.S. native parents).

To determine target sample size, I ran an *a priori* power analysis for linear regression in G*Power (Faul, Erdfelder, Lang, & Buchner, 2007) using the effect sizes for personality judgments and the number of parameters expected in my most complex models. The average effect size for the effect of behavioral expressiveness during a social interaction on personality trait judgments (as assessed by external judges) ranges from Gelman's $d = .10$ to .71 (Letzring & Human, 2014). Recruitment exceeded the minimum required number of dyads (n = 151) to reach a minimum effect size of .10 with power of .80.

---

[1] The majority of participants were undergraduate students from Washington University (97.2%), while remaining participants were university-affiliated graduate students or staff (2.8%).

To be considered eligible for the study, participants were required to be at least 18 years old and identify either as East Asian or East Asian American or White, Caucasian, or European American. Participants who identified as both Asian and White (i.e., multiracial) or with a different racial group were considered ineligible for the study. Those who passed the eligibility criteria were invited to the lab in pairs and randomly assigned into one of three types of unacquainted dyads based on their own and their partner's race: Asian-Asian (55 dyads, n = 110), White-White (52 dyads; n = 104), and mixed-race (53 dyads, n = 106). Prior to beginning study procedures, participants were asked to verbally confirm if they recognized their partner. Only those who did not know their assigned partner were allowed to participate in the study. In situations where only one person arrived for the session or participants recognized one another, they were provided an opportunity to reschedule or complete an alternate study based on whichever projects were ongoing in the lab at that time.

3.2 Procedure

Upon arriving to the lab, participants were led to separate testing rooms to review the study consent form and complete a questionnaire on their current emotional experience. Next, they were informed that they would have a ten-minute interaction with their partner in which they would introduce themselves for two minutes then complete a cooperative word-guessing activity based on the game Taboo for eight minutes. This activity was selected because it has been used in prior research to effectively elicit a combination of positive and negative emotions when interacting with an other-race partner (West, Koslov, Page-Gould, Major, & Mendes, 2017). The paradigm used in the present study was based off of the methods used by West et al. (2017). After the introductions, participants were advised that they would take two-minute turns trying to get their partner to guess words without being able to use any of the "taboo" words or

21

phrases listed on each prompt card. Participant would each play two rounds as a guesser and two as a prompter and would be provided with a timer to notify them when to switch places. The first guesser and prompter were selected via random assignment. To motivate participants to stay engaged with the task, the participants from the five dyads with the most prompts guessed correctly between the two players would each receive a $20 gift card once data collection was complete. See Appendix A for a copy of the game instructions.

After being provided with an initial description of how to play the game, participants were seated together and the experimenter asked them to introduce themselves while they "prepared the game materials" in a separate area. After two minutes elapsed, the experimenter returned to the room to provide the dyad with the game materials (i.e., prompt cards, paper and pencils for keeping score, and a timer), remind them of the game rules, and inform the first prompter and guesser of their role assignments. The experimenter then left the room to allow the participants to commence the game. The experimenter remotely observed the interaction from an adjacent room to verify that participants were following the game instructions, interrupting as needed to ensure that they stayed on task.

Once the final round of the game finished, participants were led back to their original testing rooms to complete a series of post-interaction measures. They rated their own and their partner's emotional experience, their expressivity attention during the interaction, and how likable they found their partner to be; described their social network; and provided their demographics. Participants completed several additional measures that were included in the study protocol but beyond the scope of this project prior to being debriefed on the study's purpose and compensated. The full protocol took approximately 1 hour, and participants received one course credit or $10.

3.3 Measures

All assessments were administered when the participants were in separate testing rooms.

3.3.1 Primary outcomes

3.3.1.1 Emotional experience. Participants reported the extent they were feeling ten possible emotions on a scale of 1 (*not* [feeling this emotion] *at all*) to 7 ([feeling this emotion] *a great deal*). This measure included a variety of positive and negative items from low to high arousal levels. I created two categories for analyses: positive and negative emotion. Positive emotion included the average of happy, excited, amused, relaxed, proud, and curious ($\alpha = .76$). Negative emotion included the average of disgusted, anxious, sad, frustrated, embarrassed, and bored ($\alpha = .65$). These emotions were selected from literature on basic emotions (e.g., Ekman, 1992a), the affective circumplex (Posner, Russell, & Peterson, 2005), and self-conscious emotions (e.g., Tracy & Robins, 2004) to provide a concise number of options that would capture feelings that could reasonably arise during the interaction.

3.3.1.2 Partner liking. The Reysen Likeability Scale (Reysen, 2005) measures participants' feelings toward their partner across 11 items (e.g., "This person seems warm") on a scale from 1 (*disagree strongly*) to 7 (*agree strongly*). The items were averaged to obtain a single score for each person ($\alpha = .88$). Higher scores indicate greater affiliation towards one's partner.

3.3.2 Primary predictors

3.3.2.1 Expressivity attention. Participants rated how aware they were of their partner's expressivity and their own across different expressive channels, including the face (e.g., How much did you attend to your partner's facial expressions?), body (e.g., How much did you attend to your partner's body language?), speech (e.g., How much did you attend to what your partner was saying?), and voice (e.g., How much did you attend to how your partner was speaking?).

The measure contained eight items, rated on a scale from 1 (*not at all*) to 7 (*a great deal*). Six items were averaged to create two constructs: focus on partner expressivity attention ($\alpha = .72$) and self-expressivity attention ($\alpha = .75$).[2] Higher scores indicate greater attention to one's partner's expressions or one's own, while lower scores mean decreased expressivity attention. As research on expressivity attention is limited, it is unclear if focusing on an interaction partner's expressions is orthogonal to focusing on one's own expressions. For the purposes of this project, they were treated as separate constructs in order to directly test whether partner expressivity attention moderated judgment accuracy. See Appendix B for a copy of this measure.

     3.3.2.2 Outgroup exposure. Participants described their social network using the Social Convoy Questionnaire (SCQ; Antonucci, 1986). This measure allows people to list social partners across different levels of their network, specifically their inner (i.e., "people you feel very close to"), middle (i.e., "people you feel close to"), and outer (i.e., "people whom you do not feel close to but who are still important to you") social circles. Participants could list up to ten people within each circle, for a total of 30 social partners across the full network. They also provided demographic information about each social partner, their relationship with the person (e.g., significant other, family member, coworker), and frequency of face-to-face contact with the person (i.e., in-person or video chatting). Outgroup exposure was operationalized as the ratio of Asian or White social partners (depending on the participant's own race) relative to the total size of one's social network. So for Asian participants, this ratio describes the number of White people relative to the size of the full network ($M = .29$, $SD = .25$), while the ratio focuses on the number of Asian social partners divided by the full network size for White participants ($M = .18$,

---

[2] The items examining attention to speech content were excluded given their impact on the constructs' reliability when included: partner expressivity ($\alpha = .70$) and self-expressivity ($\alpha = .72$).

SD = .23). Primary analyses examined outgroup exposure as a moderator of judgment accuracy. See Figure 1 for a graph depicting the number of in- and outgroup social partners in participants' networks.

I conducted an analysis of variance (ANOVA) to test for effects of judge and target race (coded as 0 = White, 1 = Asian) on outgroup exposure. There was a significant difference in the number of outgroup members judges reported in their network depending on their race, $F(1, 315) = 18.00$, $p < .001$, $\eta_p^2 = .05$. As expected, White judges ($M = .18$, $SD = .23$) listed fewer Asian social partners in their network compared to the number of White social partners listed by Asian judges ($M = .29$, $SD = .25$). Target race did not predict outgroup exposure, $F(1, 315) = 1.95$, $p = .164$, $\eta_p^2 = .01$, nor was the interaction between judge and target race significant, $F(1, 315) = .58$, $p = .448$, $\eta_p^2 = .00$.

3.3.3 Other measures

3.3.3.1 Demographics. Participants provided basic demographic information about themselves prior to debriefing. This measure asked about their gender, sexual orientation, age, race, citizenship, education level, relationship status, family socioeconomic status, and political orientation. Judge and target race were verified using participants' responses on the demographics form.

Chapter 4: Study 1 Results

4.1 Analysis Overview

To test my hypotheses, judges' and targets' emotional experience ratings were used to assess judgment accuracy. Dyad type (coded as 0 = same-race and 1 = mixed-race), as well as judges' attention to target expressivity during the interaction, outgroup exposure, and liking of targets were tested as moderators. Dyad gender (0 = same-gender, 1 = mixed-gender) was also tested as a potential moderator in exploratory analyses; however, there were minimal effects and no significant interactions for any of the dependent variables, so the results presented here are collapsed across gender.[3] Judgment accuracy was examined as a predictor of likeability ratings within dyads.

An inter-item correlation matrix and descriptives for key variables are provided in Table 1. Figure 2 provides a graphed summary of participants' self-reported experiences, split by dyad type and participant race.

4.2 Profile Correlations

Accuracy between judge- and target-rated emotional experience was first examined using self-other agreement assessed via within-dyad correlations computed by comparing targets' self-ratings for each item against judges' perceived ratings for the same items. Correlations provide a simple way to describe similarity between ratings, represent an index of effect size, and are a common metric of self-other agreement in psychological research (e.g., Vazire, 2010; Watson,

---

[3] Half of the of the dyads (55.6%) were same gender. Compared to those in mixed-gender dyads, participants in same-gender dyads were less effective at judging targets' negative experience, $B = -.22$, $SE = .08$, $p = .008$, $R_\beta^2 = .02$. Targets' emotional experience and judges' expressivity attention and ratings of target liking did not significantly differ between dyad types.

Hubbard, & Wiese, 2000a). Positive correlations indicate accuracy, while negative ones indicate inaccuracy. The magnitude of the correlation reflects the degree of accuracy or inaccuracy. Mean profile correlations broken down by judge and target race are graphed in Figure 3.

The Actor-Partner Interdependence Model (APIM; Cook & Kenny, 2005) was used for three sets of analyses: predicting the emotion judgment profile correlations from participants' expressivity attention and outgroup exposure and predicting partner liking from the profile correlations within dyads. This statistical method accounts for interdependence between partners while testing for actor effects (e.g., influence of the actor's condition on the actor's and partner's behavior) and partner effects (e.g., influence of the partner's condition on the actor's and partner's behavior). Here, APIM was used to simultaneously test both judges' and targets' expressivity attention and outgroup exposure as predictors of judgment accuracy (Models 1 through 4) and their emotion judgment accuracy as predictors of liking (Models 5 and 6) using indistinguishable dyads. I report semi-partial R^2 values (R_β^2) as indicators of APIM effect sizes, which reflects the proportion of variance explained (Edwards, Muller, Wolfinger, Qaqish, & Schabenberger, 2008). Estimates for theses analyses are provided in Tables 2, 3, and 4.

4.2.1 Judgment accuracy for same- versus other-race interaction partners

When comparing the mean-level profile correlations across dyad types using ANOVA, there was a marginal difference between groups, $F(1, 318) = 3.76$, $p = .053$, $\eta_p^2 = .01$. Judges paired with same-race targets ($M = .64$, $SD = .26$) were somewhat better at assessing their partners' emotional experience than were judges in mixed-race dyads ($M = .58$, $SD = .29$, $p = .053$). A two-way ANOVA was conducted to test if judge race (coded as $0 = $ White and $1 = $ Asian) moderated the effect of dyad type. While the aforementioned main effect of dyad type remained marginal, $F(1, 316) = 3.75$, $p = .054$, $\eta_p^2 = .01$, there was no effect of judge race on

27

emotion judgment accuracy, $F(1, 316) = .08$, $p = .782$, $\eta_p^2 = .00$, and the interaction between

dyad type and judge race was not significant, $F(1, 316) = .00$, $p = .956$, $\eta_p^2 = .00$.

4.2.2 Expressivity attention and judgment accuracy

Using APIM, I tested two models examining predictors of the emotion judgment profile

correlations: one testing expressivity attention between partners within each dyad (Model 1) and

a second testing if dyad type moderated the effects of the first model (Model 2). See Table 2 for

the estimates from these models.

In Model 1, monitoring the expressions of one's partner was not predictive of judgment

accuracy. Neither judges' attention to target expressivity, $p = .217$, nor targets' attention to judge

expressivity, $p = .797$, were associated with judges' ability to accurately perceive targets'

emotional experience. Dyad type did not moderate either of these effects, $ps \geq .358$.

4.2.3 Outgroup exposure and judgment accuracy

Two APIM models were ran to test for effects of outgroup exposure: Model 3 looked at

partners' outgroup exposure within each dyad and Model 4 tested whether dyad type moderated

these effects. See Table 3 for the output for these models.

In Model 3, there was a marginal effect of judges' outgroup exposure, $p = .050$; judges

with more outgroup members in their social network had slightly more difficulty accurately

perceiving the emotions of their assigned partner compared to low-exposure judges. Targets'

outgroup exposure was not predictive of judges' ability to accurately assess said targets'

emotional experience, $p = .626$. When examining dyad type as a potential moderator, there was a

marginal interaction between judges' outgroup exposure and dyad type, $p = .061$. Simple effects

for judges' outgroup exposure were tested separately for same- and mixed-race dyads. In same-

race dyads, judges with high outgroup exposure (i.e., those with a greater number of outgroup

members in their social network) had more difficulty assessing their partners' emotional experience compared to with low outgroup exposure (i.e., those with few outgroup members in their social network), $B = -.20$, $SE = .07$, $p = .004$, $R_\beta^2 = .04$. In mixed-race dyads, outgroup exposure for judges was not predictive of judges' ability to accurately assess targets' emotional experience, $B = .03$, $SE = .11$, $p = .802$, $R_\beta^2 = .00$.

An exploratory analyses was conducted to test the effect of judges' outgroup exposure separately for Asian-Asian and White-White dyads, the aforementioned outcomes were driven by participants in Asian-Asian dyads. Asian judges with greater outgroup exposure were worse at assessing Asian targets compared to Asian judges with low outgroup exposure in same-race dyads, $B = -.25$, $SE = .10$, $p = .010$, $R_\beta^2 = .07$. Judges' outgroup exposure had no effect on emotion judgment accuracy for those in White-White dyads, $B = -.14$, $SE = .11$, $p = .185$, $R_\beta^2 = .02$.

4.2.4 Partner liking and judgment accuracy

To examine the relationship between liking and emotion judgment accuracy, I tested two APIM models: one using the within-dyad profile correlations as predictors of partner liking (Model 5) and a second testing dyad type as a moderator of the Model 5 predictors (Model 6). See Table 4 for the output for these models.

The more accurate judges were at assessing targets' emotional experience, the more they liked their assigned partner, $p < .001$, but being accurately judged by their assigned partner did not predict judges' liking ratings, $p = .159$. Dyad type did not moderate these effects, $ps \geq .547$.

4.3 Truth and Bias Analyses

Next, I used the Truth and Bias Model (TBM; West & Kenny, 2011) to predict judges' ratings of target emotional experience from targets' own self-reports. This statistical technique

allows researchers to simultaneously test for accuracy and bias and has been used in many prior studies (e.g., Eldesouky, English, & Gross, 2017; Holoien, Bergsieker, Shelton, & Alegre, 2015; Overall, Fletcher, & Kenny, 2012; West, Magee, Gordon, & Gullett, 2014). While profile correlations are beneficial for obtaining information on the overall magnitude of emotion judgment accuracy, TBM yields more detailed information about specific patterns in these judgments, indicating where misperceptions arise. Following the guidelines of TBM, I grand-mean centered targets' own ratings and centered judges' assessment on this rating to examine three pieces of information: accuracy (i.e., the association between the target's own rating and the judge's assessment), directional bias (i.e., overestimation vs. underestimation of the assessment), and assumed similarity (i.e., the association between the judge's own rating and their assessment). This method was used to test two models predicting positive and negative emotional experience. TBM allows for more nuanced analyses compared to correlations, specifically the inclusion of moderators. Dyad type and judges' partner expressivity attention, outgroup exposure, and target liking were tested as potential moderators by including the variable of interest (which tests for moderation of directional bias) and its interactions with the parameters for accuracy and assumed similarity as predictors in the model. The models testing partner expressivity attention and outgroup exposure as moderators examined outcomes for mixed-race dyads only. Estimates and effect sizes (as semi-partial R^2 values) for the basic model (Model 1) and moderator models (Models 2 through 5) are provided in Tables 5 and 6. In this section, I report all statistics for Models 1 and 2, but for Models 3 through 5 I focus on effects testing moderation of judgment accuracy. Statistics for moderators of directional bias and assumed similarity are available in the aforementioned tables.

4.3.1 Judgment accuracy for other-race interaction partners

Model 1 examined accuracy, directional bias, and assumed similarity for positive and negative emotion judgments. Model 2 tested dyad type as a moderator of these three parameters. See Table 5 for the estimates for these models.

4.3.1.1 Judgment accuracy. Judges were able to accurately assess targets' positive emotion experience, $p < .001$, but they had difficulty judging targets' negative experience, $p = .116$. Dyad type had no moderating effect on accuracy for positive, $p = .497$, or negative judgments, $p = .634$.

4.3.1.2 Judgment biases. Judges tended to underestimate targets' positive experience, $p < .001$, but overestimated their negative emotions, $p < .001$. Regardless of the valence of the emotions being assessed, judges relied heavily on their own emotional experience as a reference when determining what their partners' felt, $ps < .001$. Dyad type did not moderate directional bias or assumed similarity for positive, $ps \geq .189$, or negative judgments, $ps \geq .367$.

4.3.2 Expressivity attention as a moderator of judgment accuracy

Model 3 tested partner expressivity attention as a possible moderator of judgment accuracy among mixed-race dyads. See Table 5 for the estimates from this model.

Judges' focus on partner expressivity had a marginal effect on their ability to assess negative experience, $p = .075$. Judges who paid more attention to targets' expressivity during the interaction were slightly better at deciphering their negative emotions compared to judges who were less focused on partner expressivity. Judges' focus on partner expressivity had no moderating effect on accuracy for positive emotion judgments, $p = .502$.

4.3.3 Outgroup exposure as a moderator of judgment accuracy

Model 4 tested judges' exposure to outgroups as a moderator of judgment accuracy among mixed-race dyads. See Table 6 for the estimates from this model.

Judges with more outgroup members in their social network were better able to assess other-race targets' negative experience compared to judges with low outgroup exposure, $p = .031$; however, judges' outgroup exposure had no moderating effect on their ability to accurately assess positive experience, $p = .212$.

4.3.4 Target liking as a moderator of judgment accuracy

Model 5 tested whether judges' ratings of target liking moderated judgment accuracy across all dyad types. See Table 6 for the estimates from this model.

Judges' liking ratings had no moderating effect on accuracy for positive, $p = .167$, or negative emotion judgments, $p = .492$. Regardless of how much judges liked their assigned partner, it did was not linked to their ability to accurately perceive their partners' emotional experience.

Chapter 5: Study 1 Discussion

Prior research on intergroup interactions has yet to conclusively determine the mechanisms behind our tendency to misinterpret and misattribute the characteristics of outgroup members. Two factors may play a central role in this process: judges' attentiveness to targets' expressive behaviors and judges' knowledge about targets' group-specific behavioral dialects. In the present study, we examined people's ability to detect their partner's emotional experience in the context of an in-person interaction to see if interacting with a same- or other-race partner impacted accuracy. Our findings indicate that people are able to accurately judge certain outgroup emotions, however these assessments were susceptible to bias.

