AN INVESTIGATION OF CULTURAL SENSITIVITY AND RACIAL BIAS IN THE MOVIE FOR THE ASSESSMENT OF SOCIAL COGNITION

by
Veronica Jane McLaren

A dissertation submitted to the Department of Psychology,
College of Liberal Arts and Social Sciences
In partial fulfillment of the requirements for the degree of

Doctor of Philosophy
In Clinical Psychology

Chair of Committee: Carla Sharp, Ph.D.

Committee Member: Julia Babcock, Ph.D.

Committee Member: Yusra Ahmed, Ph.D.

Committee Member: Quenette Walton, Ph.D.

University of Houston
October 2021

Abstract

Mentalizing, or one's ability to understand the mental states of oneself and others, is vital for social functioning and is predictive of a wide range of social and psychological outcomes. Mentalizing ability develops via social learning and is dependent upon the environment in which one is raised. The Movie for the Assessment of Social Cognition is a widely used and ecologically valid measure of social cognition in which participants answer multiple-choice questions about the thoughts and feelings of characters in a short movie. While we know that there are cross-cultural differences in mentalizing, no study has yet investigated the impact of race within the USA on MASC scores. Thus, the aims of the proposed study are to, in a sample of young adults (N = 1,230), 1) investigate the factor structure of the MASC, 2) evaluate the MASC for bias based on race and ethnicity by evaluating its measurement invariance across racial and ethnic groups, 3) investigate the contribution of cultural sensitivity and racial ingroup effects to any cross-group differences in MASC performance using qualitative and quantitative exploratory analyses, and 4) evaluate cross-racial differences in patterns of mentalizing. Factor analytic results support a unidimensional factor structure, and measurement invariance analysis provides evidence for racial invariance. Non-Hispanic White individuals are more likely to select the "correct" answer than other racial groups, suggesting that optimal mentalizing strategy and mentalizing values and norms may differ based on racial group membership. However, limitations of measurement invariance in assessing sources of common method variance limit the strength of these conclusions.

An investigation of racial bias and cultural sensitivity in the Movie for the Assessment of Social Cognition

Mentalizing is a social cognitive process that is vital for one's functioning in social contexts. Importantly, one's framework of mentalizing is dependent upon the environment in which one is raised, resulting in cross-cultural differences in mentalizing. Despite the importance of culture to mentalizing, very few studies of mentalizing from a cross-cultural perspective exist. A barrier to cross-cultural investigation of mentalizing is a lack of culturally sensitive assessment tools. The Movie for the Assessment of Social Cognition (MASC), one of the most frequently used measures of mentalizing, has not yet been evaluated for cultural sensitivity. Because this measure was developed in the context of White Germans and Americans, it may rely upon norms and expectations specific to those cultural groups and not account for differences in the norms and expectations used for mentalizing in other cultural groups. Additionally, the MASC has also not been evaluated for racial bias. Because all the characters in the MASC are White, it may be subject to racial ingroup effects which bias it against non-White individuals. Therefore, the current study has four aims. First, we aim to investigate the factor structure of the MASC. Second, we aim to evaluate the MASC for racial and ethnic bias using measurement invariance (MI) analyses, which investigates bias by identifying differences in individual item performance based on group membership and therefore items which are differentially related to the latent trait based on group membership. Third, we aim to analyze the nature of any cross-group differences using qualitative and quantitative exploratory methods. Fourth, we aim to investigate cross-racial differences in mentalizing after adjusting for sources of bias.

Mentalizing

Mentalizing, also called Theory of Mind, is one's social cognitive ability to consider and understand the mental states of oneself and of others (Allen et al., 2008). The concept of mentalization developed from the psychoanalytic tradition (Lecours & Bouchard, 1997; Marty &

M'Uzan, 1963). Theory of Mind developed roughly in parallel from the work of primatologists and gained traction in the realm of social psychology (Premack & Woodruff, 1978). There are some distinctions between Theory of Mind and mentalizing; for instance Theory of Mind is often criticized for having a narrow focus on specific tasks (i.e. the false belief task—a task in which individuals must understand that a character in a story can believe something that the individual knows to be false) and age groups (i.e. young children; Sharp & Fonagy, 2008). However, they are frequently, and will currently be, used interchangeably. One's ability to mentalize optimally in a variety of contexts is vital for navigating a social world.

Mentalizing is an essential skill for social development and overall wellbeing. Childhood mentalizing ability is associated with increased prosocial behavior (Imuta et al., 2016), popularity (Slaughter et al., 2015), friendships (Fink et al., 2015), as well as later increased peer acceptance and reduced peer rejection (Caputi et al., 2012). Additionally, in close relationships, one's mentalizing ability improves the quality of the relationship as well as the other's wellbeing; neural activity associated with theory of mind during interactions with one's romantic partner predicts that romantic partner's well-being and relationship satisfaction (Dodell-Feder et al., 2016), and increased parental mentalizing improves children's attachment security and emotion regulation ability (Camoirano, 2017). In the context of a single interaction, mentalizing improves interpersonal rapport (Blatt et al., 2010; Todd et al., 2011), increases positive evaluations of the mentalizer (Goldstein et al., 2014), and leads to many other positive social outcomes (see Todd & Galinsky, 2014 for a review). Mentalizing is also a source of resilience; increased mentalizing reduces the strength of the association between childhood abuse and symptoms of psychopathology (Weijers et al., 2018) and reduces the risk of developing psychosis for those at-risk (Kim et al., 2011). Further, mentalizing can be used as a mechanism of change in psychotherapy; improvements in mentalizing ability during mentalization based treatments have been associated with symptom reduction for many psychopathologies (Oehlman Forbes et al.,

2020). Given its importance in social outcomes and wellbeing, a full understanding of mentalizing and social cognition across populations is important in order to more effectively target and improve these outcomes.

Mentalizing from a cross-cultural perspective

We have long understood that the development of mentalizing ability is contingent upon the environment in which one is raised. Mentalizing is thought to be passed down generationally through social learning via a secure attachment and epistemic trust with a caregiver (Kim et al., 2011). As a result, differences in environment during development can lead to differences in reasoning about others' minds. Much work has been done to understand the impact of one's parents, specifically attachment and parental mentalizing ability, on the development of mentalizing. For example, more secure attachment to caregivers has been linked to improved emotion recognition and mental state reasoning in children (Repacholi & Trapolini, 2004). However, differences in one's environment beyond parental figures can also produce differences in mentalizing ability. Number of siblings (Jenkins & Astington, 1996; Perner et al., 1994) and access to language at home (Peterson & Siegal, 1997), for example, have both demonstrated an influence on the development of one's mentalizing ability. Therefore, broad differences in environments, such as those that occur across cultural groups, likely also lead to differences in mentalizing across those groups.

A recent systematic review by Aival-Naveh and colleagues brought attention to the impact of culture on mentalizing as a gap in our study of mentalizing that has been too long overlooked (2019; Jurist & Perez Sosa, 2019). These authors posit that, while mentalizing is universally important, different cultures emphasize different aspects of mentalizing and therefore have distinct developmental trajectories of and strengths in mentalizing. Specifically, they argue that differential emphasis on the self and the other leads to cultural differences in the developmental trajectory of mentalizing for individuals from collectivist and individualist

cultures. For example, children from individualist cultures tend to develop an understanding that one's own beliefs may differ from those of others earlier than children from collectivist cultures (Slaughter & Perez‐Zapata, 2014), a difference which has been linked to neural differences during theory of mind tasks (Kobayashi et al., 2006; Koelkebeck et al., 2011), while children from collectivist cultures develop an understanding of differing opinions between self and others faster than children from collectivist cultures (Shahaeian et al., 2011; Wellman et al., 2006). This difference in developmental trajectory indicates cross-cultural differences in emphasis on elements of mentalizing.

The framework proposed by Aival-Naveh is an important step forward in that it acknowledges that different cultures emphasize different aspects of mentalizing and therefore have different patterns of strengths; however, this approach still attempts to apply an existing framework of mentalizing to different cultures. In other words, by measuring mentalizing ability as it presents within one's own culture, one is imposing one's own version of "optimal mentalizing" to other cultures. Importantly, predominant theories of mentalizing were largely developed by and in the context of White Europeans and Americans and therefore applying these frameworks outside of the cultural contexts of these groups may be imposing norms and values specific these cultures that are not shared by other cultural groups. Attempting to apply a framework developed in the context of one culture to the mentalizing of another culture may lead to conclusions that the second culture is worse at mentalizing or does not value mentalizing, when in reality they may have a different set of values and norms that change the way "good mentalizing" looks. Jurist and Perez Sosa (2019) liken this to Gilligan's criticism of Kohlberg's theory of moral development; when attempting to understand women's moral development using a framework proposed by and for men, women appear to be "morally lacking" because they reach stages late or not at all. However, when a separate theory of moral development based on the values and expectations for women is used, a more informed understanding and study of the

moral development of women is possible. Given current theories regarding the function of mentalizing, cross-cultural study of mentalizing may require a similar approach. Mentalizing is a tool which enables us to function socially within our society. Therefore, optimal mentalizing is that which aligns with the norms and expectations of one's environment. As such, cross-cultural differences in norms and expectations influence the mentalizing of individuals within that culture.

Cross-cultural differences have been discovered in which cues are relied on in mentalizing, the significance of those cues, and the way that those cues are expressed. Differential cues are relied on to make mentalizing interpretations across different cultures. For example, in facial emotion recognition, there are differences cross-culturally in preference for the eyes region or the mouth region as primary indicators of mental state; individuals from Japan are more likely to rely on the eye region to determine facial emotion while individuals from America are more likely to rely on the mouth region. As a result, when faces are modified such that eye regions and mouth regions display different emotions, individuals from different cultural groups tend to identify different emotions given the same stimulus (Yuki et al., 2007). Additionally, when identifying the emotion on a face surrounded by other faces, individuals from East Asia are influenced by the emotions expressed on surrounding faces, while individuals from America tend to only consider the emotion of the central face (Masuda et al., 2008). Individuals from different cultures may also interpret identical cues differently; while basic emotions seem to be universally identifiable (Ekman & Friesen, 1971), expressions of more complex emotions are differentially identified by individuals from Britain, Spain, and Japan (Baron-Cohen, 1996). Individuals from American and Japan also differentially rate the level of intensity expressed in the same facial expressions (Matsumoto & Ekman, 1989). There have also been differences found in the way emotions and other mental states are expressed cross-culturally, including differences in intensity of emotion expression (Friesen, 1972; Matsumoto et al., 1998) and expectations of masking

certain emotions (Matsumoto et al., 2009). These differences in mental state expression have led to the proposal of "dialect" theory, which suggests that cultures have unique styles of expression that are passed down through social learning (Elfenbein & Ambady, 2003). Taken together, these findings indicate that what is considered "accurate" mentalizing in a given situation is highly dependent on the norms and beliefs of the culture in which an individual was raised.

Importantly, much of the aforementioned work in identifying differences in mentalizing norms based on culture has focused on differences based on country of nationality, specifically between "Western" and "Eastern" countries. Authors of these studies typically cite individualism in Western countries and collectivism in Eastern countries to explain these differences. However, this simplification of cross-national cultural differences to "Western versus Eastern" and "individualism versus collectivism" has long been criticized as reductive and missing the complexity of cultures by reducing each half of the world to one ubiquitous culture (Hermans & Kempen, 1998; Vignoles et al., 2016; Voronov & Singer, 2002). While there may be similarities within these categories, there are other cross-cultural differences within these broad categories that remain largely unexplored in the context of mentalizing. Additionally, cultural differences exist not only between countries, but also within countries. Racial groups are perhaps the most important example of broad cultural differences that can exist within a single country. Unfortunately, little research has been done on the impact of race on the norms and expectations regarding mentalizing. Instead, much of the research on the impact of race on mentalizing and related constructs has focused on the racial ingroup effect.

Racial ingroup effects

Racial ingroup effects occur when individuals perform better on tasks when they are of the same race as the individuals they are observing. Racial ingroup effects have been identified across social cognitive constructs, many of which are called upon during the process of mentalizing, including facial processing, emotion recognition, and empathy. Importantly, while

some tasks used to investigate the racial ingroup effect have objectively correct answers (e.g. facial memory), others rely on universal agreement upon a subjectively correct answer (e.g. emotion recognition). As discussed in the previous section, "correct" answers on mentalizing and related constructs are often contingent upon cultural norms and expectations. Therefore, tasks which have subjective correct answers may be influenced by cross-cultural difference in the cues used to determine a "correct" answer in addition to, or rather than, a racial ingroup effect. The following section, therefore, focuses primarily on differential performance based on stimulus race for tasks without this level of subjectivity (i.e. tasks with objectively correct answers or tasks without correct answers, such as those comparing neural patterns) with the aim of discussing racial ingroup effects separate from cultural factors.

