EXAMINING MEASURES OF SCHIZOTYPY FOR RACIAL AND GENDER BIAS USING ITEM RESPONSE THEORY AND DIFFERENTIRAL ITEM FUNCTIONING

A Thesis

presented to

the faculty of the Graduate School

at the University of Missouri

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts

By

DESMOND SPANN

Dr. John Kerns, Thesis Supervisor

DECEMBER 2022

The undersigned, appointed by the dean of the Graduate School, have examined the thesis entitled

EXAMINING MEASURES OF SCHIZOTYPY
FOR RACIAL AND GENDER BIAS
USING ITEM RESPONSE THEORY
AND DIFFERENTIRAL ITEM FUNCTIONING

presented by Desmond Spann,

a candidate for the degree of master of clinical psychology

and hereby certify that, in their opinion, it is worthy of acceptance.

Professor John Kerns

Professor Denis McCarthy

Professor Edgar Merkle

Professor Francisco Sanchez

ACKNOWLEDGEMENTS

I would like to thank Dr. John Kerns for his immense amount of support and guidance during this project as the role of my academic advisor. I would also like to thank my thesis committee members, Dr. Denis McCarthy, Dr. Edgar Merkle, and Dr. Francisco Sanchez, for their integral suggestion and recommendations throughout the thesis process. Additionally, I would like to thank Dr. David Cicero for his role as lead consultant for the project where he provided essential assistance for the data analysis portion of the project. Finally, I would like to thank my fellow lab member Kelsey Straub for the many valuable role she played during the entire project.

LIST OF ILLUSTRATIONS

Figure Page

LIST OF TABLES

Tables Page

LIST OF ITEMS

Items Page

TABLE OF CONTENTS

Abstract

Schizotypy, a latent personality organization reflecting liability for schizophrenia, includes abnormalities and deficits in functioning and traits which cluster into positive, negative, and disorganized dimensions (Lenzenweger, 2018). Research within schizotypy provides useful information regarding etiological mechanisms and risk and protective factors for schizophrenia-spectrum conditions without the presence of confounds such as medication effects and pronounced cognitive deficits associated with schizophrenia-spectrum disorders. The Magical Ideation Scale (MagicId), Perceptual Aberration Scale (PerAb), Revised Social Anhedonia Scale (RSAS), and the Cognitive Slipage Scale (CSS; Chapman et al., 1976; Chapman et al., 1980; Eckblad & Chapman, 1983; Eckblad et al., 1982; Miers & Raulin, 1987) are a collection of scales which have historically been used to measure schizotypy. A critical and unresolved issue is to what extent measures of schizotypy show racial and gender bias in their assessment with recent evidence suggesting these scales possess biased items. This current study examined racial and gender biases in items of the scales with the use of Item Response Theory (IRT) to test for levels of differential item functioning (DIF). Only two previous studies have used IRT analysis to test for racial and gender bias within the MagicId, PerAb, and/or SocAnh scales and this study was the first for the CSS. I had data for the following racial/ethnic categories: Black (African-American), Asian American, Hispanic, Multiracial, and White. Overall, I found some evidence for gender bias as well as racial bias (especially when comparing Black versus White participants) for some of these scales.

Introduction

Schizotypy has been defined as a latent personality organization reflecting liability for schizophrenia (Lenzenweger, 2018). Similar to schizophrenia (Tandon et al., 2009), schizotypy is generally conceptualized as comprised of multiple dimensions (Kwapil & Barrantes-Vidal, 2015). Research has shown expression of schizotypy exists on a dynamic continuum ranging from relative psychological health and individual personality differences to subclinical deviance to full blown schizophrenia-spectrum disorders such as schizophrenia-spectrum personality disorders and schizophrenia (Gooding & Iacono, 1995; Kwapil et al. 2008; Kwapil & Barrantes-Vidal 2012, Li et al., 2020). Schizotypy traits include abnormalities and deficits in cognitive, social, emotional, and behavioral functioning and are clustered into positive, negative, and disorganized dimensions. Positive traits include psychotic like experiences such as hallucinations and magical thinking, negative traits include anhedonia and decreased emotional expression, and disorganized traits include odd speech and behavior (Fonseca-Pedrero et al., 2018).