Contrary to our predictions for Hypothesis 1, there was no clear evidence of a systematic ingroup advantage for emotion judgments. Judges paired with same-race targets were somewhat better at assessing their partners' emotional experience than those with other-race targets when looking at the overall emotion profiles, although this effect was marginal and dyad type had no moderating effect on judgment accuracy in the Truth and Bias analyses. Participants were able to successfully judge positive emotions but had difficulty assessing negative ones, regardless of the racial makeup of the dyad. While accuracy was similar for judges in same- and mixed-race dyads, these judgments were subject to bias. Judges tended to underestimate positive experience and overestimate negative experience, and they relied on their own emotional experience as a basis for their judgments. There are a few possible explanations for these tendencies. When there are few readily available cues, we tend to assume that others' experiences are similar to our own, ascribing the most relevant and conveniently available information (i.e., our own personal experience) to our judgments (Funder, Kolar, & Blackman, 1995; Ready, Clark, Watson, & Westerhouse, 2000; Watson, Hubbard, & Wiese, 2000b). If judges had trouble identifying or

33

decoding the verbal and nonverbal cues available to them, they may have either consciously or subconsciously use their own emotional experience as a reference point when determining how their partner felt (e.g., I enjoyed the game, so I think my partner must have too). If there was a detectable disconnect between targets' emotional experience and their expressions, judges' may have tried using their own experience as a comparison to resolve any contrasting information targets were communicating (e.g., telling your partner that you are enjoying the game but showing minimal smiling throughout the interaction).

Attention to partners' expressivity was not associated with more accurate emotion judgments. This unexpected outcome may have occurred in part because participants attended quite a bit to their partner across both types of dyads. There may be something about interacting with a stranger or possibly the design of the activity itself that encouraged people to make an effort to attend to their partner's expressions, no matter who their partner was. People are generally preoccupied with how they are perceived by others, often entering social situations with a goal to be accepted, liked, and respected (Baumeister, 1982; Baumeister & Leary, 1995). This desire has been shown to lead to increased monitoring for indicators of rejection during interpersonal interactions in order to attempt to adjust behavior accordingly and avoid exclusion (Leary, Tambor, Terdal, & Downs, 1995; Vorauer, 2006). Participants' tendency to pay close attention to their partner may simply reflect a strong need to belong. It should be noted that participants did not report ceiling-level attempts to attend to their partners' expressions, so lack of a moderating effect on judgment accuracy may also stem from not paying enough attention or misreporting the extent to which they believed they attended to their partner's expressions during the interaction.

Outgroup exposure was linked to accurate negative emotion judgments but was not associated with positive emotion judgment accuracy. Judges with higher outgroup exposure were more effective at determining what negative emotions their other-race partner felt during the interaction; however, this trend did not map onto positive experience judgments. There may be a few reasons for this outcome. Forming and maintaining relationships with people of other races imparts knowledge and understanding of outgroup expressive norms (Elfenbein & Ambady, 2003; Walker et al., 2008; Yankouskaya et al., 2014). As a result of these social connections, judges may eventually become adept at differentiating emotion expressions, particularly negative ones, displayed by other members of the same racial group compared to judges with low outgroup exposure. The lack of a finding for positive emotion judgments is not too surprising, given that positive emotion is associated with a more consistent set of easily identified behavioral expressions compared to negative emotion (Ekman, 1992b; Montillaro, Mehu, & Scherer, 2011). Additionally, people generally are willing to openly express positive emotion in interpersonal contexts (Gross et al., 2006; Matsumoto, Yoo, Hirayama, & Petrova, 2005), while there may be a myriad of personal, social, and cultural reasons for people to try to minimize or avoid displaying negative expressions in these settings.

This study also found evidence suggesting that outgroup exposure could be detrimental to certain types of interactions, namely ones between Asian participants. Why might Asian judges with more White people in their social network have difficulty assessing Asian targets' emotions? These judges likely have more frequent contact with White people compared to those with low exposure (i.e., fewer Whites in their social network) and may in fact have fewer opportunities to form relationships with other Asians beyond family members. These judges could be using facial processing heuristics that have essentially been calibrated with White faces

more so than Asian ones (McLin & Malpass, 2001; Meissner & Brigham, 2001), particularly if they grew up in areas that were predominantly White or currently live in an area with a higher ratio of Whites than Asians. As a result, high-exposure Asian judges may be inadvertently decoding Asian targets' expressions in an ineffective manner given the race of their interaction partner, while low-exposure Asian judges used more appropriate facial processing strategies for Asian targets and were subsequently more accurate in comparison.

As predicted, judges liked targets who they could readily judge; however, judges' liking of targets was not contingent on targets' ability to accurately perceive judges' emotional experience. In some ways, this outcome is unsurprising. If we find it easy to identify our partner's emotions, we may make more of an effort to share our feelings and respond accordingly to theirs throughout the interaction (Colvin, 1993; Laurenceau, Barrett, & Peitromonaco, 1998; Reis & Shaver, 1988). What is unusual is that this relationship was one-sided in this study. One would expect that developing rapport with your partner would require you to both feel as though you are being perceived accurately; however, our findings indicate that in this particular context, being able to understand your partner's emotions may be enough to promote affiliative feelings towards them. It may be the case that having a transitory interaction with a stranger and completing this particular activity encourages people to use their ability to accurately judge their partner as a metric of a "successful" interaction. In more traditional social settings, people often have the goal of getting to know the other person and in turn feel that their partner knows and understands them (Baumeister, 1982; Baumeister & Leary, 1995). In the present study, the goal was essentially to get as many correct answers as possible based on clues provided by one's partner. Accurately judging the other person's emotions may have been a byproduct of increased attention to what their partner was communicating in order to score more

points, while their partners' ability to accurately judge their own emotional experience was less pertinent to the task.

There were a few additional limitations to this study. Participants felt relatively high levels of positive emotion and minimal negative emotion. The almost floor-level self-reports for negative experience means that there was little to no variability among participants' experiences. Since targets were experiencing minimal negative emotion during the interaction, judges were provided with very few indicators of negative experience to base their judgments on. This issue likely stems from the task itself. Participants really liked the game. At the end of the session, one or both participants frequently commented about how unique the study design was and how much they enjoyed meeting their partner. In some ways this makes sense. They were given a novel task during a research study: playing a card game that is relatively popular with young adults. Additionally, the game rules were set up to encourage participants to get to know and communicate with their partner. In order to succeed, you had to provide your partner with useful clues and interpret the clues given to you. The gift cards also likely incentivized dyads as a reward for working together more effectively. Having the experimenter outside of the testing room meant that some dyads failed to strictly adhere to playing the game in the allotted time. While all dyads completed four rounds of the game, small aberrances in the study's protocol sometimes occurred, such as failing to switch roles between rounds, accidentally starting a fifth round, or trying to initiate conversation in between rounds. The experimenters were always able to get participants quickly back on track after these instances, but having participants manage the gameplay themselves may have encouraged them to feel less pressure to follow the game's rules and stay attentive to the game and their partner throughout the interaction. These issues were

addressed in Study 2 to increase consistency across study sessions and encourage participants to

have more varied emotional experiences during the interaction.

Chapter 6: Study 2 Introduction

Study 2 expanded on Study 1 by taking an experimental approach to manipulate expressivity attention among intergroup dyads to test a causal relationship between attention to expressivity and judgment accuracy while also examining the effects of outgroup exposure. Asian and White participants were paired with White and Asian partners respectively and played the same game from Study 1 together for ten minutes. Prior to meeting their partner, participants described their social network then were randomly assigned to one of two conditions: an experimental condition that asked them to attend to their partner's expressions during the interaction or a control condition that asked them to focus on the task itself. Afterwards, participants completed assessments about their own and their partner's emotional experience, how much they attended to their partner's expressions, and how much they liked their partner. Using similar analysis techniques from Study 1, targets' self-rated emotional experience was compared against judges' perceived ratings to assess accuracy. Condition assignment and judges' race, target expressivity attention, outgroup exposure, and liking of targets were tested as moderators of accuracy. Judgment accuracy was again examined as a predictor of target liking.

Chapter 7: Study 2 Methods

7.1 Sample

Participants ($N = 328$) were recruited from Washington University and the St. Louis community.[4] The sample was 69.7% female and 30.3% male. Age ranged from 18 to 30 years old ($M = 19.68$, $SD = 1.69$). Participants' race was evenly split across groups: 50.0% East Asian or East Asian American and 50.0% White, Caucasian, or European American. Almost all White participants (98.7%) identified as U.S. citizens, while 58.6% of Asian participants identified as U.S. citizens. Most Asians were first generation Americans (58.6%), followed by second generation (40.1%) and third-and-higher generation (1.3%). Target sample size was calculated using a procedure in G*Power similar to the one described for Study 1. Recruitment for this study was approximate to the minimum required number of dyads (n = 166) to reach a minimum effect size of .10 with power of .80 for my most complex models.

The eligibility requirements for Study 2 were the same as those in Study 1. Participants who met these criteria were invited to the lab in pairs to form unacquainted, mixed-race dyads. Sessions in which participants recognized one another or one person failed to arrive for the study were cancelled, and the participants were asked to reschedule.

7.2 Procedure

Study 2 had a similar design to the procedure used in Study 1 with a few changes added in order to address the limitations of the previous study. Participants were invited to the lab to play the card game with a stranger for ten minutes, then completed ratings about their own

---

[4] Similar to Study 1, most participants were Washington University undergraduate students (94.5%), while the rest were university-affiliated graduate students (5.5%).

experiences during the interaction and their perceptions about their partner's experiences. A condition assignment was added to manipulate expressivity attention during the game. After reviewing the game rules, Asian and White participants received additional instructions based on random assignment. Those assigned to the experimental condition were asked to attend to their partner's emotion expressions (including facial, bodily, and vocal cues) as much as possible during the interaction, while participants in the control condition were advised to attend to the game itself. The unassigned partners of these condition-assigned participants received the control instructions. See Appendix C for a copy of the condition-specific instructions.

The social network questionnaire was moved to the beginning of the study. Given the importance of using this measure to calculate one of the study's variables of interest, I wanted to administer the survey before participants received their condition-specific instructions or met their interaction partner. The goal of this change was to obtain clean network data that would be less prone to being unintentionally skewed due to other factors, particularly for those assigned to an experimental dyad. Participants completed the social network survey immediately after consenting to participate in the study, while they and their partner were seated in separate rooms and prior to meeting one another.

The remaining changes implemented in the Study 2 protocol were intended to make the game more difficult and the interactions between partners less affiliative. These updates were piloted to ensure that they could be feasibly implemented and that participants would consistently find the game more challenging while experiencing and expressing a greater level of negative emotion during the interaction. Given the somewhat one-sided emotional experiences participants reported in Study 1, I wanted to increase the difficulty of the game so targets would feel and express a wider array of emotions during the interaction. Doing so would also allow

41

judges to have more opportunities to assess negative emotion alongside positive emotion. The two-minute introduction from Study 1 was removed in Study 2 to decrease rapport between partners prior to starting the game. The experimenter, who had been monitoring the game from outside the testing room in Study 1, was instead seated in the room with the participants. In addition to ensuring that participants adhered to the game instructions and promoting uniformity in participants' experiences across interactions, the experimenter used a buzzer to notify prompters when they said "taboo" words aloud and docked points from their game score. The card deck in Study 1 contained an even number of easy, medium, and hard clues, while in Study 2 the deck was preset to have a higher ratio of medium and hard clues. In Study 1, participants were allowed to skip difficult cards at any time, regardless of whether or not they had started giving clues for the prompt word. In Study 2, participants were limited to skipping one card per turn and were forced to play any card they started giving clues for until their partner got the correct answer or the timer ran out. Finally, the gift card incentive was removed in Study 2. The gift cards were initially meant to motivate participants to fully engage with the activity and communicate with their assigned partner, but given participants' enjoyment of the task, I felt that it was unnecessary to include them in Study 2. Additionally, the idea of potentially winning a gift card may have interfered with participants' emotional experiences and attentiveness during the game. No alternate incentive was used in Study 2.

Upon arriving to the lab, participants were led to separate testing rooms to review the consent form, complete the social network questionnaire, and rate their baseline emotional experience. Next, they received instructions for the game and were randomly assigned to the experimental or control conditions, with their unassigned partner receiving the control instructions. The experimenter then led them to the same room where they reviewed the game

42

instructions once more and played four rounds together, with the experimenter keeping track of participants' scores and the timer while monitoring the prompter with a buzzer. After the end of the last round, participants returned to their original rooms to complete post-interaction measures on their emotional experience, expressivity attention, and likeability of their partner, along with their perceptions of their partner's experiences. Following completion of the demographics form, participants were debriefed on the study's purpose and compensated one course credit or $10.

7.3 Measures

All assessments were administered when the participants were in separate testing rooms.

7.3.1 Manipulation check

In total there were 58 control dyads (n = 116) and 106 experimental dyads (n = 212), with 53 Asian participants and 53 White participants assigned to the experimental condition. Self-reported partner expressivity focus was used to verify the experimental condition assignment. The same measure of expressivity focus from Study 1 was used in Study 2. Three items relating to attention towards partners' facial expressions, body language, and vocal quality were averaged to create a construct (α = .80; self-expressivity focus: α = .75).[5] Condition-assigned participants who scored below a three (i.e., reporting minimal to no attempt to attend to their partner's expressivity) were considered to have failed the manipulation check. Using this cutoff value, 12 experimental dyads were excluded from analyses due to failing the check, specifically six dyads where the Asian participant was assigned to the experimental condition and six dyads where the White participant was assigned to the experimental condition). This left 47 Asian experimental

---

[5] The items examining attention to speech were excluded given their impact on the constructs' reliability when included: partner expressivity (α = .71) and self-expressivity (α = .70).

and 47 White experimental dyads for analysis. The demographics of participants in the excluded

dyads were similar to those of participants who were in dyads that passed the manipulation check

(50.0% Asian and 50.0% White; M_{age} = 19.63, SD = 2.22; 58.3% female and 41.7% male).

After excluding the dyads that failed the manipulation check, one-way ANOVAs were

used to confirm if there were differences in self-rated partner expressivity attention between

participants assigned to the experimental and control conditions. These analyses compared those

who received the experimental condition instructions against the condition-assigned person who

received the control instructions as a result of this random assignment in each control dyad (i.e.,

excluding the unassigned participants from the experimental and control dyads). Condition-

assigned participants paid more attention to their partner's emotion expressions when they

received the experimental instructions (M = 4.78, SD = 1.19) than did control-assigned

participants (M = 4.19, SD = 1.44), F(1, 150) = 7.43, p = .007, η_p^2 = .05. An additional,

exploratory test was conducted to compare expressivity attention among the unassigned partners;

there was no difference in self-rated attention between unassigned partners in experimental (M =

4.07, SD = 1.38) and control dyads (M = 3.98, SD = 1.69), F(1, 150) = .15, p = .699, η_p^2 = .00.

The post-interaction measures also included four items asking participants about their

engagement with the activity (e.g., "I was focused on playing the game," α = .70). No

participants reported an engagement score below three (i.e., reporting minimal to no attempt to

attend to the game itself). There were no significant differences in task engagement between

condition-assigned participants who received the experimental (M = 6.05, SD = .75) or control

instructions (M = 6.03, SD = .73), F(1, 150) = .02, p = .880, η_p^2 = .00. I ran a similar exploratory

test for the unassigned partners in experimental and control dyads. There was a marginal effect

of dyad condition for unassigned partners, F(1, 150) = 3.18, p = .076, η_p^2 = .02; unassigned

participants in control dyads ($M = 6.25$, $SD = .65$) were slightly more engaged with the game

than were their counterparts in experimental dyads ($M = 6.03$, $SD = .77$).

7.3.2 Primary outcomes

7.3.2.1 Emotional experience. The measure used in Study 1 was expanded to include 16

emotions rated on a scale from 1 (*not* [feeling this emotion] *at all*) to 10 ([feeling this emotion] *a*

great deal). Two constructs, positive and negative emotion, were calculated by averaging

participants' ratings. Positive emotion included happy, satisfied, excited, enthusiastic, amused,

entertained, relaxed, calm, proud, and confident ($\alpha = .90$). Negative emotion included anxious,

nervous, frustrated, annoyed, embarrassed, self-conscious, bored, tired, sad, and disappointed (α

$= .81$). The newly added emotion items were included in this measure to provide participants

with a greater variety of options they could endorse regarding their emotional experiences during

the interaction. The scale was increased from a seven-point to a ten-point Likert scale to increase

the variance and sensitivity of their ratings.

7.3.2.2 Partner liking. The same measure was used from Study 1 to measure how much

participants liked their partner ($\alpha = .90$).

7.3.3 Primary predictors

7.3.3.1 Outgroup exposure. The same measure from Study 1 was used to examine

diversity across participants' social networks. Outgroup exposure was again operationalized as

the ratio of Asian or White social partners relative to the total size of one's social network. A t-

test was conducted to compare outgroup exposure between Asian and White participants, $t(302)$

$= -6.32$, $p < .001$, Cohen's $d = .67$. Asian participants listed more White social partners in their

network ($M = .20$, $SD = .18$) compared to the number of Asian social partners listed by White

participants ($M = .10$, $SD = .11$). See Figure 4 for a graph depicting the number of in- and outgroup social partners in participants' networks.

7.3.4 Other measures

      7.3.4.1 Demographics. The same measure from Study 1 was used to obtain basic demographic information about the participants prior to debriefing. Judge and target race were verified using participants' responses to the demographics questionnaire.

Chapter 8: Study 2 Results

8.1 Analysis Overview

The same type of analyses that were used in Study 1 were employed here, incorporating the expressivity condition assignment as a predictor of judgment accuracy. Judge- and target-ratings of emotional experience were used to determine judgment accuracy through profile correlations and Truth and Bias analyses. Within-dyad expressivity attention and outgroup exposure were tested as predictors of the emotion judgment profile correlations, and judgment accuracy was tested as a predictor of partner liking ratings using Actor-Partner Interdependence analyses with indistinguishable dyads. Judge and target condition (coded as 0 = control and 1 = experimental) and judges' race (0 = White, 1 = Asian), outgroup exposure and liking of targets were tested as moderators. Dyad gender was also tested as a moderator; however, there were minimal differences across groups.[6] The results presented here are collapsed across gender.

The inter-item correlation matrix and descriptives for Study 2's key variables are in Table 1. Figure 5 provides a graphed summary of participants' self-reported experiences during the interaction, split by dyad condition and participant race.

8.2 Profile Correlations

Mean profile correlations broken down by dyad condition and judge race are graphed in Figure 6. APIM estimates and effect sizes for Models 1 through 8 are provided in Tables 7, 8, and 9.

---

[6] Two thirds of the of the dyads (60.5%) were same-gender. Judges in mixed-gender dyads were better able to assess positive emotion compared to those in same-gender ones, $B = .14$, $SE = .07$, $p = .048$, $R_\beta^2 = .02$. Targets' emotional experience and judges' expressivity attention and liking of targets did not significantly differ between dyad types.

8.2.1 Condition assignment and judgment accuracy

Two APIM models were ran testing the effects of partners' condition assignment: Model 1 examined judge and target condition as predictors of the emotion judgment profile correlations, while Model 2 tested judge race as a moderator of these effects. See Table 7 for the estimates from these models.

In Model 1, neither judges', $p = .230$, nor targets' condition assignment, $p = .141$, predicted judges' perceptions of targets' emotional experience. Regardless of the instructions you and your partner received, your ability to accurately decode your partner's emotions was unaffected. Judge race had no moderating effect on these outcomes, ps $\geq .153$.