Facial processing and emotion recognition. Differences in processing faces have been identified based on the target face's race. When processing a face from a different racial group than one's own, the face is processed less holistically (Michel et al., 2006) and less deeply (Levin, 1996). These differences in face processing are exhibited early in life; 3-month-olds demonstrate longer looking time at novel faces of their own race than novel faces of different races (Bar-Haim et al., 2006). However, this difference is associated with exposure to racial groups; infants who were raised in environments with predominantly individuals of other races did not demonstrate reduced looking-time for those races (Bar-Haim et al., 2006). This provides support for differential exposure as a mechanism for racial outgroup effects.

Potentially as a result of this difference in facial processing based on race, facial memory and facial emotion recognition have also demonstrated a racial in-group effect. A meta-analytic study aggregating data from 39 articles provides strong support for the existence of an own-race bias in memory for faces; individuals are significantly better at recognizing faces of individuals from their own race than individuals of a different race (Meissner & Brigham, 2001). When identifying an emotion based on facial expression, a racial ingroup effect has been identified;

individuals demonstrate decreased neural activation when identifying the emotion displayed by faces of other races when compared to their performance identifying the emotion displayed by faces of their own race (Adams et al., 2009; Dodell-Feder et al., 2016; Pinkham et al., 2017). These racial ingroup effects persevere even in in individuals with training in emotion recognition; trained coders tend to rate individuals from their own racial group as more neutral and more interested than individuals from other racial groups (Babcock & Banks, 2019).

Empathy. A racial ingroup effect was hypothesized for empathy as early as 1978 when it was suggested that empathetic response will be increased when the perceiver and target share membership in a common social category or group (Hornstein, 1978). Subsequent investigation of such an effect has provided substantial support for its existence. Johnson et al. found that when evaluating a defendant in a criminal case, White individuals reported feeling greater empathy for White defendants than for Black defendants (2002). Neural responses associated with empathy when viewing others in pain are significantly reduced when the other is of a different racial group (Avenanti et al., 2010; Mathur et al., 2010; Xu et al., 2009). These neural differences correspond to decreased empathy and altruistic motivation for individuals outside of one's own racial group (Mathur et al., 2010). Additionally, individuals with higher levels of implicit racial biases demonstrated further reductions in neural response to the pain of individuals of racial groups different from their own (Avenanti et al., 2010).

Mentalization. Mentalization relies on facial processing, emotion recognition, and empathy as sources of cues by which to ascribe mental state attributions. Given that each of these constructs has demonstrated a racial in-group effect, it is likely that mentalization will be subject to such an effect as well. Preliminary studies of differences in mentalizing based on target race have supported this possibility. When inferring intentions of an actor, an individual's mirror neuron system is more highly activated when the actor is of the same race as the observer (Liew et al., 2010). However, it is not clear how this difference in neural activity translates to

differences in mentalizing ability for individuals of different races. Additionally, research using implicit association tasks has demonstrated that individuals tend to attribute more complex emotions to individuals of their own race than to individuals of other races (Paladino et al., 2002).

Overall, more complex, thorough, and holistic social cognition seems to be related to observing individuals of one's own race compared to of other races. Thus, mentalizing of individuals from racial groups different from one's own may be hindered. Given the importance of mentalizing to social outcomes and wellbeing as well as its dependence on cultural factors, it is vital to have measures of mentalizing that are not only valid and reliable, but also culturally sensitive.

The MASC as a measure of mentalizing

The MASC is a measure of mentalizing developed as a collaboration between German and American psychologists (Dziobek et al., 2006). This task consists of a film in which four characters meet for dinner. The movie was filmed in German and dubbed in English, so it could then be administered in both languages. Throughout the movie, the video is stopped and the participant is presented with a multiple choice question about what a character is thinking or feeling. For each question, there is one "correct" answer and one answer attributed to each of three types of incorrect mentalizing- hypermentalizing (i.e. an overattribution of mental states beyond what there is evidence for), undermentalizing (i.e. a failure to fully consider mental states), and no mentalizing (i.e. a complete lack of consideration of mental states). This measure is commonly used—since its development, the MASC has been used in more than 100 publications.

The popularity of the MASC can likely be attributed to its many strengths. First, the MASC is an experimental measure, therefore it does not rely on an individual's perception of their own ability but rather requires that the individual demonstrate their ability. Second, it is

ecologically valid; the mentalizing an individual is asked to perform in the MASC closely resembles mentalizing in a real-world setting. Third, the MASC is challenging enough to detect differences in mentalizing ability among adults, which is a disadvantage of many previously developed measures of mentalizing. Fourth, because it is multiple choice and participant-led, it does not require many resources to administer and score. Finally, the MASC was developed thoughtfully and from a strong theoretical basis; the mental states it asks participants to infer span a range of modalities (thoughts, emotions, intentions), valences, types and quality of cues available for making inferences, difficulties, and social cognitive concepts (e.g. false belief, faux pas, metaphor, sarcasm). These theoretical considerations have led to postulations that the MASC is a promising assessment across dimensions of mentalizing.

Evidence from neural studies has been used to suggest that the MASC is a theoretically valid measure of mentalizing. Neural evidence suggests that mentalizing exists across four poles: automatic/controlled, internal/external, self/other, and cognitive/affective (Luyten et al., 2012). Importantly, optimal mentalizing spans the range of these dimensions. Therefore, comprehensive measures of mentalizing must span ranges of these dimensions. Luyten et al. suggest that the MASC, while only measuring controlled mentalizing of the other, demonstrates good coverage of the internal/external dimensions and cognitive/affective dimensions (2012). Given the possibility of multidimensionality, Fossati et al. investigated the factor structure of the MASC and found strong evidence of a unidimensional factor structure (2018). This unidimensionality suggests that, while the MASC may cover these dimensions of mentalizing, its total score represents a unidimensional mentalizing construct. In other words, each item on the MASC may require an integration of internal and external cues and cognitive and affective information rather than individual reliance on, for example, internal cues for certain items and external cues for others. However, despite this evidence for unidimensionality, some have used multidimensional interpretations of the MASC; most commonly, the thoughts/feelings/intentions split is used to

represent the cognitive/affective mentalizing dimension (e.g. Vaskinn et al., 2018; Wastler & Lenzenweger, 2021). However, there is an important theoretical distinction between the cognitive/affective mentalizing dimension (i.e. the use of cognitive knowledge vs affective information when determining one's mental representation) and mentalizing of thoughts, feelings, and intentions as measured by the MASC (i.e. whether the mental state being inferred is a thought, a feeling, or an intention). As such, there is need to replicate the results of Fossati et al and confirm the unidimensionality of the MASC. While the MASC's factor structure may still lack broad agreement, the MASC has demonstrated strong psychometric properties a range of populations.

Dziobek et al. (2006) initially developed the MASC in the interest of creating an experimental measure of mentalizing sensitive and challenging enough to detect differences in mentalizing ability among adults with and without Autism Spectrum Disorder (ASD). Therefore, initial evidence for its validity centers around individuals with ASD. The MASC was found to distinguish between individuals with and without ASD using the English version (Dziobek et al., 2006), German version (Müller et al., 2016), and later developed Spanish (Lahera et al., 2014) and French (Martinez et al., 2017) versions. Additionally, its Italian translation has demonstrated reliability and associations with other measures of mentalizing in populations of healthy adolescents and clinical and nonclinical adults (Fossati et al., 2018). Further, the MASC has been administered in populations from settings including inpatient (Sharp et al., 2016), forensic (Mayer et al., 2018), and academic (Taubner et al., 2017); with ages ranging from adolescent (Hart et al., 2017) to elderly (Lecce et al., 2018); and with psychopathologies including depression (Wilbertz et al., 2010), BPD (Preißler et al., 2010), and schizophrenia (Montag et al., 2011). Despite its considerable strengths, broad use, and strong evidence for validity, the MASC has noteworthy gaps and limitations in our understanding of its cultural sensitivity.

The MASC and cultural sensitivity

As discussed, mentalization is socially bound, therefore measurement must be sensitive to culture. Gaps in the research on the MASC prevent the evaluation of its cultural sensitivity. No study has evaluated the MASC's performance across racial groups, few studies using the MASC report the racial demographics of the participants, and those that do report race include predominantly White samples (McLaren et al., 2022). Of the 115 studies identified by this meta-analysis as having administered the MASC on unique datasets, only 10 report racial demographics. Of those 10, three reported a sample of more that 90% Caucasian (Buhlmann et al., 2015; Duval et al., 2018; Sharp, 2016), three report a sample of majority Caucasian and do not report racial demographics of the remainder of the participants (Divilbiss, 2009; Hassenstab et al., 2007; Schneider et al., 2016; Zainal & Newman, 2018), and three report full racial demographics but do not perform any analyses investigating the relationship between race and MASC performance (Park, 2010; Washburn et al., 2016; Wastler & Lenzenweger, 2019). Therefore, the MASC's generalizability across racial groups has not been fully investigated. This investigation is especially important due to two limitations that threaten the MASC's cultural sensitivity: an all-White cast potentially invoking racial ingroup effects and development in a White German and American context discriminating against the culture-specific mentalizing of other cultural contexts.

Cultural differences in social cognition. Given the development in a White German and American context, cues used to determine mental state on the MASC may be specific to this culture may not translate to other cultures. The mental states of the characters in the MASC are actually unknown; therefore, relying on a designated "correct" answer rather than an individual's own perception of their ability prioritizes agreement with social norms and beliefs regarding mental states of the measure developers rather than "accurate" mentalizing. Additionally, while the movie that the MASC relies on appears ecologically valid, the social setting, conventions, and interactions it demonstrates may not be generalizable to all cultural groups.

Differences in social cognition as a function of target race. Given the strong evidence of racial ingroup effects on constructs closely related to mentalizing and preliminary evidence of racial ingroup effects on mentalizing itself, it is likely that mentalizing is indeed subject to racial ingroup effects. All characters in the MASC are White. Therefore, the MASC is at risk of being influenced by a racial ingroup effect; when asked to mentalize only White characters, non-White individuals will likely perform worse than White individuals with similar mentalizing abilities. Therefore, the MASC may exhibit a bias against non-White participants. Such a bias has been observed on both the Reading the Mind in the Eyes task (RMET; Baron-Cohen, 1996), a measure of emotion recognition using the eye region of the face as stimulus, and The Awareness of Social Inferences Task (TASIT; McDonald et al., 2003), a video-based test of mental state attribution. The RMET features only White faces, and non-White participants tend to exhibit poorer performance than White individuals (Dodell-Feder et al., 2020; Pinkham et al., 2017). The TASIT features predominately White actors and has preliminarily demonstrated a similar bias against non-White participants (Pinkham et al., 2017).

Current study

Against this background, the overall objectives of the current study are to evaluate the cultural sensitivity and racial generalizability of the MASC and provide recommendations about the cross-cultural use of the MASC. These objectives will be investigated in a sample of undergraduates using measurement invariance (MI) methods. MI is ideal for the investigation of bias due to its ability to separate performance on a measure from actual levels of the latent variables. This means that MI methods cannot be used to draw conclusions about differences in actual ability outside of the context of the measure itself. Rather, MI identifies sources of bias in a measure that contribute to systematic differences in performance relative to actual ability level for specific groups. These differences might result in worse performance on the measure, but importantly do not indicate differential levels in the latent trait assessed by the measure. MI

assesses the psychometric equivalence of a measure across groups for the purposes of determining whether that measure has the same meaning to those groups (Putnick & Bornstein, 2016). MI is based upon the idea that for scores on a measure (i.e. latent factor means) to be comparable across groups, the measurement structure of the latent factor and the individual items must be invariant across those groups (Van De Schoot et al., 2015). This means that the association between items and latent factors should not depend on group membership (Mellenbergh, 1989). In particular, the factor structure, factor loadings, item intercepts, and item unique variances must be equivalent across groups to claim that the items and latent traits perform equivalently across these groups and for their mean scores on the measure to be compared (Putnick & Bornstein, 2016). Only if all of these equivalences hold true can differences in scores across groups be interpreted as meaningful.