Schizotypy has become an important focus within schizophrenia-spectrum disorder research. It has useful explanatory power for etiological mechanisms, developmental pathways, expression, and risk and protective factors for schizophrenia-spectrum conditions and cluster A personality disorders (Cicero, Martin, Krieg, 2019; Fonseca-Pedrero et al., 2018; Kerns, 2020; Kwapil & Barrantes-Vidal, 2015). For example, markedly elevated levels of a common aspect of negative schizotypy social anhedonia (SocAnh), which is decreased interest in and number of social interactions, has been shown to predict significantly increased risk of developing schizophrenia-spectrum disorders (e.g., in a college student sample, 28% at 10-year follow-up versus 1% in controls; Kwapil 1998; also Gooding et al., 2005). Hence, research on schizotypy facets such as SocAnh can provide important evidence to inform better prevention efforts,

assessments, and interventions. Schizotypy symptoms also provide a model for examination of schizophrenia-spectrum symptoms without the presence of the usual confounds found in patient research such as medication effects and pronounced cognitive deficits associated with schizophrenia-spectrum disorders (Fonseca-Pedrero et al., 2018; Kwapil, Barrantes-Vidal, 2015; Li et al., 2021). Assessment of schizotypy allows for identifying individuals possessing liability for psychosis, including in the prodromal stage of schizophrenia-spectrum disorders, before the presence of any clinical manifestations (Kwapil, Barrantes-Vidal, 2015). This provides the opportunity to use non-clinical samples, including college students, to be able to investigate important risk and protective factors attributed to schizophrenia-spectrum disorders. Finally, assessment of schizotypy is relatively inexpensive and noninvasive which increases utility in screening both clinical and non-clinical samples (Kwapil et al., 2008).

A collection of scales which have been widely used to measure schizotypy since the 1970s are the Magical Ideation Scale (MagicId), Perceptual Aberration Scale (PerAb), and the Revised Social Anhedonia Scale (RSAS; Chapman et al., 1976; Chapman et al., 1980; Eckblad & Chapman, 1983; Eckblad et al., 1982). Research has provided a significant amount of evidence which supports the reliability and validity of these scales within clinical, community, and undergraduate samples and in both cross-sectional and longitudinal studies (Blanchard et al., 2011; Chapman et al., 1994; Gooding & Tallent, 2005; Horan et al., 2008; Kwapil et al., 2008). Also, scores on these scales have been associated with higher likelihood of psychotic symptoms, poorer overall functioning, and risk for development of psychopathology including schizophrenia spectrum-disorders (Chapman et al., 1994; Kwapil et al., 2007). There is evidence that these scales predominantly load onto positive (MagicId and PerAb) and negative (SocAnh)

schizotypy dimensions (Cicero & Kerns, 2010; Kerns, 2006; Kwapil et al., 2007; Winterstien et al., 2011).

Many previous schizotypy studies have focused on positive and negative schizotypy scales. Recently, more investigations into the disorganized dimension have shed light on the importance of better understanding this facet of schizotypy. Disorganized schizotypy traits include those similar to disorganization symptoms in schizophrenia such as disorganized speech (Cicero & Kerns, 2010; Hewitt & Claridge 1989; Kerns & Becker, 2008). It has been commonly identified as a separate schizotypy dimension through factor analysis and now is viewed as an important dimension of focus for new schizotypy scale creation (Cicero & Kerns, 2010; Kerns, 2006; Kwapil et al., 2018). The disorganized dimension of schizotypy has been found to be associated with disorganized schizophrenia-spectrum symptoms, depression and elevated negative affect, attentional deficits, and impaired functioning. The Cognitive Slippage Scale (CSS) was designed to be able to measure disorganized schizotypy by assessing characteristics such as thought disorder, abnormal cognitive associations, and inability to keep track of thoughts (Gooding, Tallent, Hegyi, 2001; Miers & Raulin, 1987). Evidence from a longitudinal study found characteristics of disorganized schizotypy, such as thought disorder, are associated with higher lifetime risk for developing schizophrenia-spectrum disorders (Gooding et al., 2013).

However, a critical and as yet unresolved issue is to what extent measures of schizotypy show racial and gender bias in their assessment. For schizophrenia-spectrum disorders, there is consistent evidence that disorder and symptom prevalence vary by race and gender. Critically, there is also evidence that assessment bias contributes to variance in prevalence.