8.2.2 Outgroup exposure and judgment accuracy

Using APIM, I tested three models for outgroup exposure: Model 3 examines outgroup exposure within each dyad, while Models 4 and 5 test judge race and judge condition as moderators respectively. See Table 8 for the output of these models.

Outgroup exposure did not predict emotion judgment accuracy; judges' ability to accurately perceive targets' emotions was not associated with the number of outgroup members in judges', $p = .152$, and targets' social networks, $p = .271$. Neither judge race, ps $\geq .639$, nor judge condition assignment, ps $\geq .320$, moderated these effects.

8.2.3 Partner liking and judgment accuracy

I examined three APIM models testing partner liking. Model 6 uses the within-dyad profile correlations as predictors, while Models 7 and 8 respectively test judge race and judge condition as moderators. See Table 9 for the estimates from these models.

The more accurate judges were at assessing targets' emotional experiences, the more they liked the targets, $p < .001$, and the more these targets reported liking them, $p = .006$. In Model 7,

there was a marginal effect of judge race on judges' emotion judgment accuracy, $p = .099$. When examining the simple effects separately for Asian and White judges, all judges liked their partner more when they were better able to decode how their partner felt during the interaction; however, White judges' liking of Asian targets, $B = 1.77$, $SE = .27$, $p < .001$, $R_\beta^2 = .22$, increased more sharply in response to judgment accuracy compared to Asian judges' liking of White targets, $B = .96$, $SE = .20$, $p < .001$, $R_\beta^2 = .14$. In other words, accurate White judges liked their partner more compared to accurate Asian judges, and inaccurate White judges disliked their partner more than did inaccurate Asian judges. Judge race had no moderating effect on targets' ability to accurately perceive judges' emotional experience, $p = .996$.

In Model 8, there was a significant interaction between judges' accuracy and their condition assignment, $p = .049$. When examining the simple effects separately by judge condition, there were main effects of judgment accuracy on liking for judges in the control, $B = 1.32$, $SE = .20$, $p < .001$, $R_\beta^2 = .19$, and experimental conditions, $B = 1.06$, $SE = .31$, $p = .001$, $R_\beta^2 = .11$. Regardless of condition assignment, higher judgment accuracy was associated with increased partner liking; however, accurate judges in the control condition liked their partner more while inaccurate ones disliked their partner more compared to judges in the experimental condition, who comparatively held less polarized opinions of their partner. There was also a marginal interaction between targets' assessments of judges' emotional experience and judges' condition assignment, $p = .058$. When testing the simple effects split by judge condition, experimental and control judges who were accurately perceived by their partner liked their partner more and comparatively reported similar magnitudes of partner liking; however, judges in the experimental condition, $B = 1.06$, $SE = .31$, $p = .001$, $R_\beta^2 = .20$, were more sensitive to

being perceived inaccurately compared to control judges, $B = 1.32$, $SE = .20$, $p < .001$, $R_{\beta}^2 = .06$. Experimental judges disliked inaccurate targets much more than control judges did.

8.3 Truth and Bias Analyses

Estimates and effect sizes for the basic model (Model 1) and moderator models (Models 2 through 5) are presented in Tables 10 and 11.

8.3.1 Judgment accuracy for other-race interaction partners

Model 1 examined accuracy, directional bias, and assumed similarity for positive and negative emotion judgments collapsed across dyad types. Model 2 tested judge race as a moderator of these parameters. See Table 10 for the estimates for these models.

8.3.1.1 Judgment accuracy. Judges were able to accurately assess targets' positive and negative emotions, *p*s < .001. When testing judge race as a moderator of accuracy, there was a marginal effect for positive emotion judgments, $p = .068$. Asian judges were slightly better at assessing positive experience than were White judges when paired with an other-race target; however, judge race did not moderate accuracy for negative judgments, $p = .418$.

8.3.1.2 Judgment biases. Judges tended to underestimate targets' positive experience, $p < .001$, but did not systematically over- or underestimate their negative emotions, $p = .725$. Regardless of the valence of the emotions being assessed, judges relied heavily on their own emotional experience when determining how their partners were feeling, *p*s ≤ .003. When testing judge race as a moderator, there were effects for judgments of negative experience. Asian judges tended to overestimate negative emotion compared to assessments made by White judges, $p < .001$, and they marginally assumed White partners felt similar levels of negative emotion compared to their own experience, $p = .062$. Judge race did not moderate directional bias or assumed similarity for positive judgments, *p*s ≥ .110.

8.3.2 Condition assignment as a moderator of judgment accuracy

Model 3 tested judge condition as a possible moderator of judgment accuracy. See Table 10 for the output for this model.

Judges' condition assignment had no moderating effect on accuracy for positive, $p = .760$, or negative emotion judgments, $p = .636$. Regardless of the instructions they received, judges' ability to decipher targets' emotional experience was similar across conditions.

8.3.3 Outgroup exposure as a moderator of judgment accuracy

Model 4 tested judges' exposure to outgroups as a moderator of emotion judgment accuracy. See Table 11 for the estimates from this model.

Judges with more outgroup members in their social network were better able to assess outgroup targets' positive emotions compared to judges with low exposure, $p = .036$; however, judges' outgroup exposure had no moderating effect on their ability to accurately assess negative experience, $p = .695$.

8.3.4 Target liking as a moderator of judgment accuracy

Model 5 tested whether judges' ratings of target liking moderated judgment accuracy across all dyads. See Table 11 for the estimates from this model.

Judges' liking ratings had no moderating effect on accuracy for positive, $p = .590$, or negative emotion judgments, $p = .855$. The extent to which judges liked assigned partner was not associated with their ability to accurately assess their partners' emotional experience.

Chapter 9: Study 2 Discussion

Following the outcomes observed in Study 1, Study 2 was designed to take a more experimental approach to testing the factors relevant to judgments of outgroup members' emotions. In this study, I manipulated the extent to which participants monitored their other-race partners' expressions to see if increased attention to partner expressivity would improve judgment accuracy. Participants were able to successfully interpret their other-race partners' emotional experiences. This outcome partially replicates what we found in Study 1. Previously, only positive emotion judgments were accurate, while assessments made for negative emotion were not. The changes implemented to this study's design are a possible cause. While their negative affect was not quite as strong as their positive affect, participants in Study 2 reported feeling a greater variety of emotions in comparison to the Study 1 sample. Targets who felt more negative likely displayed more expressions relevant to these emotions during the interaction, providing judges with a better opportunity to detect and interpret verbal and nonverbal cues associated with negative emotion. Additionally, given the near floor-levels of negative emotion reported by Study 1 participants, the lack of variance in targets' self-ratings limited my ability to test judges' true assessment skills. This issue was minimized in Study 2 by making the interaction less affiliative and the game more difficult.

As was the case in Study 1, participants' judgments of their interaction partners were shaped by directional bias and assumed similarity. Judges' underestimated targets' positive experience and relied on their own experiences as a reference point when formulating positive and negative emotion judgments, replicating findings from Study 1. As one would expect, targets likely used multiple channels to convey emotional information to judges more effectively (Wieser & Brosch, 2012); however, these biases may indicate that judges failed to notice certain

valid behavioral cues or relied on some invalid ones when formulating their assessments. Despite these tendencies, Asian and White judges were still quite successfully at perceiving other-race targets' emotional experiences.

Although participants assigned to the experimental condition attended more to their partners' expressions relative to those in the control condition, this boosted expressivity focus did not benefit emotion judgment accuracy. In general, judges were quite attentive to their partners' expressions, as was the case in Study 1. Low coherence between targets' internal experience and external behavior may explain why condition-assigned partner expressivity focus had no moderating effect on judgment accuracy. If targets' observable behavior did not align well with their emotional experience, then paying more attention to target expressivity would have no meaningful effect on judgment accuracy. It would be interesting to see if increased monitoring of partner expressivity might be beneficial in situations where targets tried to genuinely express or perhaps exaggerate their emotions, as greater attention should be linked to noticing more relevant emotional cues (Funder, 1995; Nusseck et al., 2008). In the present study, however, being instructed to focus on partner expressivity attention was irrelevant to judgment accuracy.

Outgroup exposure was once again linked to judgment accuracy, but in this context it was associated with accurate judgments of positive emotion rather than negative emotion (as found in Study 1). In Study 2, judges with more outgroup members in their social network were more effective at decoding positive experience compared to judges with low exposure to outgroups. One possible explanation may be that participants in Study 1 generally felt high levels of positive emotion, making it more noticeable when targets expressed any negative emotion. If targets in that study tried to internalize their negative emotions, there may have been brief but obvious

instances where they openly showed negative expressions or perhaps displaying fewer positive expressions may have been unusual enough for judges to quickly notice.

Finally, judges liked targets who they could accurately perceive. However, unlike the outcomes from Study 1, this effect was bidirectional in Study 2. That is, judges also liked targets who in turn were able to accurately decode judges' emotional experiences. This outcome makes sense, given that being able to effectively read someone's thoughts and feelings should allow you to communicate with them more effectively and increase rapport. This process could be described as a positive feedback loop, in which judges continue to glean information from targets because they can tell that targets are responding favorably to their assessments and targets are actively providing more emotional information as judges are genuinely trying to understand targets' feelings. Interestingly, the effects of judgment accuracy on partner liking differed based on judges' condition assignment. Control judges liked targets more as a result of making more accurate assessments and disliked targets more when they were inaccurate compared to experimental judges. Although experimental judges were not more accurate at interpreting targets' emotional experience, their increased attention towards their partner may have provided them with a more complete picture of who their partner was (Hall et al., 2009), such as their personality or motives during the game. This added knowledge may have evened out experimental judges' liking ratings such that judges with more detailed information about their partner had difficulty deciding whether or not they actually liked their partner, leading to somewhat more balanced responses among this group.

Chapter 10: General Discussion

10.1 Project Summary

The purpose of this project was to test people's ability to accurately perceive the emotions of other-race partners following an in-person interaction compared to assessments made about same-race partners and to test two potential moderators of accuracy: expressivity attention and outgroup exposure. The findings from Studies 1 and 2 demonstrate that in this context people tend to be quite good at assessing outgroup members' emotional experience. In contrast to my expectations, Asian and White participants across both studies failed to show evidence of a persistent bias in favor of interpreting ingroup emotions over outgroup ones. People reported being moderately to highly attentive to their partners' expressivity regardless of the racial makeup of the dyad, but this monitoring did not play a key role in the judgment process. Outgroup exposure was associated with judgment accuracy, but the effects were inconsistent, as increased exposure was linked to better emotion judgments in some cases but not in others. Finally, people tended to like those who were more readily judged; however, partner liking was primarily associated with being able to correctly perceive one's partner and not necessarily with being accurately judged in return.

10.1.1 Explaining participants' judgment accuracy

Prior research on intergroup emotion communication has shown a consistent effect such that people are better able to judge the expressions of ingroup members over outgroup ones (e.g., Elfenbein & Ambady, 2002; Elfenbein et al., 2007; Gray, Mendes, Denny-Brown, 2008; Kang & Lau, 2013; Wickline et al., 2009; Young, Hugenberg, Bernstein, & Sacco, 2012). My findings instead demonstrate that this advantage does not necessarily translate to all types of interracial interactions. Across Studies 1 and 2, Asian and White participants were surprisingly effective at

judging positive and negative emotional experience, regardless of the race of their assigned partner. There may be a few reasons why I failed to find a boost in judgment accuracy among same-race dyads relative to mixed-race ones. The ingroup effect may be less pronounced in contexts featuring racial groups of similar status. Based on the dimensions described in Racial Position Model (Zou & Cheryan, 2017), the stereotypes Asians face are in part associated with the assumption that they are part of a high-status group (i.e., perceived to be competent, skilled, and high-achieving) relative to other non-White racial groups. Perceptions of racial superiority may be closely tied to individual group members' motivations to accurately perceive outgroup emotions, particularly if one's partner is also a member of a high-status group. In this project, Asian and White participants may have been very motivated to pay attention to and try to understand their other-race partners' emotions due to the similar status ascribed to these groups. If group status is influential to the magnitude of the ingroup advantage, it would be beneficial to test judgment accuracy among other types of interracial dyads. Doing so would allow researchers to see if this lack of an ingroup advantage occurs in interactions featuring members of racial groups that are similarly stereotyped as low-status (e.g., Blacks and Latinos) and whether participants from racial groups that differ in their perceived status (e.g., Blacks and Asians, Latinos and Asians) tend to have more difficulty making accurate emotion judgments.

Alternatively, the outcomes observed here may stem from the study design. Research examining the ingroup advantage in the context of a live interaction frequently makes race salient through the activity completed after meeting one's partner or the questionnaires participants fill out following the interaction. By doing so, researchers are increasing the likelihood that they activate concerns about group-specific stereotypes, leading to downstream consequences that can affect participants' ability to be emotionally and cognitively present in the

interaction and feel comfortable getting to know their other-race partner (Blascovich et al., 2001; Page-Gould et al., 2008; Richeson & Shelton, 2007; Shelton et al., 2006). Research has also demonstrated a tendency for individuals to exhibit a stereotype-congruent bias when race is salient during decoding of emotional other-race faces, such that same-race stimuli tend to be associated more readily with positive emotion and favorable characteristics while other-race stimuli are instead classified more negatively within subjects (Bijlstra, Holland, & Wigboldus, 2010; Hugenberg, 2005). This project was designed to avoid priming participants' racial identity, such that the consent forms, activity instructions, and measures used in Studies 1 and 2 purposefully omitted any mention of participants' race. This was done intentionally to ensure that participants were not overtly aware that they and their partner had been paired together based on their race. By avoiding priming race, the saliency of the racial makeup of the dyad may not have been as pressing to participants, allowing them to engage in the interaction without being burdened by race- and stereotype-specific concerns. An unintentional consequence of this design may have been minimization of the ingroup advantage as well. In the context of a live interaction, obtaining stronger effects for the ingroup advantage may require participants to be actively thinking about their own racial identity relative to that of their partner. One advantage to this study design may be that the interactions observed here are more reflective of a typical interracial interaction outside of laboratory settings, as real-world intergroup contact does not necessarily require those involved to be acutely aware of or purposefully discuss their racial group membership.

Another explanation may relate to group memberships among the study samples. For the purposes of this project, I focus on racial identity as the primary distinguishing variable between participants; however, there may have been other overarching groups whose salience superseded

that of participant race. The study samples I recruited primarily came from Washington University's campus and the immediate surrounding area. Almost all participants were affiliated with the university in some way. Given that participants were not primed to actively think about racial group membership, their connection to the university itself may have become the most prominent shared group identity between partners. People have multiple identity categories beyond race and ethnicity, and situational factors influence which identity is most salient (Hogg & Turner, 1987). Someone who is an outgroup member in one context could be recategorized as an ingroup member by finding a shared group membership. This transformation can eliminate distinctions between groups, leading people to treat outgroup members as if they were part of their ingroup all along (Gaertner & Dovidio, 2014). By recognizing that they share a common ingroup identity related to university affiliation, participants may have essentially overridden any effects that would have been associated with racial group membership and approached the situation as if it were an ingroup interaction rather than an intergroup one. Prior studies have found that people are quick to mentally recategorize outgroup members as part of their ingroup when given the chance; arbitrarily defined and assigned social groups among participants and stimuli demonstrate this effect (Bernstein, Young, & Hugenberg, 2007; Sheng & Han, 2012), as do personality types (Young & Hugenberg, 2010) and, importantly, university affiliation (Hehman, Mania, & Gaertner, 2010; Stevenson, Soto, & Adams, 2012).

10.1.2 Explaining participants' judgment biases

Across both studies, I found that people's emotion judgments of ingroup and outgroup members were typically influenced by directional bias and assumed similarity. These effects may have occurred in part due to the nature of the task used in both studies. It can be very difficult to fully get to know someone after ten minutes, particularly if most of that time is spent playing a

card game. This context provides partners with limited opportunities to observe and share

information about their personal emotions. Also if there was the low coherence between targets'

internal experience and their external expressions, judges would have needed to rely on a

combination of verbal and nonverbal cues to formulate their assessments and possibly discount

certain behaviors to accurately perceive how their partners truly felt (Jack, 2013; Nusseck et al.,

2008; Wieser & Brosch, 2012). Any observable mismatch between felt and displayed emotions

may have also caused judges to rely more on their own experience as a reference on which they

based their assessments. Cognitive literature describes our tendency to overemphasize initially

available information in subsequent judgments as the *anchoring effect* (Furnham & Boo, 2011;

Tversky & Kahneman, 1974). In this context, judges may have used their own emotional

experience as a readily available and simple heuristic for determining a reasonable comparison

for targets' experiences, particularly if targets failed to display enough valid emotional cues

(Funder et al., 1995; Watson et al., 2000b).

10.1.3 Explanations for the expressivity attention findings

There are a few possible explanations behind the lack of a moderating effect for partner

expressivity focus on judgment accuracy. As mentioned previously, the task used in this project

may have boosted participants' willingness to attend to their partner and respond to them

accordingly. The game requires two people to play, and both partners must actively exchange

information to be able to successfully guess correct answers. Having to complete this activity

with a stranger for an extended period of time may have provided the right conditions for

participants to generally default to paying more attention to their partner. Additionally, low

coherence between targets' experience and expression may again be to blame. If targets failed to

display observable, valid expressive cues throughout the interaction or if judges were focusing

on a high proportion of invalid cues over valid ones, then judges' attentiveness to target expressivity would not benefit their ability to formulate accurate assessments. To determine if either of these situations were at play, we would need to have data examining the behaviors targets showed (or failed to show) during the interaction.

10.1.4 Explanations for the outgroup exposure findings

Outgroup exposure was partially linked to accurate emotion judgments of other-race interaction partners, but it did not have a consistent, strong effect across Studies 1 and 2. Prior work has shown that increased exposure to people of other races, be it through individualized contact or simply living in a diverse region, decreases the ingroup advantage of emotion recognition (Elfenbein & Ambady, 2003; Soto & Levenson, 2009; Yan et al., 2016). Regular encounters with people of other races should provide multiple opportunities for one to observe group-specific expressive cues, gaining knowledge about groups' idiosyncratic behavioral norms and proficiency at decoding group members' emotions. In the present study, social network diversity was demonstrated to be a valid proxy of outgroup exposure. Participants with more outgroup members in their social network were more effective at assessing emotional experience when paired with a stranger from the same racial outgroup. However, it is important to note that the effect of outgroup exposure on judgment accuracy differed across Studies 1 and 2. Why might this outcome have occurred? The definition I used to operationalize outgroup exposure may have been too broad. Familiarity with outgroup behavioral norms may not be best characterized by the number of outgroup members in one's social network. Rather, relationship quality might be more telling of how knowledgeable and effective people are at identifying and decoding outgroup emotions (Letzring et al., 2006). Participants reported a variety of different relationships with people of other races in their networks, including friends, coworkers, and

family members. Relationship type may restrict the applicability of witnessed outgroup

behaviors. For example, the behaviors you observe with an other-race roommate will likely be

different from the behaviors you would see if you had a coworker of the same racial background.

Familiarity with outgroup emotion experience and expression is likely contingent on the quality

of the relationship and whether it allows for chances to receive feedback. One would expect that

maintaining close relationships with outgroup members provides more opportunities to witness

examples of outgroup behavior in a range of situations and greater exchange of feedback

compared to less close relationships (Beaupré & Hess, 2006; Elfenbein & Ambady, 2003;

Hugenberg et al., 2007; Meissner & Brigham, 2001). Going forward, it may be helpful to focus

more on perceived closeness with outgroup members or potentially ask participants to rate the

quality of their relationships as an alternate way to characterize exposure.