For the purposes of the current study, MI will be investigated from a structural equation modeling (SEM) framework using a series of multiple group confirmatory factor analyses (CFA) with increasing constraints. Each of these increasing constraints represents an additional equivalence that is expected between groups. At each step, if the model fit remains acceptable and does not fit significantly worse than the previous model, invariance at that step is supported. First, the configural model is tested, which constrains factor structure (i.e. for each latent variable, the pattern of items with estimated loadings and those with loadings fixed at 0) to be equivalent across groups. Configural invariance indicates that the same items make up the same latent variables in each group, while configural non-invariance indicates that some items load onto different factors depending on group membership. Second, the metric model is tested, which constrains the factor loadings of each item to be equivalent across groups. Metric invariance indicates that in each group, each item contributes to the latent construct in a similar degree, while metric non-invariance indicates that some items are more or less indicative of the latent construct depending on group membership. Third, the scalar invariance model is tested,

which constrains the item intercepts to be equivalent across groups. Scalar invariance indicates that differences in items scores represent differences in the latent trait, while scalar non-invariance indicates that differences in item scores are independent of differences in the latent trait. At each step, adjustments can be made (e.g. omitting items, releasing constraints) to achieve a better fitting model. These adjustments can be used to make recommendations about future use of the measure across groups.

Given the MASC's development in a White context and use of all-White characters as stimulus, non-Hispanic White individuals will be considered part of the "development group." The impact of racial ingroup effects will be investigated using individuals whose race differs from the characters, specifically Black and Asian Americans. Cultural factors will be investigated using race and ethnicity as a proxy for cultural groups. In addition to the groups previously mentioned, White Hispanic individuals will be included as a cultural group. While none of these groups are culturally homogenous, racial and ethnic groups have shared cultural backgrounds which contribute to shared norms, beliefs, and expectations regarding mentalizing. As such, performances on the MASC of individuals from the same racial or ethnic group are likely impacted in similar ways by similarities and differences between their cultural background and that of the developers. However, it is important to note that, while White Americans collaborated on the development of the MASC, the measure was created in Germany and filmed using German actors. Therefore, White Americans may also be impacted by a lack of cross-cultural generalizability, but we expect these differences to be less pronounced given the participation of White Americans in the development. Additionally, all groups may be impacted by the task being dubbed; while an investigation of the impact of dubbing on the MASC indicated that participants do not find it interfering (Dziobek et al., 2006), no study has evaluated the impact of this dubbing on performance.

The first aim is to evaluate the factor structure of the MASC. This aim will be investigated using confirmatory factor analysis (CFA) to evaluate a unidimensional, 3-factor (thoughts, feelings, and intentions), and bifactor model and determine the best fitting model. It is hypothesized that the unidimensional model will be the best fit, replicating the findings of Fossati et al. (2018).

The second aim is to evaluate the possibility of racial and ethnic bias on the MASC. This aim will be investigated using MI methods to evaluate whether items perform differentially in individuals from the same racial and ethnic groups of the developers and typical administration population of the MASC (i.e. non-Hispanic White) than in individuals from different racial (i.e. Black and Asian) and ethnic (i.e. Hispanic) groups. We hypothesize that the MASC will not demonstrate full invariance based on race and ethnicity.

The third aim is to evaluate the influence of racial in-group effects and cultural norms on differences performance on the MASC between racial and ethnic groups. This aim will be evaluated based on the results of the MI methods discussed in aim two. Items and factors which differ based on ethnic or racial group will be examined qualitatively based on their content, the context in which they appear, and specific considerations made by the developers in item creation. Differences in item or factor performance specific to differences in racial group may be influenced by cultural differences as well as by racial ingroup effects, while differences in item performance additionally or specifically related to differences in ethnic group are likely dependent on the norms and conventions of the culture in which the measure was developed. We hypothesize that when item performance differs, non-Hispanic White individuals will be more likely to select the “correct” answer than individuals from other racial and ethnic groups. However, this aim is primarily exploratory, so no additional a priori hypotheses were made.

The fourth aim is to evaluate cross-racial differences in mentalizing performance after correcting for sources of measurement variance. This aim will be investigated using Analysis of Variance (ANOVA) with adjustments made to MASC scoring based on the results of MI analyses conducted in previous aims. We hypothesize that, after adjusting for sources of measurement variance, there will not be cross-racial differences in mentalizing ability.

Methods

Participants and procedures

The sample consists of 1,230 undergraduate students (316 non-Hispanic White, 414 Hispanic White, 151 Black/African American, and 349 Asian/Pacific Islander) recruited through an online recruitment site from a large public university in the southwestern United States for participation in two larger studies. Inclusion criteria for study participation consist of (1) being 18-25 years of age, (2) sufficient fluency in English to complete all research, (3) acceptable reliability indicators, (4) completion of all study materials, and (5) self-identification as non-Hispanic White, Hispanic White, Black, or Asian/Pacific Islander.

Data for the first study was collected from February 2013 to April 2014 and inquired about race using a single item asking participants to self-identify as White- not Hispanic, Black – not Hispanic, Hispanic, Asian or Pacific Islander, Native American or Alaskan Native, or Other. Data for the second study was collected from June 2019 to May 2021 and inquired about race and ethnicity separately, asking participants to first identify as Hispanic or non-Hispanic, then to identify as White, Black, Asian/Pacific Islander, Native American/Alaskan Native, Multiracial, or Other. Individuals identifying as both Hispanic and non-White were excluded from the current study. Additionally, participants' countries of origin are unknown.

Measures

Movie for the Assessment of Social Cognition (MASC; Dziobek et al., 2006). The MASC is a video-based assessment that evaluates mentalizing ability. Participants watch a

15-minute film about four people getting together for dinner. The film is stopped at 45 points, during which the participant answers a multiple-choice question regarding a character's thoughts and feelings. Participants choose from 4 answer options: an "accurate" choice, a "hypermentalizing" choice, a "hypomentalizing" choice, and a "no mentalizing" choice. The number of times each type of answer was selected is calculated, resulting in 4 mentalizing scores: one indicating the frequency of "accurate" mentalizing and three indicating the frequency of each type of error. Additionally, four control questions unrelated to mentalizing are asked. The control score can be calculated by adding together the number of control items answered correctly and indicates general comprehension and attention. The MASC has demonstrated adequate internal consistency and test-retest reliability in healthy adults (Dziobek et al., 2006) and adults and adolescents with psychopathology (Fossati et al., 2018), associations with other measures of social cognition (Dziobek et al., 2006), and associations with fixations on eyes and pupil dilations during task administration (Müller et al., 2016). For the current study, group differences in "accurate" mentalizing score will be investigated and MI will be evaluated for "correct" versus "incorrect" responses on all questions.

Data analytic strategy

Descriptive statistics were calculated in SPSS (Version 29; *IBM SPSS Statistics for Windows*, 2017). For all aims, analyses were conducted in MPlus (Version 8; Muthén & Muthén, 2017).

Factor structure (Aim 1). Confirmatory factor analysis (CFA) was conducted on three models: a unidimensional model, a 3-factor model, and a bi-factor model. Model fit was evaluated using multiple fit indices (Kline, 2015): the root mean square error or approximation (RMSEA), with values of less than .08 indicating reasonable fit, and values above .10 suggesting poor fit (Browne & Cudeck, 1993); the comparative fit index (CFI; Bentler, 1990) with values between .95 and 1.00 indicating excellent fit and values between .90 and .95 indicating

acceptable fit (Hu & Bentler, 1999); and the standardized root-mean-square residual (SRMR), with values less than .08 indicating acceptable fit (Hu & Bentler, 1999). Importantly, as fit indices are affected by types of models, sample size, and other factors, these indices are considered holistically; models were considered good fit if these indices closely approach acceptable criteria (Beauducel & Herzberg, 2006; Clark & Bowles, 2018). Results of chi-square tests will be reported; however, the chi-square test is sensitive to sample size, so these results will not be given as much weight as the fit indices listed above (Fan et al., 1999).

In the event of multiple acceptable model fits, we calculated additional statistical indices following best practices for evaluating bi-factor models to evaluate the presence of multidimensionality (Rodriguez et al., 2016). These indices were computed using Dueber's Bifactor Indices Calculator (2017). We computed omega coefficients, which estimates the proportion of variance in the observed total score attributable to all modeled sources of common variance. Omega hierarchical (omegaH) estimates the proportion of variance in total scores that can be attributed to a single general factor, while omegaH subscale (omegaHS) reflects the unique variance associated with a subscale score after controlling for the variance due to the general factor. Factor determinacy (FD) is the correlation between factor scores; factors with FD > .90 are considered acceptable for use (Gorsuch, 1983). Construct reliability (H) indicates how well a latent variable is defined by the items; H values higher than .7 indicate well-defined latent variables (Hancock & Mueller, 2001). Explained common variance (ECV) and percent uncontaminated correlations (PUC) taken together give information about the dimensionality of the measure; unidimensionality can be accepted if ECV and PUC are both greater than .7 or if PUC is less than .8, ECV greater than .6, and OmegaH less than .7 (Reise et al., 2013; Rodriguez et al., 2016). Finally, relative parameter bias is the difference between an items loading in the unidimensional solution and its general factor loading divided by the general factor loading in

the bifactor model; average relative parameter bias less than 15% is considered acceptable (B. Muthén et al., 1987).

Racial and ethnic invariance (Aim 2). For aim 2, MI was examined across racial and ethnic groups (Hispanic White, non-Hispanic White, Black, and Asian) using the best fitting model identified in Aim 1, and model fit was evaluated using the fit indices discussed in Aim 1. A series of three models were evaluated with increasing constraints. First, the baseline model tested configural invariance to determine whether the factor structure of MASC scores is the same across racial and ethnic groups. Next, metric invariance was tested to evaluate whether the magnitude of factor loadings is equal across ethnic groups. Finally, scalar invariance was tested to evaluate whether item intercepts are equal across ethnic groups. In the event of poor model fit at any step, modification indices were examined and, when theoretically warranted, employed iteratively to improve model fit (Steenkamp & Baumgartner, 2000), which is common in cases of constraining several model parameters.

Overall model fit at each step was examined using the fit indices described above. In the context of measurement invariance, we followed Rutkowski and Svetina's recommendations for cutoffs of greater than .95 for CFI and TLI, while prioritizing these indices over the SRMR in the event of disagreement (2014). A cutoff of .08 was maintained for the RMSEA. To assess relative fit, the Satorra-Bentler χ^2 difference test was employed to assess differences in model fit with a non-significant difference in model comparison indicating model invariance for the more constrained model (Anderson & Gerbing, 1988). Considering χ^2 difference tests are susceptible to similar problems as the χ^2, including sample size dependency (Kline, 2015), additional fit indices were used to evaluate difference in model fit. Specifically, CFI change of less than .01 and RMSEA change of less than .015 (Chen, 2007) provided statistical evidence for invariance between the less constrained and more constrained model; for instances in which a discrepancy for invariance was observed across evaluative statistics, CFI change and RMSEA change were

considered more accurate tests for model comparison and were used as primary indicators of invariance.

Qualitative racial and cultural differences (Aim 3). Aim 3 was planned to be evaluated both qualitatively and quantitatively. First, items and factors identified as demonstrating cross-group variance will be evaluated for difference in performance based on agreement with correct answer and based on error type using χ^2. Then, items and factors will be evaluated wholistically for cultural considerations in the mentalizing context and characteristics specific to the development of the item or factor.

Cross-racial differences in mentalizing (Aim 4). Aim 4 was evaluated using ANOVA to compare means of MASC total and error scores across racial groups after adjusting these scores for sources of variance based on the findings of MI analyses conducted in aim 2. In the event of significant omnibus tests, follow up tests were conducted using Tukey's Honestly Significant Difference (HSD).

Results

Descriptive statistics

The sample consisted of 2,888 undergrads who were recruited across two studies. Of the 1,447 participants recruited in the first study, 41 did not consent to participating in the study, 184 did not have acceptable reliability indicators, 607 did not complete study measures, and 64 did not self-identify as non-Hispanic White, Hispanic White, Black, or Asian/Pacific Islander resulting in a subsample of 551. Of the 1,141 participants recruited in the second study, 32 were not between the ages of 18 and 25, 331 did not have acceptable reliability indicators, 8 did not complete study measures, and 92 did not self-identify as non-Hispanic White, Hispanic White, Black, or Asian/Pacific Islander resulting in a subsample of 679.

Of the 1,230 participants across both studies 238 (19.3%) were men, 987 (80.2%) women, 3 (0.2%) other/non-binary, and 2 (0.2%) declined to report their gender. Participants

were an average age of 20.79 (SD = 2.02). Regarding race and ethnicity, 316 were non-Hispanic White, 414 Hispanic White, 151 Black/African American, and 349 Asian/Pacific Islander. Participants in the first study were slightly older (M = 21.09, SD = 2.22) than those in the second study (20.56, SD = 1.81; $t(1036.19) = 4.50$, $p < .01$), but were not significantly different by gender ($\chi^2(2) = 2.55$, $p = .28$). Racial makeup also differed by study ($\chi^2(3) = 31.48$, $p < .01$); participants in the first study are more likely to be non-Hispanic White and less likely to be Asian/Pacific Islander than those in the second study.