For instance, race/ethnicity has been found to be a strong predictor of schizophrenia-spectrum disorders with Black and Latino(a) Americans found to be between 3–4 times more

likely to be diagnosed on the spectrum compared to White (non-Latino) Americans (Barnes, 2004; Blow et al, 2004; Boa et al, 2008; Eack, 2012; Minskey et al, 2003; Neighbors et al, 2003, Schwartz et al, 2009, Perry et al, 2013). A recent meta-analysis suggests racial diagnostic disparity in schizophrenia is a robust heterogenous clinical phenomenon with greater proportion of White participants and studies with younger participants showing consistently higher racial diagnostic disparities which has been stable over the last three decades (Olbert, Nagendra, & Buck, 2018). Consistent with assessment bias being a contributing factor, previous research has found Black and Latino(a) Americans are more likely to be misdiagnosed with schizophrenia-spectrum disorders compared to White Americans (Schwartz & Blankenship, 2014). In addition, potentially consistent with clinical diagnostic bias contributing to race/ethnicity differences in prevalence, racial diagnostic disparities are larger in studies with more White participants (Olbert et al., 2018).

As for gender, schizophrenia-spectrum disorder research has often found differences between men and women. This includes differences in age of onset, premorbid and overall functioning, symptom presentation, and substance use (Giordano et al., 2021). In particular, males have been found to have an earlier age of onset, worse functioning, and higher rates of comorbid substance abuse. In contrast, females have been found to have a higher frequency of comorbid affective disorders and a higher rate of suicide attempts. Gender differences in schizophrenia-spectrum disorder symptoms are frequently reported, but research findings have also been somewhat inconsistent. In particular, there is evidence that males have a higher severity of negative symptoms while females have a higher frequency of positive symptoms. However, there is evidence that whether these gender differences in symptoms are found could be due to using outdated measures, which contain bias for psychosocial or cultural factors

(Giordano et al., 2021). For example, a study examining sex differences in schizophrenia across diverse regions of the world did find a greater severity of positive symptoms in females, but in only some regions examined (Novick et al., 2016). This suggests there is the potential for specific psychosocial or cultural factors of these regions to influence the severity of symptoms with a differential effect on gender. Overall, there is evidence of racial and gender disparities across schizophrenia-spectrum disorders, with evidence that assessment biases contribute to these differences.

Given the importance of schizotypy as a way of understanding and assessing risk for schizophrenia-spectrum disorders, an important question in understanding racial and gender bias in schizophrenia-spectrum disorders is whether assessments of schizotypy also exhibit racial and gender bias. Similar to research on schizophrenia-spectrum disorders, racial and ethnic minorities have also been found to report higher levels of schizotypy symptoms such as magical thinking/delusional ideation and SocAnh compared to their White counterparts (Goulding, McClure-Tone, & Compton, 2009; Lincoln et al., 2022; Sharpley & Peters, 1999). Also similar to schizophrenia-spectrum disorders, males have been found to report higher levels of negative schizotypy such as SocAnh while females have been found to report higher levels of positive schizotypy. An important question is whether these racial and gender differences in schizotypy are related to assessment bias. Understanding the nature of these racial and gender differences could have important implications for schizotypy assessment as well as attempts to improve our understanding of schizophrenia-spectrum etiology, assessment, and prevention. For instance, it has been argued that understanding etiological factors that result in racial differences in schizophrenia-spectrum disorders could have critical implications for disorder prevention

(Anglin et al., 2020). However, this might depend critically on the extent to which differences reflect assessment bias.

Again, the Magical Ideation, Perceptual Aberration, and Revised Social Anhedonia Scales have been used extensively and do display good evidence for reliability and validity in cross-sectional and longitudinal studies including normal, at-risk, and clinical samples (Chapman et al., 1995; Kwapil et al., 2008; Winterstien et al., 2011). Though strong psychometric properties have been found, recent evidence suggests they also might possess racial and gender bias (Winterstein, 2010). An important factor in potentially contributing to bias with these scales that these scales were created using Classic Test Theory because modern statistical techniques, like Item Response Theory (IRT), were not available at the time of scale construction (Cicero et al., 2019; Winterstien et al., 2011).