One other item to consider regarding this topic is alternate sources of intergroup contact,

namely "passive" exposure. Global changes in sociodemographic factors, such as increased

interconnectedness and technological advancements, are constantly reshaping the way we

interact with and think about people of other cultures (Greenfield, 2016). People may see

outgroup members on a regular basis across different contexts but do not directly interact with

them. For example, if you live in a culturally diverse region you may see people of different

races or ethnicities during your daily commute or while grocery shopping. Time spent living in

regions with sizable outgroup populations (e.g., Rhodes et al., 2009; Soto & Levenson, 2009)

and experimentally induced exposure (e.g., Sporer, 2001; Walker et al., 2008; Yankouskaya et

al., 2014) both lead to better recognition of emotional outgroup faces. Alternatively, watching

television shows featuring people from different cultural backgrounds or foreign movies can

provide viewers with examples of outgroup members interacting with one another or with fellow

ingroup members. Media is a growing source of remote acculturation that is likely to become more pervasive across cultures as time continues (Ferguson, Tran, Mendez, & van de Vijver, 2017). Depending on where you live or the activities you are involved in, you may have a plethora of opportunities to witness outgroup behavioral cues without needing to go out of your way to find outgroup members. It would be beneficial for researchers to account for these factors by asking participants to characterize the racial demographics of the neighborhood they grew up in and where they currently live or perhaps describe their media consumption.

10.1.5 Explaining participants' liking of interaction partners

Finally, I demonstrated across both studies that being able to accurately judge one's partner is associated with greater liking of this partner. Interestingly, this outcome persisted across dyad types, and partner liking did not moderate emotion judgment ability. Interpersonal responsiveness is key to relationship development; attending to and responding accordingly to one's partner facilitates interconnectedness and increases rapport among partners (Laurenceau et al., 1998; Reis & Shaver, 1988). By making accurate assessments of targets' emotions, judges signal to targets that they are paying attention to what targets are communicating and actively engaged in the interaction. In Study 1, being able to determine what your partner was feeling was enough to promote liking, but in Study 2, I observed a more bidirectional association between being able to accurately judge one's partner and being accurately judged in turn by said partner. Given the nature of the task used in both studies, these effects may have partially been driven by the context, particularly since being able to accurately perceive what your partner was communicating was essential to succeeding at the game. Had the interactions featured an alternate activity, perhaps one requiring personal disclosure from each participant such as sharing a recent negative event or a favorite memory, participants' liking ratings may have been

more dependent on their partners' ability to accurately decode emotion as well as their own assessment skills (Willems, Finkenauer, & Kerkhof, 2020).

10.2 Critiques of the Project Design

10.2.1 Advantages of the project

10.2.1.1 Studying judgment accuracy using a live interaction. Using an in-lab study session to test judgment accuracy provides a realistic context for people to gather information about their interaction partner prior to formulating their judgments. Movement improves our ability to judge emotion expressions (Krumhuber, Kappas, & Manstead, 2013; Nusseck et al., 2008). Additionally, studies examining the ingroup advantage of emotion recognition often feature posed or exaggerated expressions instead of spontaneous ones. There is conflicting evidence as to whether the ingroup advantage is improved by presenting disingenuous expressions (e.g., Matsumoto, Olide, & Willingham, 2009) or naturally occurring ones (e.g., Kang & Lau, 2013). However, there are more cases where using photos of unnatural expressions yielded stronger effects for the ingroup advantage (Elfenbein & Ambady, 2002). By providing judges with dynamic stimuli depicting realistic expressions, this project increases the ecological validity and generalizability of its findings. Focusing on accuracy in the context of an in-person interaction also increases both judge and target engagement. Judges are more likely to be motivated to look for and interpret targets' emotional cues given that they need to respond to targets in real time. Targets are more likely to feel stronger emotions during the interaction and may be more motivated to openly express these feelings. This context also allows judges and targets to change their approaches if things are not working out, such that judges can ask targets for more information to confirm if their assessments are accurate and targets can correct judges if they realize that judges have made an erroneous assessment. This design also allows researchers

to collect data for both partners, as opposed to requiring one person to take on the role of judge and the other to be the target. Having both perspectives provides a much more holistic and complex look into the bidirectional exchange of information occurring during the interaction.

The design of the activity itself is flexible; depending on the experiences you ideally want participants to undergo when meeting their partner, certain aspects of the study can be easily changed to achieve these goals. Another benefit of this activity is that it allows us to test intergroup dynamics without overtly involving race. Research on intergroup interactions frequently primes people to think about race (e.g., being instructed to discuss a race-related topic with an other-race interaction partner). As one would expect, making race salient can make a situation more difficult to navigate, particularly if people feel uncomfortable sharing their emotional experience, have trouble empathizing with their partner, or are distracted by concerns about stereotype confirmation and threat (Leary et al., 1995; Richeson & Shelton, 2007; Shelton et al., 2006; Vorauer, Martens, & Saski, 2009). By examining judgment accuracy outside of these contexts, this project tests people's ability to assess outgroup emotions in a more generalizable and socially neutral context.

10.2.1.2 Using multiple analysis techniques to assess accuracy. Analyzing emotion judgments via profile correlations and Truth and Bias analyses provides different types of information about accuracy. Profile correlations demonstrate the overall magnitude of accuracy or inaccuracy, but they fail to describe underlying patterns in the data. Truth and Bias analyses solve this issue by allowing researchers to simultaneously examine judgment accuracy and biases within the same model and easily test moderators of accuracy. Also, correlations focus on relative, rank-order position in a way that ignores level, whereas TBM looks at discrepancies in a way that accounts for level. Using a combination of both is an ideal solution to the limitations of

each specific method and helps us better understand how effective people are at making interpersonal judgments and potential sources of inaccuracy. Additionally, research examining intergroup judgments has primarily focused on testing mean-level ratings or absolute difference scores to compare judge and target ratings across racial groups (e.g., Kang & Lau, 2013; Ma-Kellams & Blascovich, 2012; Rogers & Biesanz, 2014). Few studies use techniques such as TBM to analyze judgment accuracy between interracial partners (e.g., Holoien, et al., 2015; West, Dovidio, & Pearson, 2014). This project may be one of the first to employ both profile correlations and TBM in a systematic way to test judgment accuracy across all studies.

 10.2.1.3 Examining a novel type of intergroup interaction. Dyads featuring East Asian and White participants, particularly those residing in close proximity to each other, are less commonly studied. Historically, researchers examining the ingroup advantage of emotion recognition have focused on international comparisons using groups such as Caucasians from the U.S., Canada, Europe, or Australia and indigenous people in Asia (e.g., China: Zhang, Parmley, Wann, & Cavanagh, 2015; Hong Kong: Zhu, Ho, and Bonnano, 2013; Japan: Dailey et al., 2010) or Africa (e.g., Gabon: Elfenbein et al., 2007; Namibia: Gendron et al., 2014). In recent decades, studies have begun focusing more on testing groups living within the same nation; however, the majority of work based in the U.S. tests judgment accuracy among Black and White Americans (e.g., Butz & Plant, 2006; Kunstman, Tuscherer, Trawalter, & Lloyd, 2016; Malloy et al., 2011; Shelton, Richeson, & Salvatore, 2005; Shelton, Richeson, Salvatore, & Trawalter, 2005). By focusing on interactions featuring Asians and Whites, the present study provides researchers with more information about a relatively understudied type of interracial interaction. Additionally, these groups have their own idiosyncrasies regarding emotional experience, expression, and regulation along with stereotypes they face when interacting with an other-race partner (Gross et

al., 2006; Sue et al., 2007; Tsai, 2007). This project may be particularly relevant to future work testing the ingroup advantage of emotion recognition between Asians and Whites, given that my findings demonstrate that these groups can be quite accurate at decoding outgroup emotions given the right context.

10.2.2 Limitations of the project

There are a few additional limitations to this project that should be discussed. First, for the purposes of this study, participants from a wide range of ethnic and cultural backgrounds were grouped under the umbrella term "East Asian." The demographics of the samples of Asian participants in Studies 1 and 2 were not identical. Lumping together a diverse population may have masked effects that only applied to certain subgroups within the sample. There are likely to be other relevant factors, such as nationality, generational status, and acculturation, that shape the way people who could be classified as East Asian experience and express emotion and the stereotypes they face during interracial interactions.

While the activity used in Studies 1 and 2 worked for my purposes, there were some disadvantages. People tended to get easily caught up in the game, intentionally or unintentionally, which likely affected the emotions they experienced and how much effort they put into trying to decode their partner's emotions. Additionally, some participants were familiar with the original game that the activity was based on (Taboo), while others were not, giving some people a slight advantage in their performance and the degree to which they felt stressed or cognitively taxed by the game rules. Finally, certain stereotypes may be primed by this specific activity that were not measured, including vocabulary knowledge, the ability to think on the spot, and being able to translate concepts to others in a clear and concise way. By failing to set eligibility requirements for language proficiency, dyads with participants who were more

comfortable with the English language may have been more successful at playing the game compared to dyads in which one or both participants had limited fluency. Also, while this activity may be representative of interactions requiring cooperation and communication between partners, it is a relatively unusual task to perform outside of a lab setting, somewhat limiting the generalizability of my findings to more naturalistic contexts.

To be able to test my hypotheses, this project relied on the assumption that targets would openly convey their emotions to judges during the interaction. It may be the case that people communicate emotions differently when paired with an ingroup or outgroup member, relying more heavily on verbal or nonverbal cues or a unique combination of both to express how they feel depending on the race of their interaction partner. Low behavioral expressivity may be an alternate explanation for decreased emotion judgment accuracy. If targets' expressions did not align well with their internal emotional experience, judges may have had very little or even erroneous behavioral information to work with when trying to decipher how targets were feeling. In this project, targets' subjective report of their emotions was the criterion for judgment accuracy. Our internal emotional experiences are highly individualized and the emotions we feel are not always displayed in equal measure through our external behavior (Gross, John, & Richards, 2000; Mauss, Levenson, McCarter, Wilhelm, & Gross, 2005). To best understand what kind of internal experience targets undergo during the interaction, I chose to ask them to provide a personal description of what they are actively feeling. It may be beneficial for researchers studying these topics to supplement emotion judgment data with behaviorally coded body language, vocal quality, or speech content, to have a wider range of expressive data to compare against targets' subjective responses and judges' perceptions of target emotional experience. Doing so would allow us to better understand what information is available to judges when they

make their assessments, particularly the unique contribution of different expressive channels, and whether judges rely on more than just facial expressions when decoding targets' experiences (Meeren et al., 2005; Scherer et al., 2011; Wieser & Brosch, 2012).

10.3 Future Directions

There are several research questions that should be addressed in the near future. How would these results change if we selected different racial groups? Most studies testing outgroup judgment accuracy predominantly compare Whites against non-Whites (e.g., Butz & Plant, 2006; Holoien et al., 2015; Kang & Lau, 2013; Kunstman et al., 2016; Malloy et al., 2011; Ma-Kellams & Blascovich, 2012; Shelton et al., 2005; West et al., 2014). By examining these processes among more varied populations, we can better understand how generalizable these findings are to other types of interracial dyads and contexts, particularly ones featuring interactions among non-White racial groups. It would also be interesting to test judgment accuracy of outgroup members in more complex interactions and relationships, such as acquainted dyads, groups featuring three or more people, or longitudinal interactions where participants develop a relationship with their partner over time. These studies may elucidate how people learn about and develop competency with decoding outgroup behavioral norms. I am also interested in testing other types of accuracy criterion, particularly with detailed behavioral coding of bodily and vocal cues. Having a clearer sense of the information judges obtain from different expressive channels and the number of valid and invalid cues involved in emotion judgments would help us better understand sources of inaccuracy in this context. Finally, I would recommend incorporating additional self-other measures to examine different kinds of interpersonal judgments. Emotions are ideal for this particular project given that they are relatively easy to elicit, occur rapidly, and are often characterized by distinct behavioral cues (Lewis, Haviland-

Jones, & Barrett, 2010). Comparing the expressive cues associated with other types of state- and trait-level judgments during interracial interactions would provide more information about the types of judgments that may be subject to ingroup biases.

10.4 Conclusion

This project examines how we perceive the emotions of racial outgroup members during social interactions and the factors that shape our ability to form accurate assessments. Prior research on intergroup interactions and interpersonal judgments demonstrates that we tend to be more proficient at decoding the emotions of ingroup members; however, this work has not yet established which steps in the judgment process are linked to our tendency to misinterpret and misattribute outgroup characteristics. Across two studies, I demonstrated that people can be surprisingly proficient at judging other-race interaction partners' emotional experience and that certain factors, such as the number of outgroup members in one's social network, are more relevant to accurate judgments than others. This project advances our understanding of emotion communication during intergroup interactions by providing a nuanced examination of the mechanisms involved in judgment formation and describing important implications for the way we conceptualize and study intergroup emotion judgments.

References

Adams, R. B., Ambady, N., Macrae, C. N., & Kleck, R. E. (2006). Emotional expressions

forecast approach-avoidance behavior. *Motivation and Emotion, 30*(2), 177-186.

https://doi.org/10.1007/s11031-006-9020-2

Al Ramiah, A., & Hewstone, M. (2013). Intergroup contact as a tool for reducing, resolving, and

preventing intergroup conflict: Evidence, limitations, and potential. *American

Psychologist*, *68*(7), 527-542. https://doi.org/10.1037/a0032603

Allport, G. W. (1954). *The nature of prejudice.* Reading, MA: Addison-Wesley.

Ambady, N., Bernieri, F. J., & Richeson, J. A. (2000). Toward a histology of social behavior:

Judgmental accuracy from thin slices of the behavioral stream. *Advances in Experimental

Social Psychology, 32*(1), 201-271. https://doi.org/10.1016/S0065-2601(00)80006-4

Amir, Y. (1976). The role of intergroup contact in change of prejudice and race relations. In P.

Katz & D. A. Taylor (Eds.), *Towards the elimination of racism* (pp. 245-308). New York,

NY: Pergamon Press.

Antonucci T. C. (1986). Measuring social support networks: Hierarchical mapping technique.

Generations: The Journal of the Western Gerontological Society, *10*(1), 10-12.

Anzures, G., Quinn, P. C., Pascalis, O., Slater, A. M., Tanaka, J. W., & Lee, K. (2013).

Developmental origins of the other-race effect. *Current Directions in

Psychological Science, 22*(3), 173–178. https://doi.org/10.1177/0963721412474459

Atkinson, A. P., Tunstall, M. L., & Dittrich, W. H. (2007). Evidence for distinct contributions of

form and motion information to the recognition of emotions from body gestures.

Cognition, 104(1), 59-72. https://doi.org/10.1016/j.cognition.2006.05.005

Barrett, L. F. (2006). Solving the emotion paradox: Categorization and the experience of emotion. *Personality and Social Psychology Review*, *10*(1), 20-46. https://doi.org/10.1207/s15327957pspr1001

Baumeister, R. F. (1982). A self-presentational view of social phenomena. *Psychological Bulletin, 91*(1)*,* 3-26. https://doi.org/10.1037/0033-2909.91.1.3

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin, 117*(3)*,* 497-529. https://doi.org/10.1037/0033-2909.117.3.497

Beaupré, M. G., & Hess, U. (2006). An ingroup advantage for confidence in emotion recognition judgments: The moderating effect of familiarity with the expressions of outgroup members. *Personality and Social Psychology Bulletin*, *32*(1), 16-26. https://doi.org/10.1177/0146167205277097

Bergsieker, H. B., Shelton, J. N., & Richeson, J. A. (2010). To be liked versus respected: Divergent goals in interracial interactions. *Journal of Personality and Social Psychology*, *99*(2), 248-264. https://doi.org/10.1037/a0018474

Bernstein, M. J., Young, S. G., & Hugenberg, K. (2007). The cross-category effect. Mere social categorization is sufficient to elicit an own-group bias in face recognition. *Psychological Science, 18*(8), 706-712. https://doi.org/10.1111/j.1467-9280.2007.01964.x

Bijlstra, G., Holland, R. W., & Wigboldus, D. H. (2010). The social face of emotion recognition: Evaluations versus stereotypes. *Journal of Experimental Social Psychology, 46*(4)*,* 657-663. http://doi.org/10.1016/j.jesp.2010.03.006

Blais, C., Jack, R. E., Scheepers, C., Fiset, D., & Caldara, R. (2008). Culture shapes how we look at faces. *PloS One*, *3*(8), e3022. https://doi.org/10.1371/journal.pone.0003022

Blascovich, J., Mendes, W. B., Hunter, S. B., Lickel, B., & Kowai-Bell, N. (2001). Perceiver
threat in social interactions with stigmatized others. *Journal of Personality and Social
Psychology*, *80*(2), 253-267. https://doi.org/10.1037/0022-3514.80.2.253

Brielmann, A. A., Bülthoff, I., & Armann, R. (2014). Looking at faces from different angles:
Europeans fixate different features in Asian and Caucasian faces. *Vision Research,
100*(1), 105-112. https://doi.org/10.1016/j.visres.2014.04.011

Bukach, C. M., Cottle, J., Ubiwa, J., and Miller, J. (2012). Individuation experience predicts
other-race effects in holistic processing for both Caucasian and Black participants.
Cognition, *123*(2), 319-324. https://doi.org/10.1016/j.cognition.2012.02.007

Butz, D. A., & Plant, E. A. (2006). Perceiving outgroup members as unresponsive: Implications
for approach-related emotions, intentions, and behavior. *Journal of Personality and
social Psychology*, *91*(6), 1066-1079. https://doi.org/10.1037/0022-3514.91.6.1066

Cacioppo, J. T., Berntson, G. G., Larsen, J. T., Poehlmann, K. M., & Ito, T. A. (2000). The
psychophysiology of emotion. In R. Lewis & J. M. Haviland-Jones (Eds.), *The handbook
of emotion* (2nd ed., pp. 173-191). New York, NY: Guilford Press.

Cheryan, S., & Monin, B. (2005). Where are you really from?: Asian Americans and identity
denial. *Journal of Personality and Social Psychology*, *89*(5), 717-730.
https://doi.org/10.1037/0022-3514.89.5.717

Colvin, C. R. (1993). " Judgable" people: Personality, behavior, and competing explanations.
Journal of Personality and Social Psychology, *64*(5), 861-873.
https://doi.org/10.1037/0022-3514.64.5.861

Connelly, B. S., & Ones, D. S. (2010). An other perspective on personality: Meta-analytic

    integration of observers' accuracy and predictive validity. *Psychological Bulletin*, *136*(6),

    1092-1122. https://doi.org/10.1037/a0021212

Cook, W. L., & Kenny, D. A. (2005). The actor–partner interdependence model: A model of

    bidirectional effects in developmental studies. *International Journal of Behavioral*

    Development, *29*(2), 101-109. https://doi.org/10.1080/01650250444000405

Cortes, B. P., Demoulin, S., Rodriguez, R. T., Rodriguez, A. P., & Leyens, J. P. (2005).