There was a significant difference in age based on racial group ($F(3, 1217) = 5.70$, $p < .01$). Follow up tests using Tukey's HSD revealed that Asian/Pacific Islander participants (M = 20.46, SD = 1.85) were significantly younger than both non-Hispanic White participants (M = 20.95, SD = 2.07; $p = .01$) and Hispanic White participants (M = 21.01, SD = 2.07; $p < .01$). Black/African American participants (M = 20.66, SD = 2.06) were not significantly different in age from any other groups. Gender was not significantly different by racial group ($\chi^2(6) = 7.96$, $p = .241$). Only 7 participants (0.5%) had patterns of missing data across main study measures.

Factor structure

The unidimensional model demonstrated acceptable fit according to all fit indices (χ^2 (945) = 1298.02, $p < .01$; CFI = .95, TLI = .94, RMSEA = .02, SRMR = .05), as did the three-factor model (χ^2 (942) = 1288.67, $p < .01$; CFI = .95, TLI = .95, RMSEA = .02, SRMR = .05) and the bifactor model (χ^2 (900) = 1196.91, $p < .01$; CFI = .96, TLI = .95, RMSEA = .02, SRMR = .05). Factor loadings for each item in each model are displayed in Table 1. The unidimensional model did not exhibit worse fit than the bifactor model ($\Delta\chi^2$ (45) = 115.92, $p < .01$; ΔCFI < .01, ΔRMSEA < .015) or the three-factor model ($\Delta\chi^2$ (3) = 14.46, $p < .01$; ΔCFI < .01, ΔRMSEA < .015). Additionally, the three-factor model and bi-factor model did not demonstrate significantly different fit ($\Delta\chi^2$ (42) = 106.64, $p < .01$; ΔCFI < .01, ΔRMSEA <

.015). Because no models significantly improve the fit over any other model, the most parsimonious model, the unidimensional model, was accepted.

Bi-factor reliability indices were computed to confirm that no significant utility was gained by the inclusion of additional factors. Omega for the general factor was high (.92), as was omega subscale for all three subscales (all > .76). OmegaH was .92, indicating that 92% of the variance in total scores can be attributed to the general factor. However, the small difference between omega and omegaH (.006) indicates that only 0.6% of the reliable variance in total scores can be attributed to multidimensionality, while 8% is due to random error. Additionally, OmegaHS was low for all specific factors (all <.03), indicating that once variance due to the general factor is partitioned out, the subscales each account for less than 3% of the total variance.

Only the general factor (FD = .99) met the FD cutoff of .9. FD for each of the specific factors was less than .7. Additionally, only the general factor (H = .96) met the .7 cutoff for H, while all specific factors had H values of .43 or less. This indicates that only the general factor represents a well-defined latent variable (Rodriguez et al., 2016). ECV of the general factor was .87 and PUC was .64, which indicates that the MASC can be interpreted as "primarily unidimensional" (Reise et al., 2013). Average relative parameter bias was 1%, indicating no concerns related to using the measure unidimensionally. Taken together, these results support the acceptance of the unidimensional model as the best fitting model. This model is demonstrated in Figure 1.

Measurement invariance

Using the unidimensional model, measurement invariance analyses were conducted across the four racial groups. The configural model, in which the factor structures are constrained to be equivalent, demonstrated good fit according to all fit indices except the SRMR ($\chi^2(3780) = 4010.55$, $p < .01$; CFI = .96, TLI = .95, RMSEA = .01, SRMR = .10). Therefore, following the recommendations of Rutkowski and Svetina (2014), given an SRMR below .11 and acceptable

fit on all overall fit indices, the configural model was considered invariant. The metric model, in which factor loadings are constrained to be equivalent, also demonstrated good fit according to all fit indices except the SRMR ($\chi^2(3812) = 4171.20$, $p < .01$; CFI = .95, TLI = .95, RMSEA = .02, SRMR = .11). Additionally, the metric model did not fit significantly worse than the configural model ($\Delta\chi^2(132) = 167.64$, $p < .01$; ΔCFI = .005, ΔRMSEA = .001). Therefore, the metric model was found to be invariant. The scalar model also demonstrated good fit on all indices except the SRMR ($\chi^2(4044) = 4339.78$, $p < .01$; CFI = .94, TLI = .94, RMSEA = .02, SRMR = .11). The scalar model did not fit significantly worse than the metric model ($\Delta\chi^2(132) = 236.62$, $p < .01$; ΔCFI = .007, ΔRMSEA = 0). Therefore, the metric model was considered invariant and the MASC was found to have full measurement invariance across racial groups. As such, aim 3 was not investigated.

Mean comparisons

Given the findings of invariance, mean differences in MASC scores can be examined cross-racially without adjustment. Results of these analyses are presented in Table 2. An ANOVA revealed a small but significant difference in number of "correct" answers selected on the MASC racial group ($F(3,1226) = 6.15$, $p < .01$, $\eta^2 = .02$). Follow up tests using Tukey's HSD revealed that non-Hispanic White individuals ($M = 35.05$, $SD = 4.94$) were significantly more likely to choose the "correct" answer on MASC items than Black individuals ($M = 32.31$, $SD = 6.27$, $p = .01$), Asian/Pacific Islander individuals ($M = 32.23$, $SD = 5.74$, $p < .01$), and Hispanic White individuals ($M = 32.90$, $SD = 6.30$, $p = .04$). No other between-group differences were significant (all p's > .3).

To further investigate this mean difference, differences in error scores were also examined across racial groups. Regarding hypermentalizing errors, a small but significant difference was also found ($F(3,1226) = 6.15$, $p < .01$, $\eta^2 = .01$); Asian/Pacific Islander individuals ($M = 6.18$, $SD = 3.11$) were significantly more likely to choose the hypermentalizing

response than non-Hispanic White individuals (M = 5.42, SD = 2.94, p < .01) and Hispanic White individuals (M = 5.50, SD = 3.23, p = .01), but no other between-group differences were significant, including those with Black individuals (M = 5.44, SD = 2.98). Hypomentalizing errors also demonstrated a small but significant difference (F(3,1226) = 5.94, p < .01, η^2 = .01); non-Hispanic White individuals (M = 3.66, SD = 2.30) were less likely to select the hypomentalizing response than Hispanic White individuals (M = 4.34, SD = 2.63, p < .01), Black individuals (M = 4.54, SD = 3.17, p < .01), and Asian/Pacific Islander individuals (M = 4.36, SD = 2.74, p < .01). Finally, the "no mentalizing" response also revealed a small but significant difference based on racial group (F(3,1226) = 5.36, p < .01, η^2 = .01); Black individuals (M = 2.70, SD = 2.45) were more likely to select this response than non-Hispanic White individuals (M = 1.87, SD = 1.80, p < .01), but no other between group comparisons were significant, including those with Hispanic White individuals (M = 2.26, SD = 2.37) and Asian/Pacific Islander individuals (M = 2.23, SD = 2.05).

Discussion

Review and discussion of main findings

The aims of the study were to (1) investigate the factor structure of the MASC, (2) evaluate the possibility of bias based on race and ethnicity in the MASC, (3) examine the influence of racial ingroup effects and cultural norms on any variance in MASC performance across race, and (4) evaluate cross-racial differences in mentalizing after adjusting for sources of bias. Because measurement invariance was found in aim 2, it was not possible to address aim 3. Aims were investigated using CFA, MI analysis, and ANOVA, respectively.

Factor structure. Our hypotheses regarding the factor structure of the MASC were that a unidimensional model would demonstrate a stronger fit than a bi-factor or 3-factor model. Results support this hypothesized unidimensionality; while all models demonstrated acceptable and equivalent fit according to fit indices, follow-up analyses using bi-factor reliability indicators

revealed that the three specific factors did not measure reliable latent variables and did not add any significant variance above the general factor. This finding of unidimensionality is in line with the previous factor analytic study conducted on the MASC (Fossati et al., 2018). Unidimensionality indicates that there is a core construct being measured by the MASC and suggests that mentalizing itself can be considered a single unified construct. When taken together with mentalizing theory, these results support a unidimensional conceptualization of mentalizing, although this unidimensionality may not be generalizable to clinical samples.

According to some theories of mentalizing, optimal mentalizing means that the dimensions are integrated across their poles, while disordered mentalizing relies on one or the other extreme in isolation (Sharp & Vanwoerden, 2015). Importantly, both the present study and the previous factor analytic study, factor analyses were only conducted in non-clinical samples, while many studies using subscale scores are with clinical samples. It is possible that the unidimensional model holds due to the prevalence of optimal mentalizing, and therefore of integration across mentalizing dimensions in non-clinical samples. However, as suboptimal mentalization has been linked to psychopathology (McLaren et al., 2022), a lack of integration across the poles may be more common in a clinical sample. Such a lack of integration may lead to distinct, separable factors corresponding to each pole. Therefore, replication of these factor analytic studies in clinical samples is necessary to confirm the generalizability of a unidimensional model. Specifically, replication is samples with schizophrenia and schizoaffective disorders is especially important, as studies using this population frequently use subscale scores and occasionally find differences based on mentalizing of thoughts, feelings, and intentions (e.g. Vaskinn et al., 2018; Wastler & Lenzenweger, 2021).

Racial bias. We hypothesized that measurement invariance would not hold for the MASC across racial groups. However, measurement invariance analyses support racial invariance on the MASC; when factor structures, loadings, and intercepts were sequentially

constrained to be equal across racial groups, each model did not fit significantly worse than the previous. This indicates that there are no items which function differentially across racial groups and that the unidimensional factor structure holds regardless of racial group membership. Therefore, our hypothesis was not supported, and, in fact, we found evidence that the MASC items do not function differently across racial group. As such, we did not find evidence of racial bias on the MASC, which indicates that means can be compared across groups. Therefore, differences between means indicate true differences in the construct being measured rather than measurement error as would be detected by measurement invariance. As such, we conducted additional analyses to compare performance on the MASC cross-racially. Findings and possible implications of these mean comparisons are discussed in the following section; however, limitations of MI analyses in detecting bias diminish the strength of these findings.

While MI analysis is able to detect bias at the item-level, it does not account for sources of bias which may impact all items, otherwise known as common method variance (CMV; Steenkamp & Maydeu-Olivares, 2021). The proposed sources of bias (racial in-group effects and cultural norms) may affect all items on the MASC rather than individual items. If this is the case, traditional MI analyses performed here would not detect this bias on the MASC. A paradigm has been proposed by Steenkamp and Maydeu-Olivares to examine the possibility of common method variance (2021). Future research using these methods is required to fully investigate the possibility of racial bias on the MASC. Until such research is conducted, cross-racial analysis of the MASC should be interpreted with caution.

Differential performance based on race. We hypothesized that cross-racial differences would not exist after controlling for sources of variance. However, results from ANOVA indicated that non-Hispanic White individuals more frequently selected the “correct” answer on the MASC than individuals from other racial groups. Analysis of the error scores revealed that this pattern was mirrored in hypomentalizing responses; non-Hispanic White individuals were

less likely to select the answer choice indicating reduced mentalizing than other racial groups. However, differential patterns appeared for hypermentalizing and no mentalizing responses; Asian/Pacific Islander individuals were more likely to select the hypermentalizing response than both Hispanic and non-Hispanic White individuals, while Black individuals were more likely than non-Hispanic White individuals to select the response corresponding to no mentalizing. Findings of measurement invariance provide support for the idea that these differences in MASC score indicate true differences in mentalizing, but, as previously stated, results should be interpreted with caution given the limitations of measurement invariance analysis. It is possible that these differences instead result from CMV in the MASC. The following discussion poses one possible explanation for the existence of cross-racial mean differences in mentalizing as measured by the MASC, followed by a discussion of racial in-group effect and cultural differences as potential sources of CMV in MASC scores undetected by MI analyses.

If indeed these results reflect true differences in mentalizing across racial groups, we must clarify the significance of such a difference. While MASC scores are often referred to as "mentalizing ability," we argue that it is more accurate to refer to these scores as "mentalizing tendency" and recommend that others do the same. The MASC assesses what strategies of mentalizing individuals actually use in social contexts, rather than what strategies of mentalizing they possess the ability to use. While one's ability certainly plays a role in one's mentalizing tendency at an individual level, cross-racial differences in mentalizing tendency are not influenced by differences in ability. Instead, cross-racial differences in mentalizing ability may reflect learned differences in strategies for mentalizing.