A goal for any scale or measure is to capture an estimate of the level of the theoretical attribute, trait, symptom, or construct of interest intended to be quantified by the scale. Classic Test Theory assumes that each individual item of a scale contributes equally to assessing an attribute (Cappelleri, Lundy, & Hays, 2014; Nunnally, 1978). IRT on the other hand views each individual item of a scale as providing unique information regarding the level of the underlying attribute. Hence, an IRT analysis provides information about the performance of items regarding the theoretical underlying construct. Therefore, it is more interested in individual item performance than the overall sum of the scale (Neal, 2006). An IRT model can help explain how and why changes in the latent construct causes changes in item response. In a way, it could be viewed as the latent attribute causing variation in item response.

The goal for IRT is to estimate parameters which model the relationship between the items of the scale and the latent attribute being measured. This is done with the use of item

response curves (IRC) which are probabilistic functions that relate item responses to the underlying level of the attribute (Riese & Waller, 2009). IRCs provide information about responses to an item at a given level of the underlying trait. IRT models apply different parameters to response data and these parameters effect the properties of the IRC which informs what specific information can be gathered from the IRCs.

As can be seen in Figure 1a, the first model, termed as a 1PL model, applies the location or "difficulty" parameter where items can only differ on the location or level of the underlying trait. This is represented by the location of the IRC on the x-axis with the higher the level of the underlying trait, the further along the x-axis the IRC is located. Items deemed more difficult require a higher level of the underlying trait to be answered correctly than less difficult items. In Figure 1a, one can imagine that the different lines represent different items, with some items tending to be endorsed only at higher levels of the trait. Note that these curves all have the same slope.

As can be seen in Figure 1b, a second model, a 2PL model, adds the discrimination parameter, represented by the slope of the IRC, which is a measure of how much the level of the underlying trait influences responses to the item. In Figure 1b, these curves all have different slopes as they differ in how increases in the level of the trait are related to the frequency of item endorsement. 3PL and 4PL models add parameters to protect against guessing and careless responding.

An IRT analysis can allow for valid testing of scale bias. Differential item functioning (DIF) has been coined as the term to represent bias within IRT. If an item possesses DIF, this means the item is relatively more difficult, discriminating, or less easily guessed by one group relative to other groups (Nunnally, 1978).

For instance, returning to Figure 1a, imagine that these curves are for one item for five different groups. Here, groups vary in terms of the level of the trait that leads to endorsement of the item. Hence, groups vary in their responses to the item and higher scores in one group versus another does not necessarily mean that this reflects a true difference in the level of the trait. This type of bias has been referred to as uniform bias and exists when there is no interaction between level of the underlying trait and group membership (Mellenberg, 1982; Swaminathan & Rogers, 1990). Groups may vary in their response behavior, but these differences are uniform over all levels of the underlying trait.

Returning to Figure 1b, imagine again that these reflect IRC curves for one item for five different groups. Here, groups vary in how the level of the trait is related to the frequency of item endorsement. For this item, higher scores in one group versus another does not necessarily mean that this reflects a true difference in the level of the trait. This type of bias has been referred to as non-uniform bias and exists when there is interaction between level of the underlying trait and group membership (Mellenberg, 1982; Swaminathan & Rogers, 1990). Groups may vary of in response behavior similar to uniform DIF, but within non-uniform these differences are not uniform across all levels of the underlying trait and illustrated by non-parallel IRCs. IRT analyses can be used to test for whether such biases are present in items on a scale.

Reviewing the schizotypy literature, only two previous studies have used IRT analysis to test for racial and gender bias within the MagicId, PerAb, and/or SocAnh scales. Winterstein et al. (2011) was the first to apply IRT to test these scales for the presence of DIF between racial and gender groups. Their sample included 4,664 women and 1,473 men, with 4,529 White/Caucasian participants and 1,608 Black/African American participants. Results from a 2PL model analysis found the percentage of items which exhibited DIF on the scales varied from

between 23–48%; 47% for MagID, 23% for PerAb, and 48% for SocAnh/RSAS. Effect sizes for the size of DIF for items ranged from moderate to high, with the RSAS having the highest percentage of high effect size DIF items (48%). The analysis found each scale had a significant percentage of items which possessed problematic levels of DIF for both racial and gender groups. Further, items that possessed DIF for racial and gender groups also had lower discrimination values for the traits being measured, with discrimination here referring to the item being able to differentiate between the low and high end of the trait being measured. Overall, results from Winterstein et al. suggested that a significant number of items on these scales were biased in terms of racial and gender identity and furthermore that these items were doing a poor job of measuring the traits of interest due to this bias. However, this study only examined two racial groups.