    Infrahumanization or familiarity? Attribution of uniquely human emotions to the self, the

    ingroup, and the outgroup. *Personality and Social Psychology Bulletin*, *31*(2), 243-253.

    https://doi.org/10.1177/0146167204271421

Craig, M. A., & Richeson, J. A. (2012). Coalition or derogation? How perceived discrimination

    influences intraminority intergroup relations. *Journal of Personality and Social*

    Psychology, 102()*,* 759-777. http://doi.org/10.1037/a0026481

Cuddy, A. J., Fiske, S. T., & Glick, P. (2007). The BIAS map: behaviors from intergroup affect

    and stereotypes. *Journal of Personality and Social Psychology*, *92*(4), 631-648.

    https://doi.org/10.1037/0022-3514.92.4.631

Dael, N., Mortillaro, M., & Scherer, K. R. (2011). Emotion expression in body action and

    posture. *Emotion*, *12*(5), 1085-1101. https://doi.org/10.1037/a0025737

Dailey, M. N., Joyce, C., Lyons, M. J., Kamachi, M., Ishi, H., Gyoba, J., & Cottrell, G. W.

    (2010). Evidence and a computational explanation of cultural differences in facial

    expression recognition. *Emotion*, *10*(6), 874-893. https://doi.org/10.1037/a0020019

de Gelder, B. (2006). Towards the neurobiology of emotional body language. *Nature Reviews*

    Neuroscience, *7*(3), 242-249. https://doi.org/10.1038/nrn1872

de Gelder, B., & Huis in 't Veld, E. (2016). Cultural differences in emotional expressions and body language. In J. Y. Chiao, S.-C. Li, R. Seligman, & R. Turner (Eds.), *Oxford library of psychology: The Oxford handbook of cultural neuroscience* (p. 223-234). New York, NY: Oxford University Press.

Edwards, L., J., Muller, K. E., Wolfinger, R. D., Qaqish, B. F., & Schabenberger, O. (2008). An R^2 statistic for fixed effects in the linear mixed model. *Statistics in Medicine, 27*(29), 6137-6157. https://doi.org/10.1002/sim.3429

Ekman, P. (1992a). An argument for basic emotions. *Cognition and Emotion*, *6*(3-4), 169-200. https://doi.org/10.1080/02699939208411068

Ekman, P. (1992b). Facial expression of emotion: New findings, new questions. *Psychological Science*, *3(1)*, 34-38. https://doi.org/10.1111/j.1467-9280.1992.tb00253.x

Eldesouky, L., English, T., & Gross, J. J. (2017). Out of sight, out of mind? Accuracy and bias in emotion regulation trait judgments. *Journal of Personality*, *85*(4), 543-552. https://doi.org/10.1111/jopy.12259

Elfenbein, H. A. (2013). Nonverbal dialects and accents in facial expressions of emotion. *Emotion Review*, *5*(1), 90-96. https://doi.org/10.1177/1754073912451332

Elfenbein, H., A. & Ambady, N. (2002). Is there an in-group advantage in emotion recognition? *Psychological Bulletin, 128*(2), 243-249. https://doi.org/10.1037/0033-2909.128.2.243

Elfenbein, H. A., & Ambady, N. (2003). When familiarity breeds accuracy: Cultural exposure and facial emotion recognition. *Journal of Personality and Social Psychology*, *85*(2), 276-290. https://doi.org/10.1037/0022-3514.85.2.276

Elfenbein, H. A., Beaupré, M., Lévesque, M., & Hess, U. (2007). Toward a dialect theory: Cultural differences in the expression and recognition of posed facial expressions. *Emotion*, *7*(1), 131-146. https://doi.org/10.1037/1528-3542.7.1.131

Elliott, E. S., Wills, E. J., & Goldstein, A. G. (1973). The effects of discrimination training on the recognition of white and oriental faces. *Bulletin of the Psychonomic Society, 2*(1), 71-73. https://doi.org/10.3758/BF03327717

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(1), 175-191. https://doi.org/10.3758/BF03193146

Ferguson, G. M., Tran, S. P., Mendez, S. N., & van de Vijver, F. J. (2017). Remote acculturation: Conceptualization, measurement, and implications for health outcomes. In S. J. Schwartz, & J. B. Unger (Eds.), *The Oxford handbook of acculturation and health*. New York, NY: Oxford University Press.

Fiske, S. T., Cuddy, A. J. C., Glick, P. S., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology, 82*(6)*, 878-902. https://doi.org/10.1037/pspa0000163

Fiske, S. T., Xu, J., Cuddy, A. J. C., & Glick, P. (1999). (Dis)respecting versus (dis)liking: Status and interdependence predict ambivalent stereotypes of competence and warmth. *Journal of Social Issues, 55*(3)*, 473-489. https://doi.org/10.1111/0022-4537.00128

Forbes, H. D. (1997). *Ethnic conflict: Commerce, culture, and the contact hypothesis.* New Haven, CT: Yale University Press.

Frantz, C. M., Cuddy, A. J. C., Burnett, M., Ray, H., & Hart, A. (2004). A threat in the

computer: The Race Implicit Association Test as a stereotype threat experience.

Personality and Social Psychology Bulletin, 30(12), 1611-1624.

https://doi.org/10.1177/0146167204266650

Freeman, J. B., Schiller, D., Rule, N. O., & Ambady, N. (2010). The neural origins of superficial

and individuated judgments about ingroup and outgroup members. *Human Brain*

Mapping, *31*(1), 150-159. https://doi.org/10.1002/hbm.20852

Fu, G., Hu, C. S., Wang, Q., Quinn, P. C., & Lee, K. (2012). Adults scan own- and other-race

faces differently. *PLoS One, 7*(6), e37688. https://doi.org/10.1371/journal.pone.0037688

Funder, D. C. (1995). On the accuracy of personality judgment: A realistic approach.

Psychological Review, *102*(4), 652-670. https://doi.org/10.1037/0033-295X.102.4.652

Funder, D. C., Kolar, D. C., & Blackman, M. C. (1995). Agreement among judges of

personality: Interpersonal relations, similarity, and acquaintanceship. *Journal of*

Personality and Social Psychology, 69(4), 656-672. https://doi.org/10.1037//0022-

3514.69.4.656

Furnham, A. & Boo, H.C. (2011) A literature review of the anchoring effect. *The Journal of*

Socio-Economics, *40*(1), 35-42. https://doi.org/10.1016/j.socec.2010.10.008

Gaertner, S. L., & Dovidio, J. F. (2014). *Reducing intergroup bias: The common ingroup identity*

model. New York, NY: Psychology Press.

Gagné, F. M., & Lydon, J. E. (2004). Bias and accuracy in close relationships: An integrative

review. *Personality and Social Psychology Review*, *8*(4), 322-338.

https://doi.org/10.1207/s15327957pspr0804_1

Gendron, M., Roberson, D., van der Vyver, J. M., & Barrett, L. F. (2014). Perceptions of

    emotion from facial expressions are not culturally universal: evidence from a remote

    culture. *Emotion*, *14*(2), 251-262. https://doi.org/10.1037/a0036052

Goldstein, A. G., & Chance, J. E. (1985). Effects of training on Japanese face recognition:

    Reduction of the other-race effect. *Bulletin of the Psychonomic Society, 23*(1), 211-214.

    https://doi.org/10.3758/BF03329829

Gray, H. M., Mendes, W. B., & Denny-Brown, C. (2008). An in-group advantage in detecting

    intergroup anxiety. *Psychological Science*, *19*(12), 1233-1237.

    https://doi.org/10.1111/j.1467-9280.2008.02230.x

Greenfield, P. M. (2016). Social change, cultural evolution, and human development. *Current*

    Opinion in Psychology, *8*(1), 84-92. https://doi.org/10.1016/j.copsyc.2015.10.012

Gross, J. J., John, O. P., & Richards, J. M. (2000). The dissociation of emotion expression from

    emotion experience: A personality perspective. *Personality and Social Psychology*

    Bulletin, 26(6), 712-726. https://doi.org/10.1177/0146167200268006

Gross, J. J., Richards, J. M., & John, O. P. (2006). Emotion Regulation in Everyday Life. In D.

    K. Snyder, J. Simpson, & J. N. Hughes (Eds.), *Emotion regulation in couples and*

    families: Pathways to dysfunction and health (pp. 13-35). Washington, DC, US:

    American Psychological Association.

Hall, J.A., Blanch, D.C., Horgan, T.G., Murphy, N.A., Rosip, J.C., & Schmid Mast, M. (2009).

    Motivation and interpersonal sensitivity: Does it matter how hard you try? *Motivation*

    and Emotion, *33*(1), 291-302. https://doi.org/10.1007/s11031-009-9128-2

Hancock, K. J., & Rhodes, G. (2008). Contact, configural coding and the other-race effect in face

    recognition. *British Journal of Psychology*, *99*(1), 45-56.

    https://doi.org/10.1348/000712607X199981

Harrigan, J. A. (2005). Proxemics, kinesics, and gaze. In J. A. Harrigan, R. Rosenthal, & K.

    Scherer (Eds.), *The new handbook of methods in nonverbal behavior research*. New

    York, NY: Oxford University Press.

Hehman, E., Mania, E. W., & Gaertner, S. L. (2010). Where the division lies: Common ingroup

    identity moderates the cross-race facial-recognition effect. *Journal of Experimental*

    Social Psychology, *46*(2), 445-448. https://doi.org/10.1016/j.jesp.2009.11.008

Hogg, M. A., & Turner, J. C. (1987). Intergroup behaviour, self-stereotyping and the salience of

    social categories. *British Journal of Social Psychology*, *26*(4), 325-340.

    https://doi.org/10.1111/j.2044-8309.1987.tb00795.x

Holoien, D. S., Bergsieker, H. B., Shelton, J. N., & Alegre, J. M. (2015). Do you really

    understand? Achieving accuracy in interracial relationships. *Journal of Personality and*

    Social Psychology, *108*(1), 76-92. https://doi.org/10.1037/pspi0000003

Hugenberg, K. (2005). Social categorization and the perception of facial affect: Target race

    moderates the response latency advantage for happy faces. *Emotion, 5*(3)*,* 267-276.

    http://doi.org/10.1037/1528-3542.5.3.267

Hugenberg, K., Miller, J., & Claypool, H. M. (2007). Categorization and individuation in the

    cross-race recognition deficit: Toward a solution to an insidious problem. *Journal of*

    Experimental Social Psychology, *43*(2), 334-340.

    https://doi.org/10.1016/j.jesp.2006.02.010

Human, L. J., & Biesanz, J. C. (2013). Targeting the good target: An integrative review of the characteristics and consequences of being accurately perceived. *Personality and Social Psychology Review*, *17*(3), 248-272. https://doi.org/10.1177/1088868313495593

Hyers, L. L., & Swim, J. K. (1998). A comparison of the experiences of dominant and minority group members during an intergroup encounter. *Group Processes and Intergroup Relations*, *1*(2), 143-163. https://doi.org/10.1177/1368430298012003

Ickes, W., & Hodges, S. D. (2013). Empathic accuracy in close relationships. In J. Simpson & L. Campbell (Eds.), *The Oxford handbook of close relationships* (pp. 348-373). New York, NY: Oxford University Press.

Inzlicht, M., & Ben-Zeev, T. (2003). Do High-Achieving Female Students Underperform in Private? The Implications of Threatening Environments on Intellectual Processing. *Journal of Educational Psychology*, *95*(4), 796-805. https://doi.org/10.1037/0022-0663.95.4.796

Jack, R. E. (2013). Culture and facial expressions of emotion. *Visual Cognition*, *21*(9-10), 1248-1286. https://doi.org/10.1080/13506285.2013.835367

Jack, R. E., & Schyns, P. G. (2015). The human face as a dynamic tool for social communication. *Current Biology*, *25*(14), R621-R634. https://doi.org/10.1016/j.cub.2015.05.052

Jack, R. E., Sun, W., Delis, I., Garrod, O. G., & Schyns, P. G. (2016). Four not six: Revealing culturally common facial expressions of emotion. *Journal of Experimental Psychology: General*, *145*(6), 708-730. https://doi.org/10.1037/xge0000162

Jiang, X., Paulmann, S., Robin, J., & Pell, M. D. (2015). More than accuracy: Nonverbal dialects modulate the time course of vocal emotion recognition across cultures. *Journal of*

Experimental Psychology: Human Perception and Performance, 41(3), 597-612.

http://doi.org/10.1037/xhp0000043

Johns, M., Inzlicht, M., & Schmader, T. (2008). Stereotype threat and executive resource

depletion: Examining the influence of emotion regulation. *Journal of Experimental*

Psychology: General, *137*(4), 691-705. https://doi.org/10.1037/a0013834

Johnstone, T. & Scherer, K. R. (2000). Vocal communication of emotion. In M. Lewis & J. M.

Haviland-Jones (Eds.), *The handbook of emotions* (2nd ed., pp. 220-235). New York,

NY: Guilford Press.

Judd, C. M., Ryan, C. S., & Park, B. (1991). Accuracy in the judgment of in-group and out-group

variability. *Journal of Personality and Social Psychology*, *61*(3), 366-379.

https://doi.org/10.1037/0022-3514.61.3.366

Kang, S. M., & Lau, A. S. (2013). Revisiting the out-group advantage in emotion recognition in

a multicultural society: Further evidence for the ingroup advantage. *Emotion*, *13*(2), 203-

215. https://doi.org/10.1037/a0030013

Kelly, D. J., Jack, R. E., Miellet, S., De Luca, E., Foreman, K., & Caldara, R. (2011). Social

experience does not abolish cultural diversity in eye movements. *Frontiers in*

Psychology, *2*(1), 1-11. https://doi.org/10.3389/fpsyg.2011.00095

Kleinsmith, A., De Silva, P. R., & Bianchi-Berthouze, N. (2006). Cross-cultural differences in

recognizing affect from body posture. *Interacting with Computers*, *18*(6), 1371-1389.

https://doi.org/10.1016/j.intcom.2006.04.003

Krumhuber, E. G., Kappas, A., & Manstead, A. S. (2013). Effects of dynamic aspects of facial

expressions: A review. *Emotion Review*, *5*(1), 41-46.

https://doi.org/10.1177/1754073912451349

Kunstman, J. W., Tuscherer, T., Trawalter, S., & Lloyd, E. P. (2016). What lies beneath?

    Minority group members' suspicion of Whites' egalitarian motivation predicts responses

    to Whites' smiles. *Personality and Social Psychology Bulletin*, *42*(9), 1193-1205.

    https://doi.org/10.1177/0146167216652860

Landrine, H., Klonoff, E. A., Corral, I., Fernandez, S., & Roesch, S. (2006). Conceptualizing and

    measuring ethnic discrimination in health research. *Journal of Behavioral Medicine,*

    29(1)*,* 79-94. http://doi.org/10.1007/s10865-005-9029-0

Laurenceau, J., Barrett, L. F., & Peitromonaco, P. R. (1998). Intimacy as an interpersonal

    process: The importance of self-disclosure and partner disclosure, and perceived partner

    responsiveness in interpersonal exchanges. *Journal of Personality and Social*

    Psychology, 74(5)*,* 1238-1251. https://doi.org/10.1037/0022-3514.74.5.1238

Leary, M. R., Tambor, E. S., Terdal, S. K., & Downs, D. L. (1995). Self-esteem as an

    interpersonal monitor: The sociometer hypothesis. *Journal of Personality and Social*

    Psychology, 68(3)*,* 518-530. https://doi.org/10.1037/0022-3514.68.3.518

Letzring, T. D., & Human, L. J. (2014). An examination of information quality as a moderator of

    accurate personality judgment. *Journal of Personality*, *82*(5), 440-451.

    https://doi.org/10.1111/jopy.12075

Letzring, T. D., Wells, S. M., & Funder, D. C. (2006). Information quantity and quality affect the

    realistic accuracy of personality judgment. *Journal of Personality and Social Psychology*,

    91(1), 111-123. https://doi.org/10.1037/0022-3514.91.1.111

Lewis, M., Haviland-Jones, J. M., & Barrett, L. F. (Eds.). (2010). *Handbook of emotions*.

    Guilford Press.

Lin, M. H., Kwan, V. S., Cheung, A., & Fiske, S. T. (2005). Stereotype content model explains

  prejudice for an envied outgroup: Scale of anti-Asian American Stereotypes. *Personality*

  and Social Psychology Bulletin, 31(1)*,* 34-47. http://doi.org/10.1177/0146167204271320

Littleford, L. N., Wright, M. O., & Sayoc-Parial, M. S. (2005). White students' intergroup

  anxiety during same-race and interracial interactions: A multi-method approach. *Basic*

  and Applied Social Psychology, *27*(1), 85-94.

  https://doi.org/10.1207/s15324834basp2701_9

Ma-Kellams, C., & Blascovich, J. (2012). Inferring the emotions of friends versus strangers: The

  role of culture and self-construal. *Personality and Social Psychology Bulletin*, *38*(7), 933-

  945. https://doi.org/10.1177/0146167212440291

Malloy, T. E., Albright, L., Diaz-Loving, R., Dong, Q., & Lee, Y. T. (2004). Agreement in

  personality judgments within and between nonoverlapping social groups in collectivist

  cultures. *Personality and Social Psychology Bulletin*, *30*(1), 106-117.

  https://doi.org/10.1177/0146167203258863

Malpass, R. S., & Kravitz, J. (1969). Recognition for faces of own and other race. *Journal of*

  Personality and Social Psychology, 13(4), 330-334. https://doi.org/10.1037/h0028434

Malpass, R. S., Lavigueur, H., & Weldon, D. E. (1973). Verbal and visual training in face

  recognition. *Perception and Psychophysics, 14*(1), 283-292.

  https://doi.org/10.3758/BF03212392

Marsh, A., Elfenbein, H. A., & Ambady, N. (2003). Nonverbal "accents": Cultural differences in

  facial expressions of emotion. *Psychological Science, 14*(4), 373-376.

  https://doi.org/10.1111/1467-9280.24461

Matsumoto, D. (1990). Cultural similarities and differences in display rules. *Motivation and Emotion*, *14*(3), 195-214. https://doi.org/10.1007/BF00995569

Matsumoto, D., & Hwang, H. C. (2013). Cultural similarities and differences in emblematic gestures. *Journal of Nonverbal Behavior*, *37*(1), 1-27. https://doi.org/10.1007/s10919-012-0143-8

Matsumoto, D., Olide, A., & Willingham, B. (2009). Is there an ingroup advantage in recognizing spontaneously expressed emotions? *Journal of Nonverbal Behavior*, *33*(3), 181-191. https://doi.org/10.1007/s10919-009-0068-z

Matsumoto, D., Yoo, S. H., Hirayama, S., & Petrova, G. (2005). Development and validation of a measure of display rule knowledge: The display rule assessment inventory. *Emotion, 5*(1), 23-40. https://doi.org/10.1037/1528-3542.5.1.23

Mauss, I. B., Levenson, R. W., McCarter, L., Wilhelm, F. H., & Gross, J. J. (2005). The tie that binds? Coherence among emotion experience, behavior, and physiology. *Emotion, 5*(2), 175-190. https://doi.org/10.1037/1528-3542.5.2.175

McDuff, D., Girard, J. M., & El Kaliouby, R. (2017). A large-scale observational evidence of cross-cultural differences in facial behavior. *Journal of Nonverbal Behavior*, *41*(1), 1-19. https://doi.org/10.1371/journal.pone.0173942

McLin, O. H., & Malpass, R. S. (2001). Racial categorization of faces: The ambiguous race face effect. *Psychology, Public Policy, and Law*, 7(1), 98-118. https://doi.org/10.1037/1076-8971.7.1.98

Meeren, H. K., van Heijnsbergen, C. C., & de Gelder, B. (2005). Rapid perceptual integration of facial expression and emotional body language. *Proceedings of the National Academy of*

Sciences of the United States of America, *102*(45), 16518-16523.

https://doi.org/10.1073/pnas.0507650102

Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in

memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law, 7*(1), 3-

35. https://doi.org/10.1037/1076-8971.7.1.3

Mendes, W. B., Blascovich, J., Lickel, B., & Hunter, S. (2002). Challenge and threat during

social interactions with White and Black men. *Personality and Social Psychology*

Bulletin, *28*(7), 939-952. https://doi.org/10.1177/014616720202800707

Mortillaro, M., Mehu, M., & Scherer, K. R. (2011). Subtly different positive emotions can be

distinguished by their facial expressions. *Social Psychological and Personality Science*,

2(3), 262-271. https://doi.org/10.1177/1948550610389080

Ng, W., & Lindsay, R. C. L. (1994). Cross-race facial recognition: Failure of the contact

hypothesis. *Journal of Cross-Cultural Psychology, 25*(2), 217-232.

https://doi.org/10.1177/0022022194252004

Nusseck, M., Cunningham, D. W., Wallraven, C., & Bülthoff, H. H. (2008). The contribution of

different facial regions to the recognition of conversational expressions. *Journal of*

Vision, *8*(8), 1-1. https://doi.org/10.1007/s10919-006-0020-4

O'Brien, L. T., & Major, B. (2005). System-justifying beliefs and psychological well-being: The

roles of group status and identity. *Personality and Social Psychology Bulletin, 31*(12)*,*

1718-1729. http://doi.org/10.1177/0146167205278261

Omi, M., & Winant, H. (2014). *Racial formation in the United States*. New York, NY:

Routledge.