"Optimal" mentalizing is only optimal if it leads to desired social outcomes. Through social learning, individuals create mental representations of how frequently and in which contexts various strategies for mentalizing lead to desired social outcomes. It follows that the frequency in which they use different mentalizing strategies corresponds to the frequency at

which individuals have found those strategies to be effective. Therefore, differential performance on the MASC may indicate a difference in the likelihood that different strategies of mentalizing have been found to be optimal for an individual. Cross-racial differences in optimal mentalizing strategy, therefore, may indicate that for different racial groups, mentalizing strategies are differentially effective. This possibility introduces a need for future research investigating whether mentalizing tendency is differentially associated with social outcomes based on one's race.

Possible mechanisms for a cross-racial difference in optimal mentalizing include differential treatment based on one's race in America, differential cultural expectations associated with different racial groups, and a combination of the two. The distinct patterns of mentalizing strategy for different racial groups (hypermentalizing and hypomentalizing for Asian/Pacific Islander individuals, hypomentalizing and no mentalizing for Black/African American individuals, and only hypomentalizing for Hispanic White individuals) indicate a contribution of unique cultural factors for each racial group. Exposure to trauma, especially interpersonal trauma such as racial trauma and chronic attachment trauma such as adverse childhood events, has long-term effects on one's mentalizing strategies (Bateman & Fonagy, 2012). In traumatic situations, the amygdala is activated, leading to hypervigilance and hypermentalizing; however, in the case of early or long-term exposure to trauma, emotional response is blunted, which is associated with reduced mentalizing (Wagner-Skacel et al., 2022; Williams et al., 2006). While Asian American individuals are exposed to less overall trauma than other racial groups and no difference in overall trauma exposure has been found between non-Hispanic White, Hispanic White, and African American individuals, differences in the prevalence types of trauma exist across racial groups (Alegría et al., 2013). Overall, members of racial and ethnic minority groups in America are uniquely exposed to racial trauma (Williams et al., 2018). Hispanic and African American individuals are additionally more likely to be exposed

to adverse childhood events (Sacks & Murphey, 2018), while Asian American individuals are more likely to be exposed to political violence (Alegría et al., 2013). This increased exposure to interpersonal trauma may help explain the hypermentalization increase with Asian/Pacific Islander individuals, while increased exposure to early and long-term trauma may help explain reduced mentalization associated with all Hispanic and African American individuals. This possibility warrants future research investigating the role of racial trauma and adverse childhood events in cross-racial mentalizing differences. While racial trauma may contribute to differences in mentalizing cross-racially, it is important to consider the contributions of cultural factors in mentalizing differences.

Preliminary research on cross-cultural mentalizing values has indicated that mentalizing values differ cross-culturally (Aival-Naveh et al., 2022); however, this work has not yet extended to America or to racial groups. Additionally, this work still attempts to apply existing frameworks of mentalizing to cultures in which they may not be applicable (Jurist & Perez Sosa, 2019); rather than examining cross-cultural differences in beliefs and values surrounding mentalizing, research has examined the extent to which different cultures value mentalizing as conceptualized in existing frameworks. Given the limited research in this field, insufficient information is available to make hypotheses about the nature of cross-cultural differences in mentalizing values and norms that may contribute to differences in mentalizing strategy. Future research investigating cross-cultural differences in values and beliefs regarding mentalizing is necessary to determine the mechanism by which cultural differences contribute to differences in mentalizing strategy. This line of research would greatly benefit from qualitative inquiry given the complexity of both mentalizing and cultural values and the expertise of individuals in understanding their own cultural experience.

Given the limitations of measurement invariance in detecting CMV, it is important to consider possible sources of CMV that may explain cross-racial differences. The two

hypothesized sources of cross-racial variance discussed previously—racial in-group effects and cultural norms—may represent sources of CMV rather than individual item-level bias. Regarding racial in-group effects, because the racial makeup of the characters is consistent throughout the MASC, it is likely that impacts of racial in-group effects would be apparent across every item, rather than in individual items. Regarding cross-cultural differences in mentalizing norms, such differences indicate that mentalizing is inextricably linked to social norms and beliefs. While traditional measurement invariance would reveal individual items that rely on familiarity with cultural norms, cross-cultural differences in mentalizing norms may systemically impact every item by influencing how familiar an individual is with the social situation, behaviors, and mental states depicted in the MASC. Therefore, rather than a handful of items exhibiting bias against individuals unfamiliar with a relevant cultural norm, there may be variance inherent in each item which reflects the degree to which one is "fluent" in the culture of those being mentalized. This distinction can be likened to the difference between taking a test in one's second language (a source of CMV) the inclusion of references to an unfamiliar concept in an individual question (a source of individual item variance); while racial and ethnic minority groups may be familiar with White American and European mentalizing norms, these may not be the norms with which they have the most exposure, experience, or fluency. As such, CMV analysis is vital to rule out the possibility of CMV due to racial in-group effects and cultural norms before moving forward with the conclusion that the MASC is not racially biased.

General limitations and future directions

Possible covariates in the sample demographic characteristics of the current study may limit the of its results. First, all participants in the current study are university students. Higher levels of education are associated with increased mentalizing ability (Li et al., 2013). Therefore, the results of the current study may not be generalizable to individuals who do not attend college. Additionally, the majority of participants in the current sample (about 80%) were women. MASC

performance has been found to differ by gender; women tend to select more "correct" answers and fewer "hypermentalizing" answers on the MASC than men in community samples, general clinical samples, and individuals with ASD (Abate et al., 2017; Müller et al., 2016; Poznyak et al., 2019; Wacker et al., 2017). Specifically, both men and women performed similarly regarding the mental states of characters whose gender did not correspond to their own, while women scored significantly higher than men regarding the mental states of characters of the same gender as their own (Wacker et al., 2017). A gender difference in MASC performance may mean that results from a sample of predominantly women cannot be generalized to other genders. However, measurement invariance of gender in the MASC has not yet been investigated and the interaction between gender and race in MASC performance is unknown. As such, it is difficult to predict the potential effect of this gender imbalance on the results of the current study. We recommend that future research on cross-gender measurement invariance and the interaction between gender and race in mentalizing tendency be conducted.

Additionally, the current study did not collect sufficient demographic information to consider the impact of socioeconomic status (SES) or verbal ability as potential covariates. SES has an impact on one's neural activity while mentalizing; specifically, individuals with lower SES demonstrate greater activation in the mentalizing network and in the amygdala when processing threatening faces (Muscatell et al., 2012). This suggests that hypervigilance and hypermentalizing may be negatively associated with SES. Thus, it is possible that cross-racial differences in socioeconomic status influence the relationship between race and mentalizing tendency. Verbal ability is additionally a predictor of mentalizing development (Astington & Baird, 2005). This is especially relevant on the MASC, a task which relies on an individual's verbal and reading comprehension ability. Thus, future studies which examine the influence of these covariates on the relationship between race and mentalizing are necessary.

Finally, racial demographic information collection differed between the two included studies; data collected for the first study does not follow current guidelines for collecting data regarding race and ethnicity. As a result, in this study, participants who identified as Hispanic were not able to select a racial group. It is therefore likely that some individuals included in the Hispanic White group from this study are non-White. However, we expect that this is a small number; in the second study, which collected from the same pool of university students, only 6 individuals were excluded due to identifying as both Hispanic and either Black/African American or Asian/Pacific Islander. Therefore, we expect the impact that this limitation has on the conclusions of the current study to be negligible.

The results of this study have important implications for clinical work, specifically regarding the use of Mentalization Based Therapy (MBT) and education about mentalization from a public health standpoint. For both of these interventions, the results of the current study emphasize the importance of allowing for cultural and individual differences in "optimal mentalizing" rather than attempting to conform individuals to a universal standard of mentalizing. MBT specifically targets an individual's mentalizing tendency with the aim of improving psychological functioning by encouraging a "mentalizing stance" in both the clinician and the client. Notably, MBT has been primarily developed and implemented in White European populations (Vogt & Norman, 2019). Additionally, clinicians practicing MBT bring their own cultural background and understanding of "optimal mentalizing." Therefore, it is vital that when practicing MBT with racially and culturally diverse populations, clinicians remain client-led and curious regarding what optimal mentalizing means for each individual in each situation. When providing public health education regarding mentalizing, it is similarly important to partner with and be led by the communities in which the intervention is taking place. These interventions must be informed by mentalizing expectations specific to that community rather than enforcing a standard of mentalizing developed for another population. One such example of culturally

informed mentalization education has been proposed using media that is representative of social interactions in the queer community (Collins, 2021). Following this example, public education regarding mentalizing for racially diverse communities can be developed using media created by and for members of these communities, leading to interventions based on social norms and mentalizing expectations specific to the community in which they are being implemented.

Conclusions

The current study provides support for the unidimensionality and racial measurement invariance of the MASC. This indicates that mentalizing as measured by the MASC is a unified construct whose structure does not change based on racial group membership. Non-Hispanic White individuals are more likely to select the "correct" answer on the MASC than other racial groups, which may indicate that the optimal mentalizing strategy and mentalizing values and norms differ based on racial group membership. However, limitations of MI impact the strength of these conclusions; CMV is not detected using traditional MI methods. Because racial in-group effects and cultural norms may represent sources of CMV, these results should be interpreted with caution and indicate a need to more closely evaluate the possibility of CMV on the MASC.

References

Abate, A., Marshall, K., Sharp, C., & Venta, A. (2017). Trauma and Aggression: Investigating the Mediating Role of Mentalizing in Female and Male Inpatient Adolescents. *Child Psychiatry & Human Development*, *48*(6), 881–890. https://doi.org/10.1007/s10578-017-0711-6

Adams, R. B., Rule, N. O., Franklin, R. G., Wang, E., Stevenson, M. T., Yoshikawa, S., Nomura, M., Sato, W., Kveraga, K., & Ambady, N. (2009). Cross-cultural Reading the Mind in the Eyes: An fMRI Investigation. *Journal of Cognitive Neuroscience*, *22*(1), 97–108. https://doi.org/10.1162/jocn.2009.21187

Aival-Naveh, E., Rothschild-Yakar, L., & Kurman, J. (2019). Keeping culture in mind: A systematic review and initial conceptualization of mentalizing from a cross-cultural perspective. *Clinical Psychology: Science and Practice*, *26*(4), e12300. https://doi.org/10.1111/cpsp.12300

Aival-Naveh, E., Rothschild-Yakar, L., Park, J., & Kurman, J. (2022). The Value of Thinking About Feelings Across Cultures: A Preliminary Investigation of the Mentalizing Values Scale. *Journal of Cross-Cultural Psychology*, *53*(3–4), 362–379. https://doi.org/10.1177/00220221221077355

Alegría, M., Fortuna, L. R., Lin, J. Y., Norris, L. F., Gao, S., Takeuchi, D. T., Jackson, J. S., Shrout, P. E., & Valentine, A. (2013). Prevalence, Risk, and Correlates of Posttraumatic Stress Disorder across Ethnic and Racial Minority Groups in the U.S. *Medical Care*, *51*(12), 1114–1123. https://doi.org/10.1097/MLR.0000000000000007

Allen, J., Fonagy, P., & Bateman, A. W. (2008). *Mentalizing in Clinical Practice*. American Psychiatric Pub.

Anderson, J. C., & Gerbing, D. W. (1988). Structural Equation Modeling in Practice: A Review and Recommended Two-Step Approach. *Psychological Bulletin*, *103*(3), 411–423.

Avenanti, A., Sirigu, A., & Aglioti, S. M. (2010). Racial Bias Reduces Empathic Sensorimotor Resonance with Other-Race Pain. *Current Biology*, *20*(11), 1018–1022. https://doi.org/10.1016/j.cub.2010.03.071

Babcock, J. C., & Banks, J. C. (2019). Interobserver agreement and the effects of ethnicity on observational coding of affect. *Journal of Social and Personal Relationships*, *36*(9), 2842–2856. https://doi.org/10.1177/0265407518803474

Bar-Haim, Y., Ziv, T., Lamy, D., & Hodes, R. M. (2006). Nature and Nurture in Own-Race Face Processing. *Psychological Science*, *17*(2), 159–163. https://doi.org/10.1111/j.1467-9280.2006.01679.x

Baron-Cohen, S. (1996). Reading the Mind in the Face: A Cross-cultural and Developmental Study. *Visual Cognition*, *3*(1), 39–60. https://doi.org/10.1080/713756728

Bateman, A. W., & Fonagy, P. (2012). *Handbook of mentalizing in mental health practice* (A. W. Bateman & P. Fonagy, Eds.; 2011-19854-000). American Psychiatric Publishing, Inc.; psyh. http://ezproxy.lib.uh.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2011-19854-000&site=ehost-live

Beauducel, A., & Herzberg, P. Y. (2006). On the Performance of Maximum Likelihood Versus Means and Variance Adjusted Weighted Least Squares Estimation in CFA. *Structural Equation Modeling: A Multidisciplinary Journal*, *13*(2), 186–203. https://doi.org/10.1207/s15328007sem1302_2

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*(2), 238–246.