More recently, Cicero et al. (2019) examined DIF with these scales with somewhat different racial groups. Their sample included 1,175 women, 1,130 men and 68 who declined to specify their gender and there were 1,051 Asian, 541 Multiracial, 416 White/Caucasian, and 303 Hispanic/Latino(a) participants. Results from the study again suggested the presence of DIF for these scales between groups both racial and gender groups. Asian and Latino(a) participants displayed significantly higher latent means on the scales compared to white participants and this result suggest a presence of DIF large enough to change interpretations of group differences. For gender, the Magical Ideation scale performed the worst with 20% of items displaying some level of DIF. Overall, the results suggest DIF is present in the scales and the DIF may be important for understanding group differences between the groups.

Hence, the two previous studies that have examined DIF for these three scales have found evidence for some racial and gender bias. However, this has still only been examined in two

studies. Further, other than White participants, this has only been examined in different American racial groups in one study each. And no study has examined DIF for the Cognitive Slippage Scale measuring disorganized schizotypy.

Therefore, the current study will conduct an IRT analysis to examine DIF within the Magical Ideation, Perceptual Aberration, Revised Social Anhedonia, and Cognitive Slippage Scales. This study will be first study to examine racial and gender bias in the Cognitive Slippage Scale. It will also be only the third study to examine gender bias in the Magical Ideation, Perceptual Aberration, Revised Social Anhedonia Scales. And it will be only the second study to examine racial bias for African American/Black, Asian/Pacific Islander, Latino/Latina, and Multiracial participants in these three scales. One limitation of the current study is that only a subset of items from these scales were used (as these subsets of items were used for screening participants) and we cannot make inferences about the full scales. However, again, IRT analyses focus on DIF in individual items. Further, previous research has found that subsets of items from these scales do form reliable and valid instruments (Cicero et al., 2019). Finally, research on bias with these scales has been so infrequent that it seems important to take advantage of the present large dataset to be able to further investigate racial and gender bias. The presence of bias amongst these subsets of items will still provide important evidence about the presence of bias in commonly used schizotypy scales.

Method

Participants

Data used in this investigation was constructed with mass pre-test data collected from the "Psych 1000" pool using the Mizzou SONA system between the years of 2005-2019. The data represents the participant screening procedure used to identify participants who are potentially

at-risk for development of psychosis/schizophrenia-spectrum disorders. This compiled data results in a sample of 21,829 participants. The sample is 53% female and 82% White (*n* = 17,831), 7% Black (n=1,514), 5% Multiracial (n=1,075), 3% Asian American/Pacific Islander (n=710), 2% Latino/Latina (n=339), 1% Other. Each participant does not possess an answer for each of the items within the data set due to the variation in the items used to screen participants from year to year. This results in the samples for the Magical Ideation Scale (n =18,881), Perceptual Aberration Scale (n =18,890), Revised Social Anhedonia Scale (n =17,370), and Cognitive Slippage Scale (n= 6,565) being different sizes compared to the overall sample.

Measures

Magical Ideation Scale

The Magical Ideation Scale is a 30-item, true-false self-report scale which measures beliefs in forms of causation that by conventional standards are invalid (Eckbad & Chapman, 1983). Magical ideation as a symptom refers to a belief in several magical influences and experiences include things such as thought transmission between people (Eckbad & Chapman 1983). The magical ideation scale has been found to possess utility in predicting future schizophrenia-spectrum disorders. Chapman et al., (1994) performed a 10-year longitudinal study and found elevated magical ideation predicted future onset of psychotic disorder. Seven items from the Magical Ideation Scale are included in this investigation.