Overall, N. C., Fletcher, G. J., & Kenny, D. A. (2012). When bias and insecurity promote accuracy: Mean-level bias and tracking accuracy in couples' conflict discussions. *Personality & Social Psychology Bulletin*, *38*(5), 642-655. https://doi.org/10.1177/0146167211432764

Page-Gould, E., Mendoza-Denton, R., & Tropp, L. R. (2008). With a little help from my cross-group friend: Reducing anxiety in intergroup contexts through cross-group friendship. *Journal of Personality and Social Psychology*, *95*(5), 1080-1094. https://doi.org/10.1037/0022-3514.95.5.1080

Paolini, S., Hewstone, M., Voci, A., Harwood, J., & Cairns, E. (2006). Intergroup contact and the promotion of intergroup harmony: The influence of intergroup emotions. In R. Brown & D. Capozza (Eds.), *Social identities: Motivational* (pp. 209-238). Hove, UK: Psychology Press/Taylor & Francis.

Pettigrew, T. F. (1998). Intergroup contact theory. *Annual Review of Psychology*, *49*(1), 65-85. https://doi.org/10.1146/annurev.psych.49.1.65

Pettigrew, T. F., & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology*, *90*(5), 751-783. https://doi.org/10.1037/0022-3514.90.5.751

Pettigrew, T. F., & Tropp, L. R. (2008). How does intergroup contact reduce prejudice? Meta-analytic tests of three mediators. *European Journal of Social Psychology*, *38*(6), 922-934. https://doi.org/10.1002/ejsp.504

Posner, J., Russell, J. A., & Peterson, B. S. (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and Psychopathology*, *17*(3), 715-734.

https://doi.org/10.1017/S0954579405050340

Ready, R. E., Clark, L. A., Watson, D., & Westerhouse, K. (2000). Self- and peer-reported
personality: Agreement, trait ratability, and the "self-based heuristic." *Journal of
Research in Personality, 34*(2)*,* 208-224. https://doi.org/10.1006/jrpe.1999.2280

Reis, H. T., & Shaver, P. (1988). Intimacy as an interpersonal process. In S. W. Duck (Ed.),
Handbook of personal relationships (pp. 367–389). Oxford, UK: John Wiley & Sons Ltd.

Reysen, S. (2005). Construction of a new scale: The Reysen Likability Scale. *Social Behavior
and Personality, 33*(2), 201-208. https://doi.org/10.2224/sbp.2005.33.2.201

Rhodes, G., Ewing, L., Hayward, W. G., Maurer, D., Mondloch, C. J., & Tanaka, J. W. (2009).
Contact and other-race effects in configural and component processing of faces. *British
Journal of Psychology*, *100*(4), 717-728. https://doi.org/10.1348/000712608X396503

Richeson, J. A., & Shelton, J. N. (2007). Negotiating interracial interactions: Costs,
consequences, and possibilities. *Current Directions in Psychological Science*, *16*(6), 316-
320. https://doi.org/10.1111/j.1467-8721.2007.00528.x

Sangrigoli, S., Pallier, C., Argenti, A.-M., Ventureyra, V. A. G., & de Schonen, S. (2005).
Reversibility of the other-race effect in face recognition during childhood. *Psychological
Science, 16*(6), 440-444. https://doi.org/10.1111/j.0956-7976.2005.01554.x

Scherer, K. R., Clark-Polner, E., & Mortillaro, M. (2011). In the eye of the beholder?
Universality and cultural specificity in the expression and perception of emotion.
International Journal of Psychology, *46*(6), 401-435.
https://doi.org/10.1080/00207594.2011.626049

Schmader, T., Johns, M., & Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychological Review*, *115*(2), 336-356. https://doi.org/10.1037/0033-295X.115.2.336

Schofield, T. J., Parke, R. D., Castañeda, E. K., & Coltrane, S. (2008). Patterns of gaze between parents and children in European American and Mexican American families. *Journal of Nonverbal Behavior*, *32*(3), 171-186. https://doi.org/10.1007/s10919-008-0049-7

Shelton, J. N. (2003). Interpersonal concerns in social encounters between majority and minority group members. *Group Processes and Intergroup Relations*, *6*(2), 171-185. https://doi.org/10.1177/1368430203006002003

Shelton, J. N., Dovidio, J. F., Hebl, M., & Richeson, J. A. (2009). Prejudice and intergroup interaction. In S. Demoulin, J. P. Leyens, & J. F. Dovidio (Eds.), *Intergroup misunderstandings: Impact of divergent social realities* (pp. 21-38). New York, NY: Psychology Press.

Shelton, J. N., Richeson, J. A., & Salvatore, J. (2005). Expecting to be the target of prejudice: Implications for interethnic interactions. *Personality and Social Psychology Bulletin*, *31*(9), 1189-1202. https://doi.org/10.1177/0146167205274894

Shelton, J. N., Richeson, J. A., Salvatore, J., & Trawalter, S. (2005). Ironic effects of racial bias during interracial interactions. *Psychological Science*, *16*(5), 397-402. https://doi.org/10.1111/j.0956-7976.2005.01547.x

Shelton, J. N., Richeson, J. A., & Vorauer, J. D. (2006). Threatened identities and interethnic interactions. In W. Stroebe & M. Hewstone (Eds.), *European review of social psychology* (pp. 321-358). New York, NY: Psychology Press.

Sheng, F., & Han, S. (2012). Manipulations of cognitive strategies and intergroup relationships

    reduce the racial bias in empathic neural responses. *NeuroImage*, *61*(4), 786-797.

    https://doi.org/10.1016/j.neuroimage.2012.04.028

Shih, M., Ambady, N., Richeson, J. A., Fujita, K., & Gray, H. M. (2002). Stereotype

    performance boosts: The impact of self-relevance and the manner of stereotype

    activation. *Journal of Personality and Social Psychology*, *83*(3), 638-647.

    https://doi.org/10.1037//0022-3514.83.3.638

Siy, J. O., & Cheryan, S. (2013). When compliments fail to flatter: American individualism and

    responses to positive stereotypes. *Journal of Personality and Social Psychology, 104*(1)*,*

    87-102. https://doi.org/10.1037/a0030183

Sorokowska, A., Sorokowski, P., Hilpert, P., Cantarero, K., Frackowiak, T., Ahmadi, K., ... &

    Blumen, S. (2017). Preferred Interpersonal Distances: A Global Comparison. *Journal of*

    Cross-Cultural Psychology, *48*(4), 577-592. https://doi.org/10.1177/0022022117698039

Soto, J. A., & Levenson, R. W. (2009). Emotion recognition across cultures: the influence of

    ethnicity on empathic accuracy and physiological linkage. *Emotion*, *9*(6), 874-884.

    https://doi.org/10.1037/a0017399

Sporer, S. L. (2001). Recognizing faces of other ethnic groups: An integration of theories.

    Psychology, Public Policy, and Law, *7*(1), 36-97. https://doi.org/10.1037/1076-

    8971.7.1.36

Stanley, J. T., Zhang, X., Fung, H. H., & Isaacowitz, D. M. (2013). Cultural differences in gaze

    and emotion recognition: Americans contrast more than Chinese. *Emotion*, *13*(1), 36-46.

    https://doi.org/10.1037/a0029209

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of

    African Americans. *Journal of Personality and Social Psychology, 69*(5), 797-811.

    https://doi.org/10.1037/0022-3514.69.5.797

Stephan, W. G., & Stephan, C. W. (1989). Antecedents of intergroup anxiety in Asian-

    Americans and Hispanic-Americans. *International Journal of Intercultural Relations*,

    13(1), 203-219. https://doi.org/10.1016/0147-1767(89)90006-0

Stevenson, M. T., Soto, J. A., & Adams, R. B. (2012). More than meets the eye: The role of self-

    identity in decoding complex emotional states. *Emotion*, *12*(5), 882-886.

    https://doi.org/10.1037/a0028628

Sue, D. W., Bucceri, J., Lin, A. I., Nadal, K. L., & Torino, G. C. (2007). Racial microaggressions

    and the Asian American experience. *Cultural Diversity and Ethnic Minority Psychology*,

    13(1), 72-81. https://doi.org/10.1037/1099-9809.13.1.72

Tajfel, H., & Turner, J. C. (2004). An integrative theory of intergroup conflict. In M. J. Hatch &

    M. Schultz (Eds.), *Organizational identity: A reader* (pp. 56-65). New York, NY: Oxford

    University Press.

Thompson, W. F., & Balkwill, L. L. (2006). Decoding speech prosody in five languages.

    Semiotica, *2006*(158), 407-424. https://doi.org/10.1515/SEM.2006.017

Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2015). Social attributions from

    faces: Determinants, consequences, accuracy, and functional significance. *Annual Review*

    of Psychology, *66*(1), 519-545. https://doi.org/10.1146/annurev-psych-113011-143831

Tracy, J. L., & Robins, R. W. (2004). Putting the self into self-conscious emotions: A theoretical

    model. *Psychological Inquiry*, *15*(2), 103-125.

    https://doi.org/10.1207/s15327965pli1502_01

Trawalter, S., Richeson, J. A., & Shelton, J. N. (2009). Predicting behavior during interracial interactions: A stress and coping approach. *Personality and Social Psychology Review*, *13*(4), 243-268. https://doi.org/10.1177/1088868309345850

Tsai, J. L. (2007). Ideal affect: Cultural causes and behavioral consequences. *Perspectives on Psychological Science*, *2*(3), 242-259. https://doi.org/10.1111/j.1745-6916.2007.00043.x

Tversky, A. and Kahneman, D. (1974). Judgment under uncertainty: heuristics and biases. *Science*, *185*(4157), 1124-1131. https://doi.org/10.1126/science.185.4157.1124

U.S. Census Bureau. (2019). *National Population by Characteristics: 2010-2019*. Retrieved from https://www.census.gov/data/tables/time-series/demo/popest/2010s-national-detail.html

Van den Stock, J., Righart, R., & De Gelder, B. (2007). Body expressions influence recognition of emotions in the face and voice. *Emotion*, *7*(3), 487-494. https://doi.org/10.1037/1528-3542.7.3.487

Vazire, S. (2010). Who knows what about a person? The self–other knowledge asymmetry (SOKA) model. *Journal of Personality & Social Psychology*, *98*(2), 281-300. https://doi.org/10.1037/a0017908

Vinciarelli, A., Pantic, M., & Bourlard, H. (2008). Social signal processing: Survey of an emerging domain. *Image and Vision Computing*, *27*(12), 1743-1759. https://doi.org/10.1016/j.imavis.2008.11.007

Vorauer, J. D. (2006). An information search model of evaluative concerns in intergroup interaction. *Psychological Review, 113*(4)*,* 862-886. https://doi.org/10.1037/0033-295X.113.4.862

Vorauer, J. D., Hunter, A. J., Main, K. J., & Roy, S. (2000). Meta-stereotype activation: Evidence from indirect measures for specific evaluative concerns experienced by members of dominant groups in intergroup interaction. *Journal of Personality and Social Psychology*, *78*(4), 690-707. https://doi.org/10.1037/0022-3514.78.4.690

Vorauer, J. D., Main, K. J., & O'Connell, G. B. (1998). How do individuals expect to be viewed by members of lower status groups? Content and implications of meta-stereotypes. *Journal of Personality and Social Psychology*, *75*(4), 917-937. https://doi.org/10.1037/0022-3514.75.4.917

Vorauer, J. D., Martens, V., & Sasaki, S. J. (2009). When trying to understand detracts from trying to behave: Effects of perspective taking in intergroup interaction. *Journal of Personality and Social Psychology*, *96*(4), 811-827. https://doi.org/10.1037/a0013411

Walker, P. M., & Hewstone, M. (2006). A developmental investigation of other-race contact and the own-race face effect. *British Journal of Developmental Psychology*, *24*(3), 451-463. https://doi.org/10.1348/026151005X51239

Walker, P. M., Silvert, L., Hewstone, M., & Nobre, A. C. (2008). Social contact and other-race face processing in the human brain. *Social Cognitive and Affective Neuroscience*, *3*(1), 16-25. https://doi.org/10.1093/scan/nsm035

Watson, R., & de Gelder, B. (2017). How white and black bodies are perceived depends on what emotion is expressed. *Scientific Reports, 7*(1), 1-12. https://doi.org/10.1038/srep41349

Watson, D., Hubbard, B., & Wiese, D. (2000a). General traits of personality and affectivity as predictors of satisfaction in intimate relationships: Evidence from self-and partner-ratings. *Journal of Personality*, *68*(3), 413-449. https://doi.org/10.1111/1467-6494.00102

Watson, D., Hubbard, B., & Wiese, D. (2000b). Self-other agreement in personality and

    affectivity: The role of acquaintanceship, trait visibility, and assumed similarity. *Journal*

    of Personality and Social Psychology, 78(3)*,* 546–558. https://doi.org/10.1037//0022-

    3514.78.3.546

West, T. V., Dovidio, J. F., & Pearson, A. R. (2014). Accuracy and bias in perceptions of

    relationship interest for intergroup and intragroup roommates. *Social Psychological and*

    Personality Science, *5*(2), 235-242. https://doi.org/10.1177/1948550613490966

West, T. V., & Kenny, D. A. (2011). The truth and bias model of judgment. *Psychological*

    Review, *118*(2), 357-378. https://doi.org/10.1037/a0022936

West, T. V., Koslov, K., Page-Gould, E., Major, B., & Mendes, W. B. (2017). Contagious

    anxiety: anxious European Americans can transmit their physiological reactivity to

    African Americans. *Psychological Science*, *28*(12), 1796-1806.

    https://doi.org/10.1177/0956797617722551

West, T. V., Magee, J. C., Gordon, S. H., & Gullett, L. (2014). A little similarity goes a long

    way: The effects of peripheral but self-revealing similarities on improving and sustaining

    interracial relationships. *Journal of Personality and Social Psychology, 107*(1), 81-100.

    https://doi.org/10.1037/a0036556

Wickline, V. B., Bailey, W., & Nowicki, S. (2009). Cultural in-group advantage: Emotion

    recognition in African American and European American faces and voices. *The Journal*

    of Genetic Psychology, *170*(1), 5-30. https://doi.org/10.3200/GNTP.170.1.5-30

Wieser, M. J., & Brosch, T. (2012). Faces in context: a review and systematization of contextual

    influences on affective face processing. *Frontiers in Psychology*, *3*(1), 1-13.

    https://doi.org/10.3389/fpsyg.2012.00471

Wojciszke, B. (2005). Morality and competence in person- and self-perception. In W. Stroebe

    (Ed.), *European review of social psychology* (Vol. 16, pp. 155-188). Hove, UK:

    Psychology Press.

Yan, X., Andrews, T. J., Jenkins, R., & Young, A. W. (2016). Cross-cultural differences and

    similarities underlying other-race effects for facial identity and expression. *The Quarterly*

    Journal of Experimental Psychology, *69*(7), 1247-1254.

    https://doi.org/10.1080/17470218.2016.1146312

Yankouskaya, A., Humphreys, G. W., & Rotshtein, P. (2014). Differential interactions between

    identity and emotional expression in own and other-race faces: Effects of familiarity

    revealed through redundancy gains. *Journal of Experimental Psychology: Learning,*

    Memory, and Cognition, *40*(4), 1025-1038. https://doi.org/10.1037/a0036259

Young, S. G., & Hugenberg, K. (2010). Mere social categorization modulates identification of

    facial expressions of emotion. *Journal of Personality and Social Psychology*, *99*(6), 964-

    977. https://doi.org/10.1037/a0020400

Young, S. G., Hugenberg, K., Bernstein, M. J., Sacco, D. F. (2012). Perception and motivation in

    face recognition: A critical review of theories of the cross-race effect. *Personality and*

    Social Psychology Review, 16(2), 116-142. https://doi.org/10.1177/1088868311418987

Yuki, M., Maddux, W.W. & Masuda, T. (2007). Are the windows to the soul the same in the

    East and West? Cultural differences in using the eyes and mouth as cues to recognize

    emotions in Japan and the United States. *Journal of Experimental Social Psychology*,

    43(2), 303-311. https://doi.org/10.1016/j.jesp.2006.02.004

Yzerbyt, V. Y., Kervyn, N., & Judd, C. M. (2008). Compensation versus halo: The unique

    relations between the fundamental dimensions of social judgment. *Personality and Social*

    Psychology Bulletin, 34(8), 1110-1123. https://doi.org/10.1177/0146167208318602

Zhang, F., Parmley, M., Wan, X., & Cavanagh, S. (2015). Cultural differences in recognition of

    subdued facial expressions of emotions. *Motivation and Emotion*, *39*(2), 309-319.

    https://doi.org/10.1007/s11031-014-9454-x

Zhu, Y. (2013). Which is the best listener group? Perception of Chinese emotional prosody by

    Chinese natives, naïve Dutch listeners and Dutch L2 learners of Chinese. *Dutch Journal*

    of Applied Linguistics, 2(2), 170-183. http://doi.org/10.1075/dujal.2.2.03zhu

Zhu, Z., Ho, S. M., & Bonanno, G. A. (2013). Cultural similarities and differences in the

    perception of emotional valence and intensity: a comparison of Americans and Hong

    Kong Chinese. *The American Journal of Psychology*, *126*(3), 261-273.

    https://doi.org/10.5406/amerjpsyc.126.3.0261

Zou, L. X., & Cheryan, S. (2017). Two axes of subordination: A new model of racial position.

    Journal of Personality and Social Psychology, *112*(5), 696-717.

    https://doi.org/10.1037/pspa0000080

Table 1

Inter-item correlation matrix for key Study 1 and 2 variables.