Blatt, B., LeLacheur, S. F., Galinsky, A. D., Simmens, S. J., & Greenberg, L. (2010). Does Perspective-Taking Increase Patient Satisfaction in Medical Encounters? *Academic Medicine*, *85*(9), 1445–1452. https://doi.org/10.1097/ACM.0b013e3181eae5ec

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In *Testing Structural Equation Models* (pp. 136–162). SAGE.

Buhlmann, U., Wacker, R., & Dziobek, I. (2015). Inferring other people's states of mind: Comparison across social anxiety, body dysmorphic, and obsessive–compulsive disorders. *Journal of Anxiety Disorders*, *34*, 107–113. https://doi.org/10.1016/j.janxdis.2015.06.003

Camoirano, A. (2017). Mentalizing Makes Parenting Work: A Review about Parental Reflective Functioning and Clinical Interventions to Improve It. *Frontiers in Psychology*, *8*, 14. https://doi.org/10.3389/fpsyg.2017.00014

Caputi, M., Lecce, S., Pagnin, A., & Banerjee, R. (2012). Longitudinal effects of theory of mind on later peer relations: The role of prosocial behavior. *Developmental Psychology*, *48*(1), 257–270. https://doi.org/10.1037/a0025402

Chen, F. F. (2007). Sensitivity of Goodness of Fit Indexes to Lack of Measurement Invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(3), 464–504. https://doi.org/10.1080/10705510701301834

Clark, D. A., & Bowles, R. P. (2018). Model Fit and Item Factor Analysis: Overfactoring, Underfactoring, and a Program to Guide Interpretation. *Multivariate Behavioral Research*, *53*(4), 544–558. https://doi.org/10.1080/00273171.2018.1461058

Collins, N. (2021). On RuPaul, drag race and mentalization. *Australasian Psychiatry*, *29*(6), 611–613. https://doi.org/10.1177/10398562211022754

Divilbiss, M. (2009). *The role of Theory of Mind as a mediator in the relationship between social functioning and schizotypy* (304911687) [M.A., Kent State University]. ProQuest Dissertations & Theses Global.

Dodell-Feder, D., Felix, S., Yung, M. G., & Hooker, C. I. (2016). Theory-of-mind-related neural activity for one's romantic partner predicts partner well-being. *Social Cognitive and Affective Neuroscience*, *11*(4), 593–603. https://doi.org/10.1093/scan/nsv144

Dodell-Feder, D., Ressler, K. J., & Germine, L. T. (2020). Social cognition or social class and culture? On the interpretation of differences in social cognitive performance. *Psychological Medicine*, *50*(1), 133–145. https://doi.org/10.1017/S003329171800404X

Dueber, D. (2017). Bifactor Indices Calculator: A Microsoft Excel-Based Tool to Calculate Various Indices Relevant to Bifactor CFA Models. *Educational, School, and Counseling Psychology Research Tools*. https://doi.org/10.13023/edp.tool.01

Duval, J., Ensink, K., Normandin, L., Sharp, C., & Fonagy, P. (2018). Measuring reflective functioning in adolescents: Relations to personality disorders and psychological difficulties. *Adolescent Psychiatry*, *8*(1), 5–20. psyh. https://doi.org/10.2174/2210676608666180208161619

Dziobek, I., Fleck, S., Kalbe, E., Rogers, K., Hassenstab, J., Brand, M., Kessler, J., Woike, J. K., Wolf, O. T., & Convit, A. (2006). Introducing MASC: A Movie for the Assessment of Social Cognition.

Journal of Autism and Developmental Disorders, *36*(5), 623–636. https://doi.org/10.1007/s10803-006-0107-0

Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, *17*(2), 124–129. https://doi.org/10.1037/h0030377

Elfenbein, H. A., & Ambady, N. (2003). When Familiarity Breeds Accuracy: Cultural Exposure and Facial Emotion Recognition. *Journal of Personality and Social Psychology*, 434–436.

Fan, X., Thompson, B., & Wang, L. (1999). Effects of sample size, estimation methods, and model specification on structural equation modeling fit indexes. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 56–83. https://doi.org/10.1080/10705519909540119

Fink, E., Begeer, S., Peterson, C. C., Slaughter, V., & Rosnay, M. de. (2015). Friendlessness and theory of mind: A prospective longitudinal study. *British Journal of Developmental Psychology*, *33*(1), 1–17. https://doi.org/10.1111/bjdp.12060

Fossati, A., Borroni, S., Dziobek, I., Fonagy, P., & Somma, A. (2018). Thinking about assessment: Further evidence of the validity of the Movie for the Assessment of Social Cognition as a measure of mentalistic abilities. *Psychoanalytic Psychology*, *35*(1), 127–141. https://doi.org/10.1037/pap0000130

Friesen, W. (1972). *Cultural differences in facial expression in a social situation: An experimental test of the concept of display rules*. https://www.proquest.com/openview/2318df49e2dc7546618305634b5a6c3d/1.pdf/advanced

Goldstein, N. J., Vezich, I. S., & Shapiro, J. R. (2014). Perceived perspective taking: When others walk in our shoes. *Journal of Personality and Social Psychology*, *106*(6), 941–960. https://doi.org/10.1037/a0036395

Gorsuch, R. L. (1983). Factor Analysis, 2nd edn Lawrence Erlbaum Associates. *Hillsdale, NJ*.

Hancock, G., & Mueller, R. O. (2001). Rethinking construct reliability within latent variable systems. *Structural Equation Modeling: Present and Future*, 195–216.

Hart, J. R., Venta, A., & Sharp, C. (2017). Attachment and thought problems in an adolescent inpatient sample: The mediational role of theory of mind. *Comprehensive Psychiatry*, *78*, 38–47. https://doi.org/10.1016/j.comppsych.2017.07.002

Hassenstab, J., Dziobek, I., Rogers, K., Wolf, O. T., & Convit, A. (2007). Knowing What Others Know, Feeling What Others Feel: A Controlled Study of Empathy in Psychotherapists. *The Journal of Nervous and Mental Disease*, *195*(4), 277. https://doi.org/10.1097/01.nmd.0000253794.74540.2d

Hermans, H., & Kempen, H. (1998). Moving cultures—The perilous problems of cultural dichotomies in a globalizing society. *American Psychologist*, *53*, 1111–1120. https://doi.org/10.1037//0003-066X.53.10.1111

Hornstein, H. A. (1978). Promotive tension and prosocial behavior: A Lewinian analysis. *Altruism, Sympathy, and Helping: Psychological and Sociological Principles*, 177–207.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

IBM SPSS Statistics for Windows (Version 25). (2017). IBM Corp.

Imuta, K., Henry, J. D., Slaughter, V., Selcuk, B., & Ruffman, T. (2016). Theory of mind and prosocial behavior in childhood: A meta-analytic review. *Developmental Psychology*, *52*(8), 1192–1205. https://doi.org/10.1037/dev0000140

Jenkins, J. M., & Astington, J. W. (1996). Cognitive factors and family structure associated with theory of mind development in young children. *Developmental Psychology*, *32*(1), 70–78. https://doi.org/10.1037/0012-1649.32.1.70

Johnson, J. D., Simmons, C. H., Jordav, A., Maclean, L., Taddei, J., Thomas, D., Dovidio, J. F., & Reed, W. (2002). Rodney King and O. J. Revisited: The Impact of Race and Defendant Empathy Induction on Judicial Decisions. *Journal of Applied Social Psychology*, *32*(6), 1208–1223. https://doi.org/10.1111/j.1559-1816.2002.tb01432.x

Jurist, E., & Perez Sosa, M. (2019). Commentary on mentalization and culture. *Clinical Psychology: Science and Practice*, *26*. https://doi.org/10.1111/cpsp.12302

Kim, H. S., Shin, N. Y., Jang, J. H., Kim, E., Shim, G., Park, H. Y., Hong, K. S., & Kwon, J. S. (2011). Social cognition and neurocognition as predictors of conversion to psychosis in individuals at ultra-high risk. *Schizophrenia Research*, *130*(1), 170–175. https://doi.org/10.1016/j.schres.2011.04.023

Kline, R. B. (2015). *Principles and Practice of Structural Equation Modeling, Fourth Edition*. Guilford Publications.

Kobayashi, C., Glover, G. H., & Temple, E. (2006). Cultural and linguistic influence on neural bases of 'Theory of Mind': An fMRI study with Japanese bilinguals. *Brain and Language*, *98*(2), 210–220. https://doi.org/10.1016/j.bandl.2006.04.013

Koelkebeck, K., Hirao, K., Kawada, R., Miyata, J., Saze, T., Ubukata, S., Itakura, S., Kanakogi, Y., Ohrmann, P., Bauer, J., Pedersen, A., Sawamoto, N., Fukuyama, H., Takahashi, H., & Murai, T. (2011). Transcultural differences in brain activation patterns during theory of mind (ToM) task performance in Japanese and Caucasian participants. *Social Neuroscience*, *6*(5–6), 615–626. https://doi.org/10.1080/17470919.2011.620763

Lahera, G., Boada, L., Pousa, E., Mirapeix, I., Morón-Nozaleda, G., Marinas, L., Gisbert, L., Pamiàs, M., & Parellada, M. (2014). Movie for the Assessment of Social Cognition (MASC): Spanish Validation. *Journal of Autism and Developmental Disorders*, *44*(8), 1886–1896. https://doi.org/10.1007/s10803-014-2061-6

Lecce, S., Ceccato, I., & Cavallini, E. (2018). Investigating ToM in aging with the MASC: From accuracy to error type. *Aging, Neuropsychology, and Cognition*, *0*(0), 1–17. https://doi.org/10.1080/13825585.2018.1500996

Lecours, S., & Bouchard, M.-A. (1997). Dimensions of mentalisation: Outlining levels of psychic transformation. *International Journal of Psycho-Analysis*, *78*(1997), 855–875.

Levin, D. (1996). Classifying Faces by Race: The Structure of Face Categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 1364–1382. https://doi.org/10.1037/0278-7393.22.6.1364

Li, X., Wang, K., Wang, F., Tao, Q., Xie, Y., & Cheng, Q. (2013). Aging of theory of mind: The influence of educational level and cognitive processing. *International Journal of Psychology*, *48*(4), 715–727. https://doi.org/10.1080/00207594.2012.673724

Liew, S., Han, S., & Aziz‐Zadeh, L. (2010). Familiarity modulates mirror neuron and mentalizing regions during intention understanding. *Human Brain Mapping*, *32*(11), 1986–1997. https://doi.org/10.1002/hbm.21164

Luyten, P., Fonagy, P., Lowyck, B., & Vermote, R. (2012). Assessment of mentalization. In A. W. Bateman & P. Fonagy (Eds.), *Handbook of mentalizing in mental health practice.* (2011-19854-002; pp. 43–65). American Psychiatric Publishing, Inc.; psyh. http://ezproxy.lib.uh.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2011-19854-002&site=ehost-live

Martinez, G., Alexandre, C., Mam-Lam-Fook, C., Bendjemaa, N., Gaillard, R., Garel, P., Dziobek, I., Amado, I., & Krebs, M.-O. (2017). Phenotypic continuum between autism and schizophrenia: Evidence from the Movie for the Assessment of Social Cognition (MASC). *Schizophrenia Research*, *185*, 161–166. https://doi.org/10.1016/j.schres.2017.01.012

Marty, P., & M'Uzan, D. (1963). La pensée opératoire. *Revue Francaise de Pschanalyse*, *27*, 1345–1356.

Masuda, T., Ellsworth, P. C., Mesquita, B., Leu, J., Tanida, S., & Van de Veerdonk, E. (2008). Placing the face in context: Cultural differences in the perception of facial emotion. *Journal of Personality and Social Psychology*, *94*(3), 365–381. https://doi.org/10.1037/0022-3514.94.3.365

Mathur, V. A., Harada, T., Lipke, T., & Chiao, J. Y. (2010). Neural basis of extraordinary empathy and altruistic motivation. *NeuroImage*, *51*(4), 1468–1475. https://doi.org/10.1016/j.neuroimage.2010.03.025

Matsumoto, D., & Ekman, P. (1989). American-Japanese cultural differences in intensity ratings of facial expressions of emotion. *Motivation and Emotion*, *13*(2), 143–157. https://doi.org/10.1007/BF00992959

Matsumoto, D., Takeuchi, S., Andayani, S., Kouznetsova, N., & Krupp, D. (1998). The Contribution of Individualism vs. Collectivism to Cross‐national Differences in Display Rules. *Asian Journal of Social Psychology*, *1*(2), 147–165. https://doi.org/10.1111/1467-839X.00010

Matsumoto, D., Willingham, B., & Olide, A. (2009). Sequential Dynamics of Culturally Moderated Facial Expressions of Emotion. *Psychological Science*, *20*(10), 1269–1274. https://doi.org/10.1111/j.1467-9280.2009.02438.x

Mayer, S. V., Jusyte, A., Klimecki-Lenz, O. M., & Schönenberg, M. (2018). Empathy and altruistic behavior in antisocial violent offenders with psychopathic traits. *Psychiatry Research*, *269*, 625–632. https://doi.org/10.1016/j.psychres.2018.08.035

McDonald, S., Flanagan, S., Rollins, J., & Kinch, J. (2003). TASIT: A New Clinical Tool for Assessing Social Perception After Traumatic Brain Injury. *The Journal of Head Trauma Rehabilitation*, *18*(3), 219–238.