Perceptual Aberration Scale

The Perceptual Aberration Scale (PerAb) is a 35-item, true-false scale which measures psychotic-like perceptual experiences and body distortions (Kwapil et al., 2008). Perceptual Aberration includes experiences such as unclear boundaries of the body, feeling of unreality or estrangement of parts of one's body, feeling of deterioration of one's body; perception of change

in the size, proportions, spatial relationship of one's body part, and changes in the appearance of the body (Chapman et al., 1978; Fornasari et al., 2015). Individuals who rate high on perceptual aberration have reported greater frequency of psychotic like experiences as well as higher rates of psychosis, schizotypy symptoms, and relatives with psychotic disorders (Chapmen, et al., 1994, Kwapil et al., 2013; Fornasari et al., 2015). Eight items from the Perceptual Aberration Scale are included in this investigation.

Revised Social Anhedonia Scale

The RSAS is a 40-item true-false questionnaire that measures an individual's lack of relationships as well as lack of pleasure derived from and interest in relationships (Eckblad et al., 1982). High endorsement of RSAS items has been found to predict a significant increase in risk for future development of schizophrenia-spectrum disorders (Gooding et al., 2005; Kwapil, 1998). 15 items from the RSAS are included in this investigation.

Cognitive Slippage Scale

The Cognitive Slippage Scale (CSS) is a 35-item, true-false self-report measure designed to identify thought disorder schizotypal characteristics such as disorganized speech and confused thinking (Gooding, Tallent, Hegyi, 2001; Miers & Raulin, 1987). The CSS has been found to strongly correlate with other measures of disorganized schizotypy (Cicero & Kerns, 2010; Kerns, 2006). 15 items from the CSS are included in this investigation.

Data Analysis

The data analysis for this investigation will follow the strategies used in Cicero, Martin, & Krieg (2019) as well as in Straub & Kerns (2020). Data analysis will be done using Mplus 8.7 (Muthén & Muthén, 1998-2021). First, I will examine whether a significant amount of DIF is present when considering each scale in total. I will conduct these global tests of measurement

invariance for each scale within a multiple indicator multiple causes (MIMIC) model. Following Cicero et al. (2019), for this I will specify a two-parameter (2PL) model for each scale. I will compare the fit of two different models. One model will regress a latent factor and all individual items on dummy-coded ethnicity or gender variables. The second constrained model will have these regression weights constrained to zero. DIF at the scale level would be indicated by the first model fitting significantly better than the second constrained model. Tests of model fit will be examined using a chi-square difference test and with Bayesian Information Criterion (BIC), Akaike Information Criterion (AIC), and Sample-Size Adjusted BIC (SABIC). Given that these scales were created with largely White samples, White participants will be the reference group. Individual dummy-coded variables will be created for each other race/ethnicity group. In addition, for gender, the reference group will be females.

Third, I will examine whether a significant amount of uniform DIF is present for each individual item. For this, I will use the same model as mentioned in the preceding paragraph that will regress a latent factor and all individual items on dummy-coded race/ethnicity or gender variables. Uniform DIF will be indicated by a significant beta weight for that item.

Fourth and finally, I will examine whether a significant amount of non-uniform DIF is present for each individual item. For this, following Woods and Grim (2011), I will create interaction variables between latent trait by race/ethnicity and gender group variables. I will then conduct an analysis where I regress each item on these interaction variables. Non-uniform DIF for an item would be indicated by a significant beta weight for an interaction variable.

To examine effect sizes for DIF, I will compute the Education Testing Service (ETS) Δ as a measure of effect size. ETS Δ is linear transformation of β which is multiplied by -2.35 (Monahan et al., 2007). Items with ETS Δ between 0 and 1 are considered to have negligible

DIF, 1 and 1.5 are considered to have slight to moderate DIF, and over 1.5 are considered to have moderate to high DIF.

Results

Group Differences in Mean Scores

First, I examined whether there were mean group differences on schizotypy scores. Based on previous research, I expected that females and racial/ethnic minorities would generally have higher schizotypy scores than males and White participants. As can be seen in the Appendix in Supplemental Tables 1-6, in general, results were consistent with expectations, with a few exceptions. Specifically, there was not a significant group difference on Cognitive Slippage scores between Hispanic and White non-Hispanic individuals [$t(5795)=0.682$, $p=.495$]. In addition, there were no gender differences for Cognitive Slippage and for Magical Ideation. But for all other group comparisons, there were significant between group differences. This included higher schizotypy scores in racial/ethnic minorities than for White participants. As expected, this also included higher Perceptual Aberration scores in females than in males. An expected exception to this pattern is that for Social Anhedonia females had lower scores than males.