	Mean (SD)	1	2	3	4	5	6	7
		5.73 (1.58)	6.08 (1.64)	3.19 (1.47)	3.16 (1.47)	--	.15 (.16)	5.18 (.94)
1. Judge PE rating	4.10 (1.00)	--	.79 **	-.17 **	-.10 †	-.05	.03	.35 **
2. Target PE rating	4.39 (.99)	.22 **	--	-.09	-.12 *	-.14 *	.04	.38 **
3. Judge NE rating	1.96 (.68)	-.01	-.07	--	.68 **	-.08	-.05	-.17 **
4. Target NE rating	1.80 (.65)	-.13 *	-.11 †	.12 *	--	-.05	-.04	-.15 **
5. Judges' target expressivity attention	3.95 (1.33)	.19 **	.01	.09 †	-.02	--	-.02	-.07
6. Judges' outgroup exposure	.23 (.25)	-.15 **	-.06	.00	.07	.00	--	-.02
7. Judges' liking of targets	5.22 (.75)	.37 **	.19 **	-.24 **	-.08	-.03	.03	--

Note. Means and standard deviations are collapsed across dyads. Study 1 variables are presented below the diagonal, while Study 2 items are above. In Study 2, Column 5 represents judge condition (0 = control, 1 = experimental); dyads who failed the manipulation check have been excluded. "PE" and "NE" refer to positive and negative emotion respectively. Judge ratings represent their perceptions of targets' emotions, while target ratings refer to their self-reported experience. In Study 1, emotional experience was rated from 1 (*not at all*) to 7 (*a great deal*), while the Likert scale used in Study 2 ranged from 1-10. Expressivity attention was rated from 1 (*not at all*) to 7 (*a great deal*). Outgroup exposure refers to the ratio of outgroup social partners relative to the total size of judges' social networks. Liking was rated from 1 (*disagree strongly*) to 7 (*agree strongly*). All correlated variables have been z-scored. † $p < .10$, * $p < .05$, ** $p < .01$.

Table 2

Study 1 Actor-Partner Interdependence analyses for Models 1 and 2.

	Emotion Judgment Accuracy	
Model 1	B (SE)	R_β^2
Intercept	.58 (.08) **	.25
P1 expressivity attention	.01 (.01)	.01
P2 expressivity attention	.00 (.01)	.00
Model 2	B (SE)	R_β^2
Intercept	.63 (.10) **	.21
P1 expressivity attention	.01 (.01)	.00
P2 expressivity attention	.00 (.01)	.00
Dyad type	-.12 (.17)	.00
Dyad type * P1 expressivity attention	.02 (.02)	.00
Dyad type * P2 expressivity attention	-.01 (.02)	.00

Note. Fixed effects are presented with standard errors in parentheses and semi-partial R^2 (R_β^2) as an index of effect size. Profile correlations were used as the measure of emotion judgment accuracy. The Partner 1 (P1) predictor indicates judges' attention to target expressivity, while the Partner 2 (P2) one represents targets' attention to the expressivity of these judges. Dyad type was coded as 0 = same race and 1 = mixed race. † $p < .10$, * $p < .05$, ** $p < .01$.

Table 3

Study 1 Actor-Partner Interdependence analyses for Models 3 and 4.

	Emotion Judgment Accuracy	
Model 3	B (SE)	R_β^2
Intercept	.66 (.03) **	.76
P1 outgroup exposure	-.12 (.06) †	.01
P2 outgroup exposure	-.03 (.06)	.00
Model 4	B (SE)	R_β^2
Intercept	.69 (.04) **	.71
P1 outgroup exposure	-.20 (.07) **	.03
P2 outgroup exposure	.00 (.07)	.00
Dyad type	-.11 (.06) †	.02
Dyad type * P1 outgroup exposure	.23 (.12) †	.01
Dyad type * P2 outgroup exposure	-.02 (.12)	.00

Note. Fixed effects are presented with standard errors in parentheses and semi-partial R^2 (R_β^2) as an index of effect size. Profile correlations were used as the measure of emotion judgment accuracy. Outgroup exposure refers to the ratio of outgroup social partners relative to the total size of participants' social networks, with Partner 1 (P1) and Partner 2 (P2) referring to judge and target respectively. Dyad type was coded as 0 = same race and 1 = mixed race. † $p < .10$, * $p < .05$, ** $p < .01$.

Table 4

Study 1 Actor-Partner Interdependence analyses for Models 5 and 6.

	Partner Liking	
Model 5	B (SE)	R_β^2
Intercept	4.61 (.12) **	.90
P1's accuracy in judging P2	.75 (.17) **	.07
P2's accuracy in judging P1	.24 (.17)	.01
Model 6	B (SE)	R_β^2
Intercept	4.55 (.16) **	.84
P1's accuracy in judging P2	.83 (.22) **	.05
P2's accuracy in judging P1	.26 (.22)	.01
Dyad type	.14 (.24)	.00
Dyad type * P1's accuracy in judging P2	-.21 (.35)	.00
Dyad type * P2's accuracy in judging P1	-.05 (.35)	.00

Note. Fixed effects are presented with standard errors in parentheses and semi-partial R^2 (R_β^2) as an index of effect size. Profile correlations were used as the measure of emotion judgment accuracy. The Partner 1 (P1) predictor indicates the accuracy of judges' assessments of their partners, while the Partner 2 (P2) one represents partners' assessments of these judges. Dyad type was coded as 0 = same race and 1 = mixed race. Partner liking refers to how much judges liked targets. † $p < .10$, * $p < .05$, ** $p < .01$.

Table 5

Study 1 Truth and Bias analyses for Models 1, 2, and 3.

	Positive Emotion		Negative Emotion	
Model 1	B (SE)	R_β^2	B (SE)	R_β^2
Accuracy	.14 (.04) **	.05	.06 (.04)	.01
Directional bias	-.29 (.04) **	.29	.16 (.03) **	.18
Assumed similarity	.76 (.04) **	.58	.74 (.04) **	.51
Model 2	B (SE)	R_β^2	B (SE)	R_β^2
Accuracy	.15 (.04) **	.04	.08 (.05)	.01
Directional bias	-.26 (.04) **	.18	.14 (.03) **	.10
Assumed similarity	.77 (.04) **	.49	.72 (.05) **	.40
Dyad type *Accuracy	-.05 (.08)	.00	-.04 (.09)	.00
Dyad type * Directional bias	-.10 (.08)	.01	.05 (.06)	.01
Dyad type * Assumed similarity	-.05 (.08)	.00	.08 (.09)	.00
Model 3	B (SE)	R_β^2	B (SE)	R_β^2
Accuracy	.10 (.07)	.02	.06 (.08)	.01
Directional bias	-.37 (.07) **	.40	.19 (.05) **	.21
Assumed similarity	.71 (.07) **	.53	.80 (.08) **	.52
Partner expressivity attention *Accuracy	.04 (.06)	.01	.13 (.04) †	.03
Partner expressivity attention * Directional bias	.05 (.05)	.00	.00 (.04)	.00
Partner expressivity attention * Assumed similarity	.03 (.05)	.01	.04 (.07)	.00

Note. Fixed effects are presented with standard errors in parentheses and semi-partial R^2 (R_β^2) as an index of effect size. Dyad type was coded as 0 = same-race and 1 = mixed-race. Partner expressivity attention refers to how much judges monitored targets' expressivity during the interaction. Positive and negative emotion refer to judgments of targets' emotional experience during the interaction. Continuous predictors have been z-scored. † $p < .10$, * $p < .05$, ** $p < .01$.

Table 6

Study 1 Truth and Bias analyses for Models 4 and 5.

	Positive Emotion		Negative Emotion	
Model 4	B (SE)	R_β^2	B (SE)	R_β^2
Accuracy	.12 (.07) †	.03	.06 (.05)	.01
Directional bias	-.47 (.06) **	.55	.20 (.05) **	.25
Assumed similarity	.73 (.07) **	.55	.50 (.05) **	.53
Outgroup exposure *Accuracy	-.08 (.07)	.02	.12 (.06) *	.05
Outgroup exposure * Directional bias	-.02 (.07)	.00	-.07 (.05)	.02
Outgroup exposure * Assumed similarity	.12 (.07) †	.03	.03 (.05)	.00
Model 5	B (SE)	R_β^2	B (SE)	R_β^2
Accuracy	.12 (.04) **	.04	.06 (.04)	.01
Directional bias	-.32 (.04) **	.30	.16 (.03) **	.18
Assumed similarity	.72 (.04) **	.55	.73 (.04) **	.49
Partner liking *Accuracy	.07 (.05)	.01	.04 (.06)	.00
Partner liking * Directional bias	.18 (.05) **	.04	-.11 (.04) **	.03
Partner liking * Assumed similarity	.08 (.05)	.01	.02 (.05)	.00

Note. Fixed effects are presented with standard errors in parentheses and semi-partial R^2 (R_β^2) as an index of effect size. Outgroup exposure refers to the ratio of outgroup social partners relative to the total size of judges' social networks. Partner liking refers to how much judges liked targets. Positive and negative emotion refer to judgments of targets' emotional experience during the interaction. Continuous predictors have been z-scored. † $p < .10$, * $p < .05$, ** $p < .01$.

Table 7

Study 2 Actor-Partner Interdependence analyses for Models 1 and 2.

	Emotion Judgment Accuracy	
Model 1	B (SE)	R_β^2
Intercept	.52 (.03) **	.60
P1 condition	-.05 (.04)	.01
P2 condition	-.07 (.05)	.01
Model 2	B (SE)	R_β^2
Intercept	.51 (.04) **	.46
P1 condition	-.07 (.06)	.01
P2 condition	-.02 (.06)	.00
Judge race	.01 (.03)	.00
Judge race * P1 condition	.04 (.07)	.00
Judge race * P2 condition	-.11 (.07)	.01

Note. Fixed effects are presented with standard errors in parentheses and semi-partial R^2 (R_β^2) as an index of effect size. Dyads who failed the experimental manipulation check have been excluded. Profile correlations were used as the measure of emotion judgment accuracy. The Partner 1 (P1) predictor indicates judges' condition assignment, while the Partner 2 (P2) one represents targets' condition assignment (both coded as 0 = control and 1 = experimental). Race was coded as 0 = White and 1 = Asian. † $p < .10$, * $p < .05$, ** $p < .01$.

Table 8

Study 2 Actor-Partner Interdependence analyses for Models 3, 4, and 5.

	Emotion Judgment Accuracy	
Model 3	B (SE)	R_β^2
Intercept	.53 (.04) **	.57
P1 outgroup exposure	-.16 (.11)	.01
P2 outgroup exposure	-.12 (.11)	.01
Model 4	B (SE)	R_β^2
Intercept	.52 (.04) **	.42
P1 outgroup exposure	-.08 (.22)	.00
P2 outgroup exposure	-.11 (.13)	.00
Judge race	.02 (.04)	.00
Judge race * P1 outgroup exposure	-.12 (.25)	.00
Judge race * P2 outgroup exposure	-.05 (.26)	.00
Model 5	B (SE)	R_β^2
Intercept	.52 (.04) **	.46
P1 outgroup exposure	-.09 (.14)	.00
P2 outgroup exposure	-.16 (.13)	.01
Judge condition	.01 (.04)	.00
Judge condition * P1 outgroup exposure	-.18 (.18)	.00
Judge condition * P2 outgroup exposure	.14 (.19)	.00

Note. Fixed effects are presented with standard errors in parentheses and semi-partial R^2 (R_β^2) as an index of effect size. Dyads who failed the experimental manipulation check have been excluded. Profile correlations were used as the measure of emotion judgment accuracy. Outgroup exposure refers to the ratio of outgroup social partners relative to the total size of participants' social networks, with Partner 1 (P1) and Partner 2 (P2) referring to judge and target respectively. Race was coded as 0 = White and 1 = Asian. Condition was coded as 0 = control and 1 = experimental. † $p < .10$, * $p < .05$, ** $p < .01$.

Table 9

Study 2 Actor-Partner Interdependence analyses for Models 6, 7, and 8.

	Partner Liking	
Model 6	B (SE)	R_β^2
Intercept	4.45 (.11) **	.91
P1 profile correlations	.99 (.19) **	.10
P2 profile correlations	.53 (.19) **	.03
Model 7	B (SE)	R_β^2
Intercept	4.26 (.14) **	.76
P1 profile correlations	1.43 (.31) **	.07
P2 profile correlations	.48 (.27) †	.01
Judge race	.35 (.18) †	.02
Judge race * P1 profile correlations	-.72 (.44) †	.01
Judge race * P2 profile correlations	.00 (.44)	.00
Model 8	B (SE)	R_β^2
Intercept	4.49 (.13) **	.86
P1 profile correlations	1.26 (.23) **	.12
P2 profile correlations	.23 (.25)	.00
Judge condition	-.03 (.20)	.00
Judge condition * P1 profile correlations	-.91 (.46) *	.02
Judge condition * P2 profile correlations	.81 (.42) †	.02

Note. Fixed effects are presented with standard errors in parentheses and semi-partial R^2 (R_β^2) as an index of effect size. Dyads who failed the experimental manipulation check have been excluded. Profile correlations were used as the measure of emotion judgment accuracy. The Partner 1 (P1) predictor indicates the accuracy of judges' assessments of their partners, while the Partner 2 (P2) one represents partners' assessments of these judges. Race was coded as 0 = White and 1 = Asian. Condition was coded as 0 = control and 1 = experimental. Partner liking refers to how much judges liked targets. † $p < .10$, * $p < .05$, ** $p < .01$.

Table 10

Study 2 Truth and Bias analyses for Models 1, 2, and 3.

	Positive Emotion		Negative Emotion	
Model 1	B (SE)	R_β^2	B (SE)	R_β^2
Accuracy	.72 (.03) **	.64	.65 (.04) **	.45
Directional bias	-.35 (.05) **	.27	.02 (.06)	.00
Assumed similarity	.21 (.03) **	.13	.13 (.04) **	.03
Model 2	B (SE)	R_β^2	B (SE)	R_β^2
Accuracy	.66 (.05) **	.39	.64 (.06) **	.28
Directional bias	-.26 (.07) **	.04	-.25 (.08) **	.03
Assumed similarity	.21 (.04) **	.07	.03 (.06)	.00
Judge race *Accuracy	.12 (.06) †	.01	.07 (.08)	.00
Judge race * Directional bias	-.18 (.11)	.02	.52 (.12) **	.11
Judge race * Assumed similarity	.01 (.06)	.00	.15 (.08) †	.01
Model 3	B (SE)	R_β^2	B (SE)	R_β^2
Accuracy	.74 (.04) **	.59	.64 (.04) **	.44
Directional bias	-.41 (.06) **	.19	.07 (.07)	.00
Assumed similarity	.21 (.04) **	.10	.13 (.04) **	.03
Judge condition *Accuracy	-.02 (.07)	.00	.03 (.07)	.00
Judge condition * Directional bias	.19 (.12) †	.01	-.14 (.13)	.01
Judge condition * Assumed similarity	-.03 (.07)	.00	-.05 (.06)	.00

Note. Fixed effects are presented with standard errors in parentheses and semi-partial R^2 (R_β^2) as an index of effect size. Dyads who failed the experimental manipulation check have been excluded. Race was coded as 0 = White and 1 = Asian. Condition was coded as 0 = control and 1 = experimental. Positive and negative emotion refer to judgments of targets' emotional experience during the interaction. † $p < .10$, * $p < .05$, ** $p < .01$.

Table 11

Study 2 Truth and Bias analyses for Models 4 and 5.

	Positive Emotion		Negative Emotion	
Model 4	B (SE)	R_β^2	B (SE)	R_β^2
Accuracy	1.18 (.05) **	.65	.94 (.06) **	.44
Directional bias	-.36 (.05) **	.28	.02 (.06)	.00
Assumed similarity	.32 (.06) **	.11	.19 (.06) **	.03
Outgroup exposure *Accuracy	.11 (.05) *	.02	.02 (.05)	.00
Outgroup exposure * Directional bias	-.03 (.05)	.00	-.05 (.06)	.00
Outgroup exposure * Assumed similarity	-.12 (.06) *	.01	.04 (.06)	.00
Model 5	B (SE)	R_β^2	B (SE)	R_β^2
Accuracy	1.18 (.06) **	.61	.93 (.06) *	.44
Directional bias	-.38 (.05) **	.25	.02 (.06)	.00
Assumed similarity	.34 (.06) **	.11	.17 (.06) **	.03
Partner liking *Accuracy	-.03 (.06)	.00	.01 (.06)	.00
Partner liking * Directional bias	.02 (.06)	.00	-.12 (.06) *	.02
Partner liking * Assumed similarity	.10 (.05) †	.01	-.03 (.06)	.00

Note. Fixed effects are presented with standard errors in parentheses and semi-partial R^2 (R_β^2) as an index of effect size. Dyads who failed the experimental manipulation check have been excluded. Outgroup exposure refers to the ratio of outgroup social partners relative to the total size of judges' social networks. Partner liking refers to how much judges liked targets. Positive and negative emotion refer to judgments of targets' emotional experience during the interaction. Continuous predictors have been z-scored. † $p < .10$, * $p < .05$, ** $p < .01$.
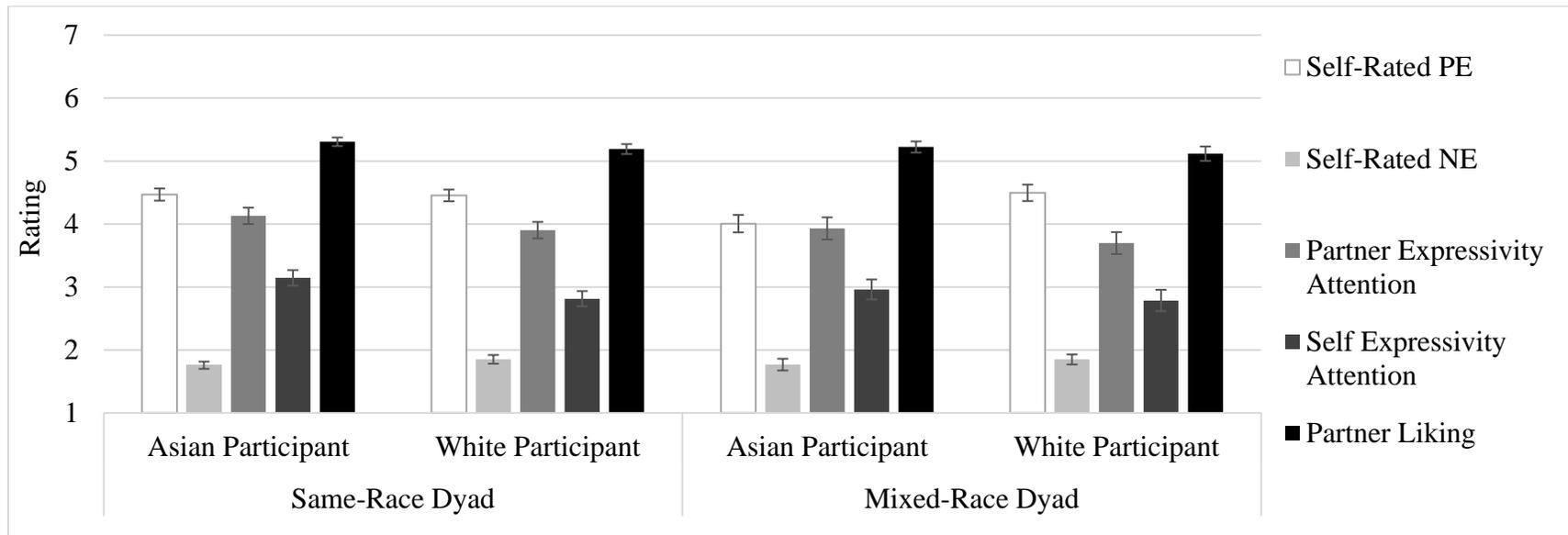
Figures