McLaren, V., Gallagher, M., Hopwood, C. J., & Sharp, C. (2022). Hypermentalizing and Borderline Personality Disorder: A meta-analytic review. *American Journal of Psychotherapy*, *75*(1).

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

Mellenbergh, G. J. (1989). Item bias and item response theory. *International Journal of Educational Research*, *13*(2), 127–143. https://doi.org/10.1016/0883-0355(89)90002-5

Michel, C., Rossion, B., Han, J., Chung, C.-S., & Caldara, R. (2006). Holistic Processing Is Finely Tuned for Faces of One's Own Race. *Psychological Science*, *17*(7), 608–615. https://doi.org/10.1111/j.1467-9280.2006.01752.x

Montag, C., Dziobek, I., Richter, I. S., Neuhaus, K., Lehmann, A., Sylla, R., Heekeren, H. R., Heinz, A., & Gallinat, J. (2011). Different aspects of theory of mind in paranoid schizophrenia: Evidence from a video-based assessment. *Psychiatry Research*, *186*(2), 203–209. https://doi.org/10.1016/j.psychres.2010.09.006

Müller, N., Baumeister, S., Dziobek, I., Banaschewski, T., & Poustka, L. (2016). Validation of the Movie for the Assessment of Social Cognition in Adolescents with ASD: Fixation Duration and Pupil Dilation as Predictors of Performance. *Journal of Autism and Developmental Disorders*, *46*(9), 2831–2844. https://doi.org/10.1007/s10803-016-2828-z

Muscatell, K. A., Morelli, S. A., Falk, E. B., Way, B. M., Pfeifer, J. H., Galinsky, A. D., Lieberman, M. D., Dapretto, M., & Eisenberger, N. I. (2012). Social status modulates neural activity in the mentalizing network. *NeuroImage*, *60*(3), 1771–1777. https://doi.org/10.1016/j.neuroimage.2012.01.080

Muthén, B., Kaplan, D., & Hollis, M. (1987). On structural equation modeling with data that are not missing completely at random. *Psychometrika*, *52*(3), 431–462.

Muthén, L. K., & Muthén, B. O. (2017). *Mplus User's Guide* (8th ed.). Muthén & Muthén. https://www.statmodel.com/orderonline/products.php?product=Mplus-Version-8-User%27s-Guide

Oehlman Forbes, D., Lee, M., & Lakeman, R. (2020). Psychotherapy The Role of Mentalization in Child Psychotherapy, Interpersonal Trauma, and Recovery: A Scoping Review. *Psychotherapy*, *58*. https://doi.org/10.1037/pst0000341

Paladino, M.-P., Leyens, J.-P., Rodriguez, R., Rodriguez, A., Gaunt, R., & Demoulin, S. (2002). Differential Association of Uniquely and Non Uniquely Human Emotions with the Ingroup and the Outgroup. *Group Processes*, 13.

Park, S. G. (2010). *A simulated live interaction to examine behavioral correlates of social cognition in individuals with social anhedonia*. University of Maryland.

Perner, J., Ruffman, T., & Leekam, S. R. (1994). Theory of Mind Is Contagious: You Catch It from Your Sibs. *Child Development*, *65*(4), 1228–1238. https://doi.org/10.2307/1131316

Peterson, C. C., & Siegal, M. (1997). *Domain specificity and everyday biological, physical, and psychological thinking in normal, autistic, and deaf children.*

Pinkham, A. E., Kelsven, S., Kouros, C., Harvey, P. D., & Penn, D. L. (2017). The Effect of Age, Race, and Sex on Social Cognitive Performance in Individuals With Schizophrenia: *The Journal of Nervous and Mental Disease*, *205*(5), 346–352. https://doi.org/10.1097/NMD.0000000000000654

Poznyak, E., Morosan, L., Perroud, N., Speranza, M., Badoud, D., & Debbané, M. (2019). Roles of age, gender and psychological difficulties in adolescent mentalizing. *Journal of Adolescence*, *74*, 120–129. https://doi.org/10.1016/j.adolescence.2019.06.007

Preißler, S., Dziobek, I., Ritter, K., Heekeren, H. R., & Roepke, S. (2010). Social Cognition in Borderline Personality Disorder: Evidence for Disturbed Recognition of the Emotions, Thoughts, and Intentions of others. *Frontiers in Behavioral Neuroscience*, *4*. https://doi.org/10.3389/fnbeh.2010.00182

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, *1*(4), 515–526. https://doi.org/10.1017/S0140525X00076512

Putnick, D. L., & Bornstein, M. H. (2016). Measurement Invariance Conventions and Reporting: The State of the Art and Future Directions for Psychological Research. *Developmental Review : DR*, *41*, 71–90. https://doi.org/10.1016/j.dr.2016.06.004

Reise, S. P., Bonifay, W. E., & Haviland, M. G. (2013). Scoring and Modeling Psychological Measures in the Presence of Multidimensionality. *Journal of Personality Assessment*, *95*(2), 129–140. https://doi.org/10.1080/00223891.2012.725437

Repacholi, B., & Trapolini, T. (2004). Attachment and preschool children's understanding of maternal versus non-maternal psychological states. *British Journal of Developmental Psychology*, *22*(3), 395–415. https://doi.org/10.1348/0261510041552693

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychological Methods*, *21*(2), 137–150. https://doi.org/10.1037/met0000045

Rutkowski, L., & Svetina, D. (2014). Assessing the Hypothesis of Measurement Invariance in the Context of Large-Scale International Surveys. *Educational and Psychological Measurement*, *74*(1), 31–57. https://doi.org/10.1177/0013164413498257

Sacks, V., & Murphey, D. (2018). The prevalence of adverse childhood experiences, nationally, by state, and by race or ethnicity—Child Trends. *Child Trends*, *20*(2018). https://www.childtrends.org/publications/prevalence-adverse-childhood-experiences-nationally-state-race-ethnicity

Schneider, A., Johnston, C., Tassone, F., Sansone, S., Hagerman, R. J., Ferrer, E., Rivera, S. M., & Hessl, D. (2016). Broad autism spectrum and obsessive–compulsive symptoms in adults with the fragile X premutation. *The Clinical Neuropsychologist*, *30*(6), 929–943. https://doi.org/10.1080/13854046.2016.1189536

Shahaeian, A., Peterson, C. C., Slaughter, V., & Wellman, H. M. (2011). Culture and the sequence of steps in theory of mind development. *Developmental Psychology*, *47*(5), 1239–1247. https://doi.org/10.1037/a0023899

Sharp, C. (2016). Bridging the gap: The assessment and treatment of adolescent personality disorder in routine clinical care. *Archives of Disease in Childhood*, *102*(1), 103–108. https://doi.org/10.1136/archdischild-2015-310072

Sharp, C., & Fonagy, P. (2008). The Parent's Capacity to Treat the Child as a Psychological Agent: Constructs, Measures and Implications for Developmental Psychopathology. *Social Development*, *17*(3), 737–754. https://doi.org/10.1111/j.1467-9507.2007.00457.x

Sharp, C., & Vanwoerden, S. (2015). Hypermentalizing in Borderline Personality Disorder: A Model and Data. *Journal of Infant, Child, and Adolescent Psychotherapy*, *14*(1), 33–45. https://doi.org/10.1080/15289168.2015.1004890

Sharp, C., Venta, A., Vanwoerden, S., Schramm, A., Ha, C., Newlin, E., Reddy, R., & Fonagy, P. (2016). First empirical evaluation of the link between attachment, social cognition and borderline features in adolescents. *Comprehensive Psychiatry*, *64*, 4–11. https://doi.org/10.1016/j.comppsych.2015.07.008

Slaughter, V., Imuta, K., Peterson, C. C., & Henry, J. D. (2015). Meta-Analysis of Theory of Mind and Peer Popularity in the Preschool and Early School Years. *Child Development*, *86*(4), 1159–1174.

Slaughter, V., & Perez‐Zapata, D. (2014). Cultural Variations in the Development of Mind Reading. *Child Development Perspectives*, *8*(4), 237–241. https://doi.org/10.1111/cdep.12091

Steenkamp, J.-B. E. M., & Baumgartner, H. (2000). On the use of structural equation models for marketing modeling. *International Journal of Research in Marketing*, *17*(2), 195–202. https://doi.org/10.1016/S0167-8116(00)00016-1

Steenkamp, J.-B. E. M., & Maydeu-Olivares, A. (2021). An updated paradigm for evaluating measurement invariance incorporating common method variance and its assessment. *Journal of the Academy of Marketing Science*, *49*(1), 5–29. https://doi.org/10.1007/s11747-020-00745-z

Taubner, S., Müller, S., & Kotte, S. (2017). Mentalizing Vocational Training increases mentalization interest in professionals and young people with the need for learning support: Results of a pilot study. *Mental Health & Prevention*, *6*, 1–11. https://doi.org/10.1016/j.mhp.2017.02.001

Todd, A. R., Bodenhausen, G. V., Richeson, J. A., & Galinsky, A. D. (2011). Perspective taking combats automatic expressions of racial bias. *Journal of Personality and Social Psychology*, 1027–1042.

Todd, A. R., & Galinsky, A. D. (2014). Perspective‐taking as a strategy for improving intergroup relations: Evidence, mechanisms, and qualifications. *Social and Personality Psychology Compass*, *8*(7), 374–387.

Van De Schoot, R., Schmidt, P., De Beuckelaer, A., Lek, K., & Zondervan-Zwijnenburg, M. (2015). Editorial: Measurement Invariance. *Frontiers in Psychology*, *6*, 1064. https://doi.org/10.3389/fpsyg.2015.01064

Vaskinn, A., Andersson, S., Østefjells, T., Andreassen, O. A., & Sundet, K. (2018). Emotion perception, non-social cognition and symptoms as predictors of theory of mind in schizophrenia. *Comprehensive Psychiatry*, *85*, 1–7. https://doi.org/10.1016/j.comppsych.2018.05.002

Vignoles, V. L., Owe, E., Becker, M., Smith, P. B., Easterbrook, M. J., Brown, R., González, R., Didier, N., Carrasco, D., Cadena, M. P., Lay, S., Schwartz, S. J., Des Rosiers, S. E., Villamar, J. A., Gavreliuc, A., Zinkeng, M., Kreuzbauer, R., Baguma, P., Martin, M., … Bond, M. H. (2016). Beyond the 'east–west' dichotomy: Global variation in cultural models of selfhood. *Journal of Experimental Psychology: General*, *145*(8), 966–1000. https://doi.org/10.1037/xge0000175

Vogt, K. S., & Norman, P. (2019). Is mentalization-based therapy effective in treating the symptoms of borderline personality disorder? A systematic review. *Psychology and Psychotherapy: Theory, Research and Practice*, *92*(4), 441–464. https://doi.org/10.1111/papt.12194

Voronov, M., & Singer, J. A. (2002). The Myth of Individualism-Collectivism: A Critical Review. *The Journal of Social Psychology*, *142*(4), 461–480. https://doi.org/10.1080/00224540209603912

Wacker, R., Bölte, S., & Dziobek, I. (2017). Women Know Better What Other Women Think and Feel: Gender Effects on Mindreading across the Adult Life Span. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.01324

Wagner-Skacel, J., Riedl, D., Kampling, H., & Lampe, A. (2022). Mentalization and dissociation after adverse childhood experiences. *Scientific Reports*, *12*(1), Article 1. https://doi.org/10.1038/s41598-022-10787-8

Washburn, D., Wilson, G., Roes, M., Rnic, K., & Harkness, K. L. (2016). Theory of mind in social anxiety disorder, depression, and comorbid conditions. *Journal of Anxiety Disorders*, *37*, 71–77. https://doi.org/10.1016/j.janxdis.2015.11.004

Wastler, H. M., & Lenzenweger, M. F. (2019). Self-referential hypermentalization in schizotypy. *Personality Disorders: Theory, Research, and Treatment*, *10*(6), 536–544. https://doi.org/10.1037/per0000344

Wastler, H. M., & Lenzenweger, M. F. (2021). Cognitive and Affective Theory of Mind in Positive Schizotypy: Relationship to Schizotypal Traits and Psychosocial Functioning. *Journal of Personality Disorders*, *35*(4), 538–553. https://doi.org/10.1521/pedi_2020_34_473

Weijers, J., Fonagy, P., Eurelings-Bontekoe, E., Termorshuizen, F., Viechtbauer, W., & Selten, J. P. (2018). Mentalizing impairment as a mediator between reported childhood abuse and outcome in nonaffective psychotic disorder. *Psychiatry Research*, *259*, 463–469. https://doi.org/10.1016/j.psychres.2017.11.010

Wellman, H. M., Fang, F., Liu, D., Zhu, L., & Liu, G. (2006). Scaling of Theory-of-Mind Understandings in Chinese Children. *Psychological Science*, *17*(12), 1075–1081. https://doi.org/10.1111/j.1467-9280.2006.01830.x

Wilbertz, G., Brakemeier, E.-L., Zobel, I., Härter, M., & Schramm, E. (2010). Exploring preoperational features in chronic depression. *Journal of Affective Disorders*, *124*(3), 262–269. https://doi.org/10.1016/j.jad.2009.11.021

Williams, L. M., Liddell, B. J., Kemp, A. H., Bryant, R. A., Meares, R. A., Peduto, A. S., & Gordon, E. (2006). Amygdala–prefrontal dissociation of subliminal and supraliminal fear. *Human Brain Mapping*, *27*(8), 652–661.

Williams, M. T., Ching, T. H. W., Printz, D. M. B., & Wetterneck, C. T. (2018). Assessing PTSD in Ethnic and Racial Minorities: Trauma and Racial Trauma. *Directions in Psychiatry*, *38*(3), 179–196.

Xu, X., Zuo, X., Wang, X., & Han, S. (2009). Do You Feel My Pain? Racial Group Membership Modulates Empathic Neural Responses. *Journal of Neuroscience*, *29*(26), 8525–8529. https://doi.org/10.1523/JNEUROSCI.2418-09.2009

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

Zainal, N. H., & Newman, M. G. (2018). Worry amplifies theory-of-mind reasoning for negatively valenced social stimuli in generalized anxiety disorder. *Journal of Affective Disorders*, *227*, 824–833. https://doi.org/10.1016/j.jad.2017.11.084

Tables and Figures

Table 1. *Standardized factor loadings for MASC items in unidimensional, three-factor, and bi-factor CFA models*

	Unidimensional	**Three-Factor**	**Bifactor**	
Item	Std(SE)	Std(SE)	General, Std(SE)	Specific, Std(SE)
Thoughts				
6	**0.6(0.04)****	**0.61(0.04)****	**0.61(0.04)****	-0.19(0.1)
15	**0.54(0.04)****	**0.54(0.04)****	**0.54(0.04)****	-0.04(0.08)
18	0.3(0.05)**	0.3(0.05)**	0.29(0.05)**	0.1(0.09)
24	0.29(0.04)**	0.29(0.04)**	0.28(0.04)**	0.34(0.12)**
28	**0.72(0.04)****	**0.73(0.04)****	**0.72(0.04)****	0.29(0.1)**
29	**0.57(0.05)****	**0.58(0.05)****	**0.57(0.05)****	0.08(0.09)
36	**0.62(0.05)****	**0.63(0.05)****	**0.62(0.05)****	-0.06(0.09)
39	**0.67(0.04)****	**0.68(0.05)****	**0.66(0.05)****	0.27(0.1)**
Feelings				
1	0.11(0.04)**	0.11(0.04)**	0.11(0.04)**	-0.05(0.07)
4	0.06(0.04)	0.06(0.04)	0.06(0.04)	-0.33(0.08)**
8	**0.4(0.05)****	0.39(0.05)**	**0.4(0.05)****	0.06(0.08)
9	0.21(0.04)**	0.21(0.04)**	0.21(0.04)**	0.23(0.08)**
10	**0.55(0.04)****	**0.54(0.04)****	**0.55(0.04)****	-0.19(0.08)*
12	0.07(0.04)	0.07(0.04)	0.07(0.04)	0.17(0.08)*
13	0.11(0.05)*	0.11(0.05)*	0.11(0.05)*	0.01(0.08)
14	**0.68(0.05)****	**0.66(0.05)****	**0.68(0.05)****	-0.2(0.09)*
20	**0.62(0.05)****	**0.61(0.04)****	**0.62(0.05)****	0.07(0.08)
22	**0.52(0.05)****	**0.51(0.05)****	**0.52(0.05)****	-0.24(0.08)**
25	**0.48(0.04)****	**0.47(0.04)****	**0.48(0.04)****	0.2(0.07)**
26	0.18(0.04)**	0.18(0.04)**	0.18(0.04)**	0.21(0.08)**
30	**0.61(0.05)****	**0.6(0.05)****	**0.61(0.05)****	-0.11(0.09)
33	**0.6(0.04)****	**0.59(0.04)****	**0.6(0.04)****	0.1(0.08)
34	0.15(0.04)**	0.15(0.04)**	0.15(0.04)**	0.17(0.07)*
38	**0.56(0.04)****	**0.55(0.04)****	**0.56(0.04)****	-0.13(0.07)
40	0.35(0.04)**	0.34(0.04)**	0.35(0.04)**	0.01(0.08)
45	**0.54(0.05)****	**0.53(0.05)****	**0.55(0.05)****	0.2(0.08)*
Intentions				
2	**0.59(0.04)****	**0.59(0.04)****	**0.59(0.04)****	0.04(0.07)
3	**0.41(0.04)****	**0.42(0.04)****	**0.41(0.04)****	0.11(0.07)
5	0.3(0.05)**	0.3(0.05)**	0.29(0.05)**	0.1(0.07)
7	0.29(0.05)**	0.3(0.05)**	0.29(0.05)**	0.16(0.07)*
11	**0.76(0.06)****	**0.77(0.06)****	**0.76(0.06)****	0.11(0.09)
16	**0.56(0.05)****	**0.56(0.05)****	**0.55(0.05)****	0.3(0.08)**
17	**0.48(0.05)****	**0.48(0.05)****	**0.47(0.05)****	0.15(0.07)*
19	0.38(0.04)**	0.38(0.04)**	0.38(0.04)**	0.03(0.07)
21	**0.45(0.04)****	**0.46(0.04)****	**0.45(0.04)****	0.12(0.07)
23	**0.48(0.04)****	**0.48(0.04)****	**0.47(0.04)****	0.05(0.07)
27	**0.67(0.05)****	**0.67(0.05)****	**0.68(0.05)****	-0.28(0.08)**
31	0.32(0.04)**	0.32(0.04)**	0.33(0.04)**	-0.17(0.07)*
32	0.23(0.04)**	0.23(0.04)**	0.22(0.04)**	**0.48(0.09)****
35	0.08(0.04)*	0.08(0.04)*	0.08(0.04)	0.04(0.07)
37	0.39(0.04)**	0.39(0.04)**	0.38(0.04)**	0.12(0.07)
41	**0.62(0.04)****	**0.62(0.04)****	**0.63(0.04)****	-0.19(0.07)*
42	0.16(0.04)**	0.17(0.04)**	0.15(0.04)**	0.28(0.07)**
43	0.38(0.04)**	0.38(0.04)**	0.38(0.04)**	-0.04(0.07)
44	**0.93(0.04)****	**0.93(0.04)****	**0.93(0.04)****	-0.04(0.08)

Note. *p<.05, **p < .01, factor loadings > .40 bolded

Figure 1. *Unidimensional factor structure and standardized loadings*

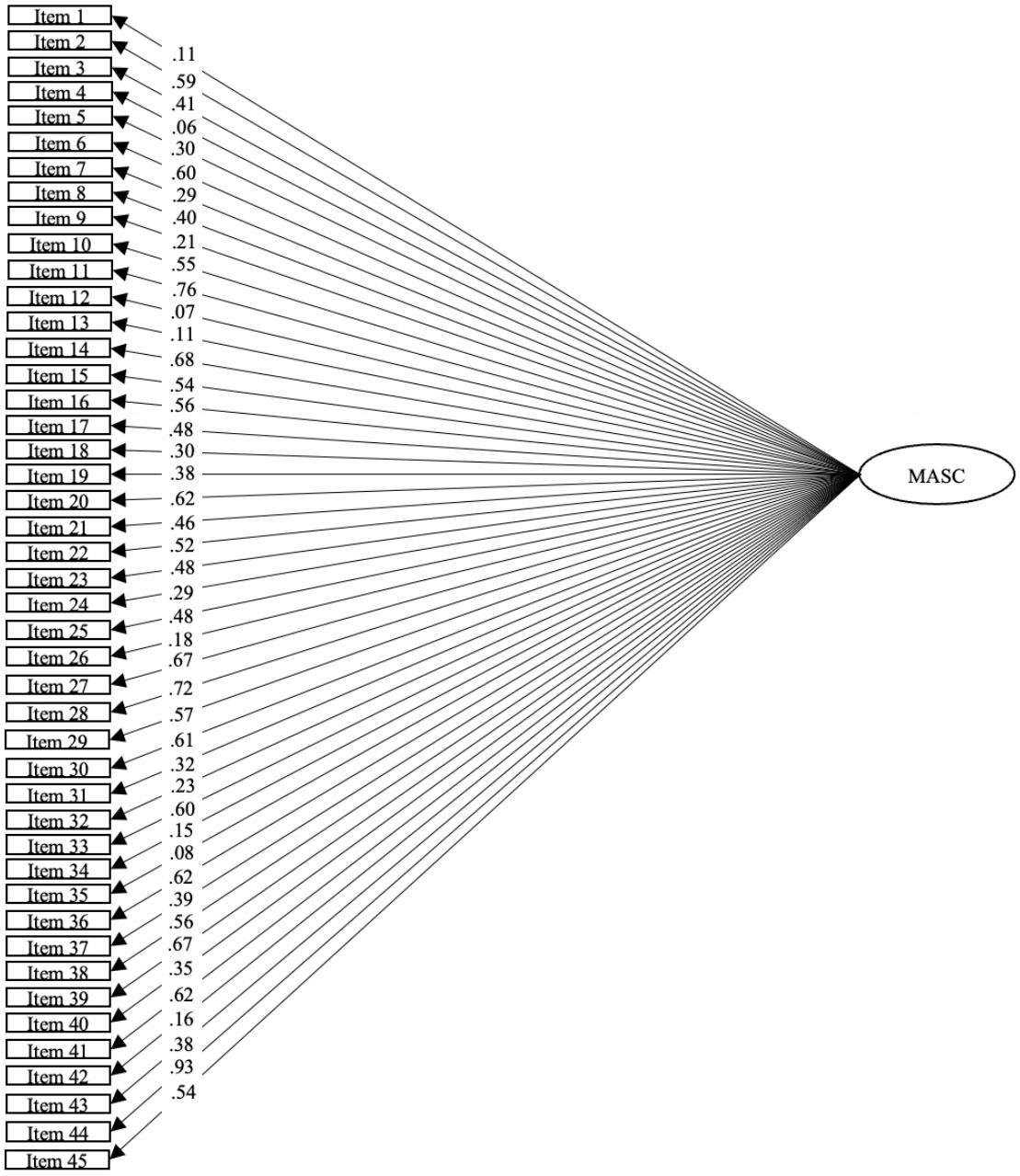

Table 2. *Cross-racial means and group comparisons for MASC total and error scores*

	Black/AA **M(SD)**	**Asian/PI** **M(SD)**	**Hispanic White** **M(SD)**	**Non-Hispanic White** **M(SD)**	**F(3, 1226)**	η^2	**Significant post hoc comparisons**
Correct	32.31(6.27)	32.23(5.74)	32.90(6.30)	35.05(4.94	6.15**	.02	Non-Hispanic White > all other groups
Hypermentalizing	5.44(2.98)	6.18(3.11)	5.50(3.23)	5.42(2.94)	6.15**	.01	Asian/PI > non-Hispanic and Hispanic White
Hypomentalizing	4.54(3.17)	4.36(2.74)	4.34(2.63)	3.66(2.30)	5.94**	.01	Non-Hispanic White < all other groups
No mentalizing	2.70(2.45)	2.23(2.05)	2.26(2.37)	1.87(1.80)	5.36**	.01	Black/AA > non-Hispanic White