Figure 1

Social network descriptives for Study 1 participants

Number of In- and Outgroup Social Partners by Participant Race
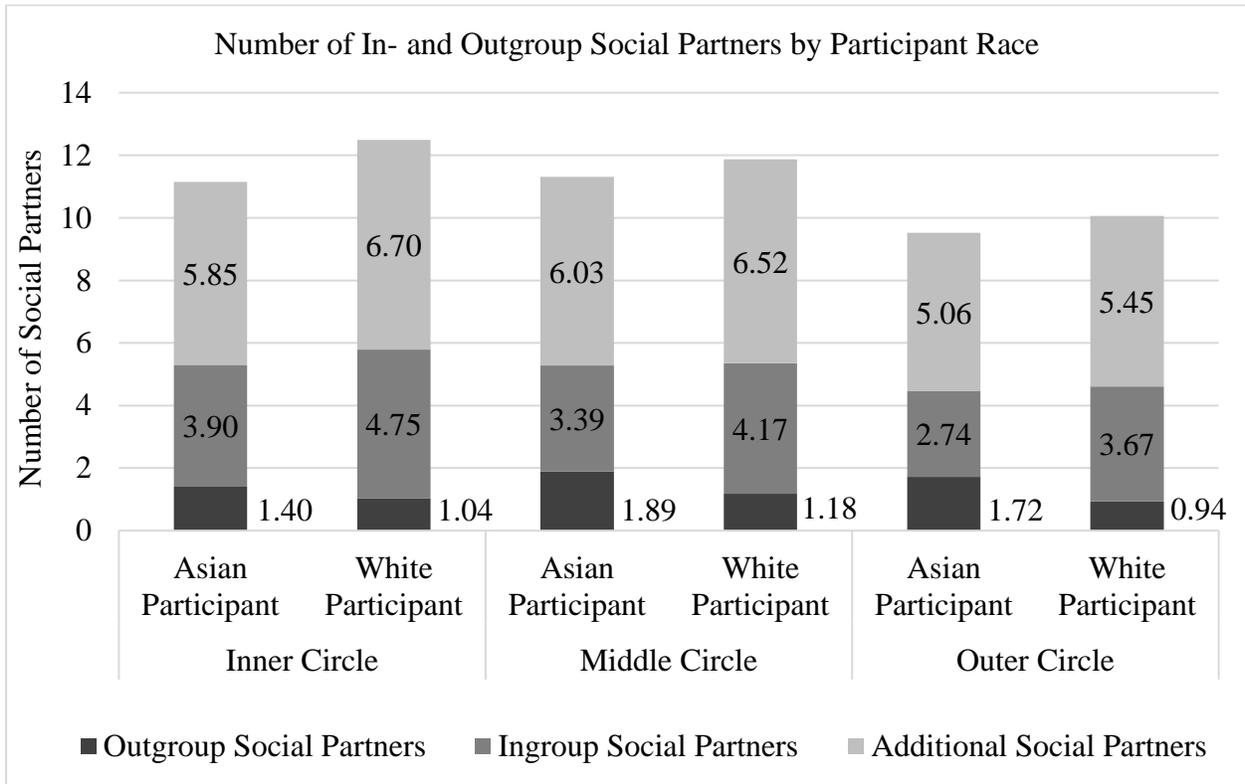
Means are presented above. "Outgroup" and "Ingroup" refer to either Asian or White social partners only based on participants' own race. "Additional" includes anyone who was mixed-race (e.g., Asian-White) or neither Asian nor White. The height of each column represents the combined total of people across these categories (i.e., the sum of outgroup, ingroup, and additional social partners in participants' inner, middle, and outer social circles).
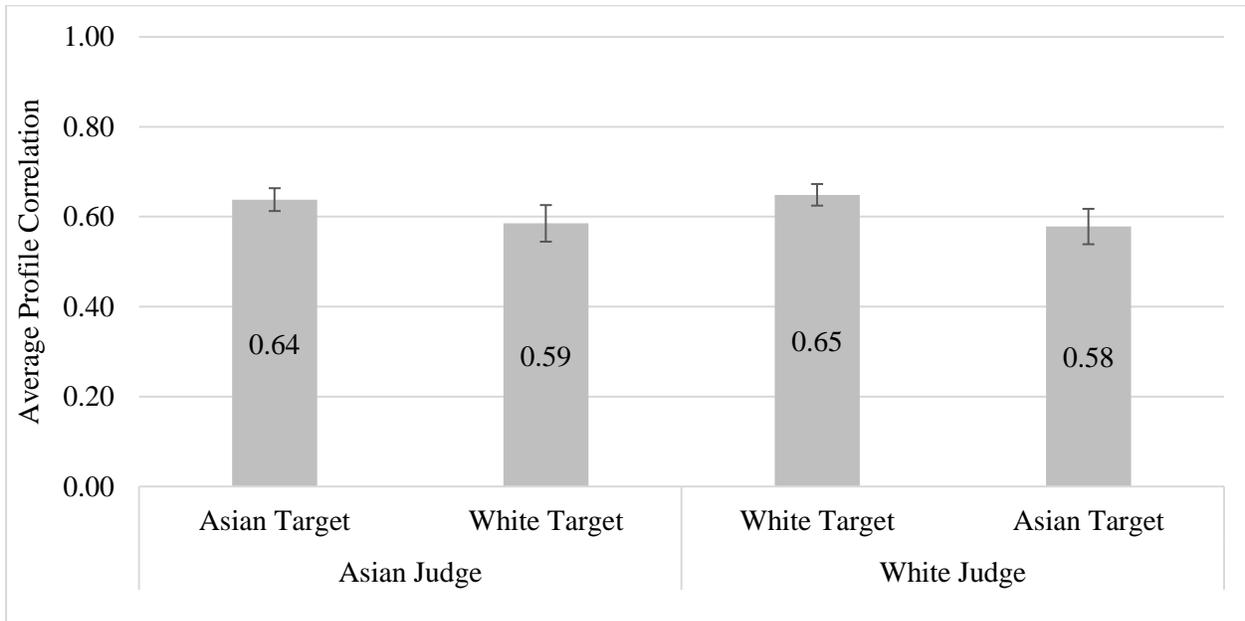
Figure 2

Study 1 participants' experiences during the interaction by dyad type and participant race.

Means are graphed above. Error bars represent SEs. Emotional experience and expressivity attention were rated on a scale from 1 (*not at all*) to 7 (*a great deal*), while liking was rated from 1 (*disagree strongly*) to 7 (*agree strongly*).
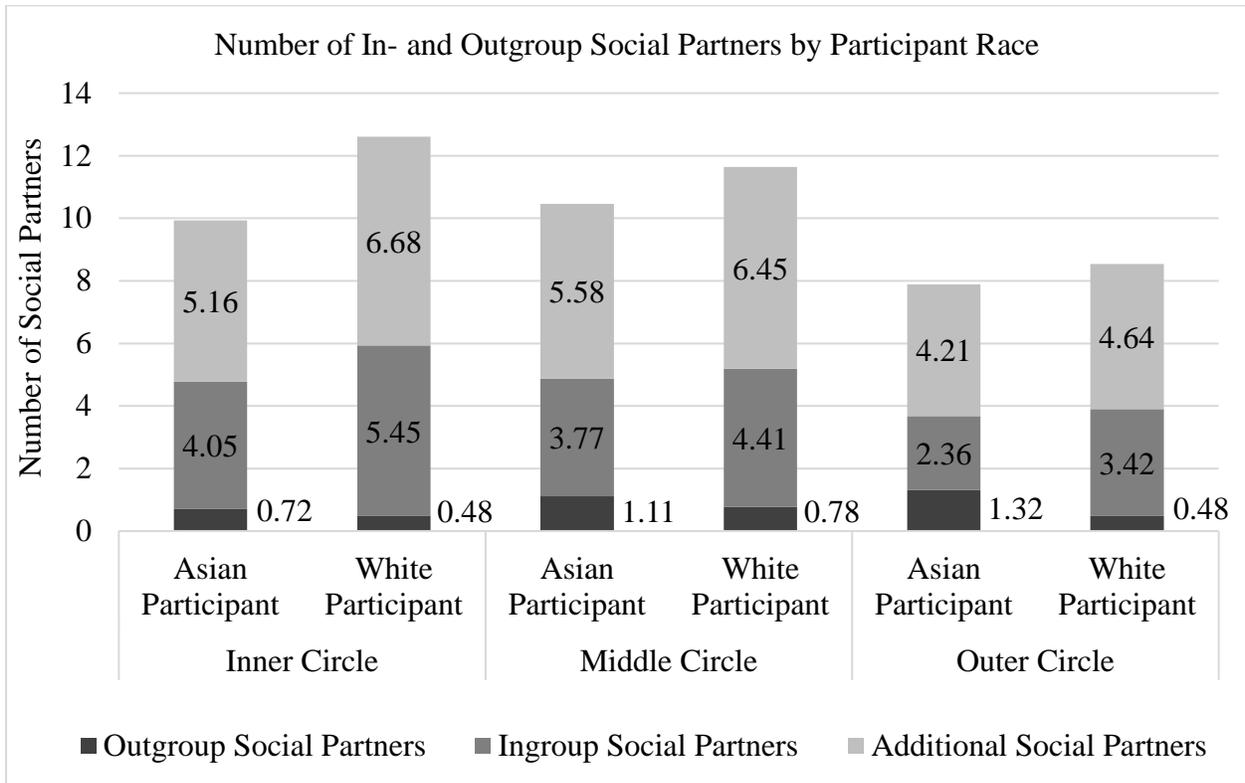
Figure 3

Study 1 mean profile correlations for emotion judgments by judge and target race.

Raw correlations are graphed above. Error bars represent SEs.

Figure 4

*Social network descriptives for Study 2 participants*

Number of In- and Outgroup Social Partners by Participant Race

Means are presented above. Dyads who failed the experimental manipulation check have been excluded. "Outgroup" and "Ingroup" refer to either Asian or White social partners only based on participants' own race. "Additional" includes anyone who was mixed-race (e.g., Asian-White) or neither Asian nor White. The height of each column represents the combined total of people across these categories (i.e., the sum of outgroup, ingroup, and additional social partners in participants' inner, middle, and outer social circles).

Figure 5

Study 2 participants' experiences during the interaction by dyad condition and participant race.
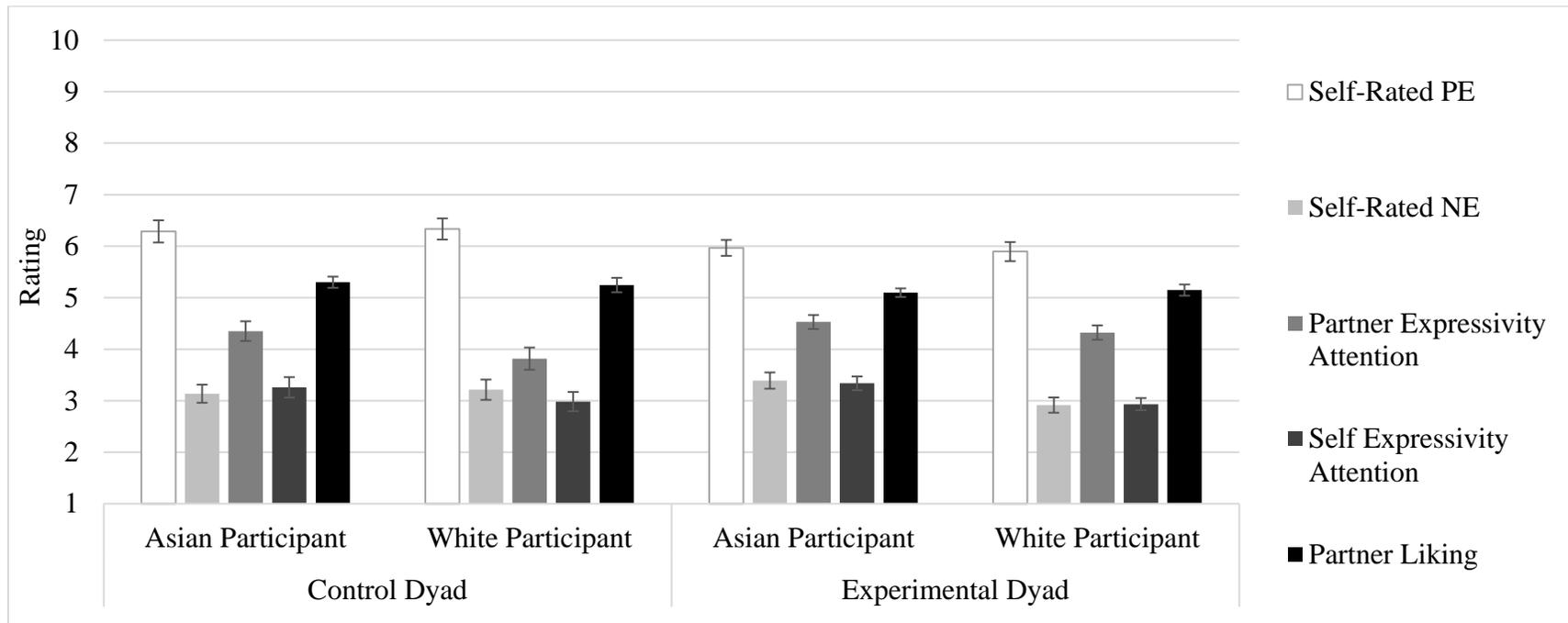
Means are graphed above. Error bars represent SEs. Dyads who failed the experimental manipulation check have been excluded. Emotional experience was rated on a scale from 1 (*not at all*) to 10 (*a great deal*). Expressivity attention was rated on a scale from 1 (*not at all*) to 7 (*a great deal*), while liking was rated from 1 (*disagree strongly*) to 7 (*agree strongly*).

Figure 6

Study 2 mean profile correlations for emotion judgments by dyad condition and judge race.
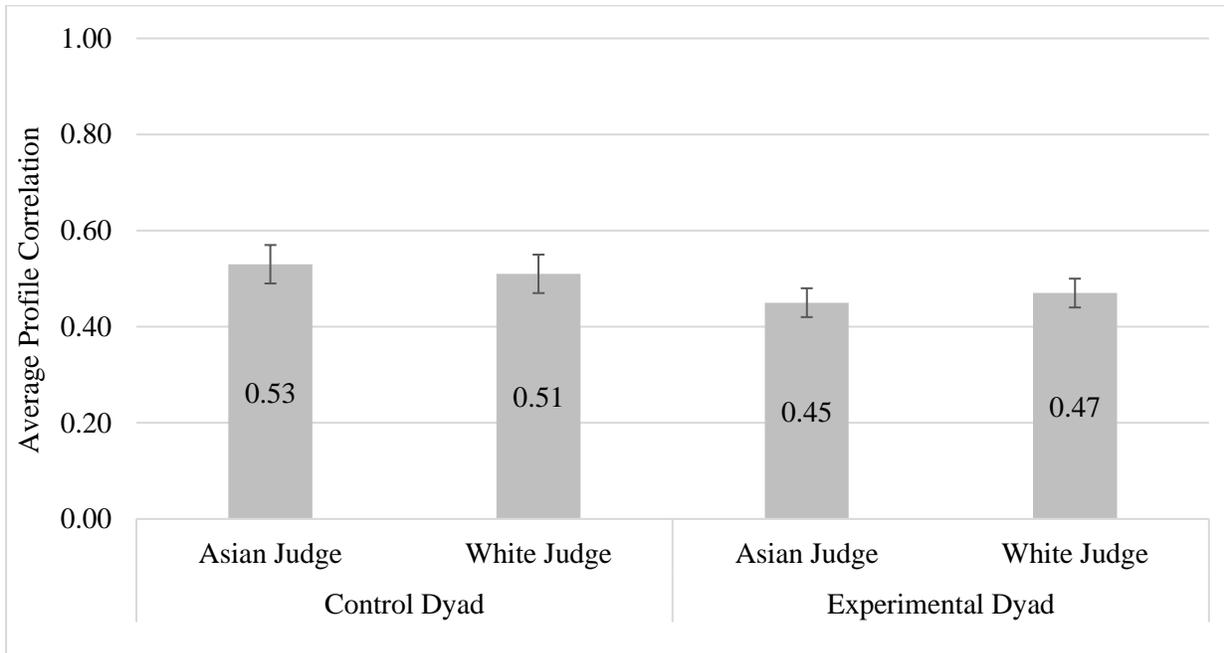
Raw correlations are graphed above. Error bars represent SEs.

Appendices

Appendix A

<center>Study 1 Game Instructions</center>

Participants received the following instructions after filling out the baseline emotional experience questionnaire.

You and your interaction partner are about to play a word-guessing game based on the card game Taboo.

*During this activity, you and your partner will each receive a deck of cards with words at the top of the card (called **prompt** words) and relevant words and phrases listed below (called **taboo** words and phrases). You will take turns trying to get the other person to guess prompt words without being able to use any of the taboo words or phrases listed on each card. For example, let's say you have a card where the prompt is "watch" and the taboo words are "look," "time," "wrist," "clock," and "wears." This means that you cannot say any of the taboo words aloud when trying to provide clues for your partner to say the correct answer ("watch"). The experimenter will press a buzzer whenever a taboo word is said aloud. <u>If a prompter says a taboo word, they will lose a point from their own score and must move on to the next card</u>.*

You and your partner will play four rounds total, with you both playing two turns as a guesser and two as a prompter. Each round will take two minutes. When you are playing as the guesser, keep track of <u>how many cards you answer correctly</u>. When you are playing as the prompter, <u>you can skip one card per round</u>; however, if you've started giving clues about a card to your partner, <u>you cannot skip it</u>. You will have to continue

<center>112</center>

playing that card until your partner gets the correct answer or until the round ends (if

you partner cannot guess the answer before the timer goes off).

This interaction will last about ten minutes. Your experimenter will randomly assign the

guesser and prompter for the first round of the game, after which you and your partner

will switch roles after each two-minute interval. Make sure to keep track of the timer and

your scores during each round.

Appendix B

Attention to Expression Questionnaire

Think back over the interaction you just had.

1	2	3	4	5	6	7
Not at all			Moderately			A great deal

1. _____ How much did you attend to **your own** facial expressions?

2. _____ How much did you attend to **your own** body language?

3. _____ How much did you attend to what **you** were saying?

4. _____ How much did you attend to how **you** were speaking (e.g., tone of voice)?

5. _____ How much did you attend to **your partner's** facial expressions?

6. _____ How much did you attend to **your partner's** body language?

7. _____ How much did you attend to what **your partner** was saying?

8. _____ How much did you attend to how **your partner** was speaking (e.g., tone of voice)?

Appendix C

Study 2 Condition Instructions

Depending on their condition assignment, participants received the following instructions after reviewing the game instructions.

<u>Control condition</u>

During the upcoming interaction, please try your best to focus on doing well at playing the game.

*You should **<u>pay careful attention to the clues you receive & try to provide useful, creative clues so you & your partner can guess as many correct answers as you can in each round</u>**.*

Being able to stay attentive to the game during the interaction could improve your performance.

<u>Experimental condition</u>

During the upcoming interaction, please try your best to focus on your partner & what they are communicating.

*You should **<u>pay careful attention to your partner's facial expressions & body language, along with what they say & how they say it</u>**.*

Being able to accurately determine what your partner is thinking & feeling during the interaction could improve your performance.