Global Measurement Invariance Analysis

I then conducted a global measurement invariance for all the scales examine whether a significant amount of DIF is present when considering each scale in total. As shown in Table 7, there was mixed evidence for whether the scales had measurement invariance among groups. For every scale, the models where regression weights were constrained to zero fit the data worse than the models where parameters were freely estimated based on the X^2 difference test. Conversely, the BIC and SABIC scores, which highlight parsimony in the model, displayed the models with

fixed parameters fit the data best, while the AIC scores implied the freely estimated models fit best. These mixed results suggest examining whether the scales have DIF at the item level.

Differential Item Functioning by Group

Gender. As can be seen in Table 8, there was substantial evidence of DIF by gender across all four schizotypy scales. In particular, for 4 of the 8 Magical Ideation items, for 4 of 7 Perceptual Aberration items, for 12 of the 15 Social Anhedonia items, and for 7 of the 15 Cognitive Slippage items there was evidence of uniform and/or non-uniform DIF by gender. For instance, as can be seen in the Appendix in Supplemental Tables 9-12 for four Magical Ideation items and for four Perceptual Aberration items there was significant evidence of uniform DIF by gender. The positive beta weights for all of these items indicate that males were less likely to endorse these items than females at lower levels of the trait. This evidence for uniform DIF for Perceptual Aberration is consistent with the view that part of the reason females might have higher scores on this scale might reflect DIF. Effect sizes for some of these group differences were in the negligible range, with one Magical Ideation item and two Perceptual Aberration items having effect sizes in the small to moderate range.

For Social Anhedonia, where females tend to have lower mean scores than males, for 5 items there was uniform DIF. Negative beta weights here indicate that for these social anhedonia items that females were less likely than males to endorse these items at lower levels of the trait. Effect sizes for three of these items were in the moderate to severe range, with one item each having effect sizes in the small to moderate range and in the negligible range. Hence, this pattern for Social Anhedonia is consistent with the view that part of the reason for higher mean scores in males might reflect DIF. In addition, there were seven other Social Anhedonia items where there was evidence of non-uniform DIF (with for some of these items also significant evidence of

uniform DIF). With these seven items, non-uniform DIF indicates that there were gender differences in the relationship between item endorsement by level of the trait. Effect sizes for non-uniform DIF for these seven items were in the negligible range (in fact, for all analyses, effect sizes for non-uniform DIF were in the negligible range with one exception).
For Cognitive Slippage, for three items there was significant evidence of uniform DIF. Specifically, males were less likely than females to endorse these items at lower levels of the trait. In addition, for four items where was significant evidence of non-uniform DIF.

Race/Ethnicity. As can be seen in Table 8 and in the Appendix in Supplemental Tables 13-20, for three of the schizotypy scales, there was substantial evidence of DIF for Black participants relative to White participants. For Magical Ideation, 5 of the 8 items exhibited DIF. For three of these items there was evidence of only uniform DIF, although beta weights for these items were not completely consistent. There were two items where Black participants were more likely and one item where Black participants were less likely than White participants to endorse items at lower levels of the trait. There were also two Magical Ideation items with non-uniform DIF.

For Perceptual Aberration, there were five items with uniform DIF, although beta weights across these items were also not consistent. There were two items where Black participants were more likely and three items where Black participants were less likely than White participants to endorse items at lower levels of the trait. Effect sizes across these items varied from negligible to the severe range, with the two items with effect sizes in the moderate to severe range involving negative beta weights.

For social anhedonia, there were 5 items with uniform DIF where Black participants were less likely than White participants to endorse these items at lower levels of the trait. Effect sizes

for these five items were all in the moderate to severe range. There were also seven Social Anhedonia items with non-uniform DIF for Black participants. Overall, there was substantial DIF for Black participants relative to White participants for three of the scales, although the pattern of DIF across items for two of the scales was somewhat inconsistent.

For other racial/ethnic groups, the only other instance of substantial DIF was for Multiracial participants for the Magical Ideation Scale. There were three items with uniform DIF where Multiracial participants were less likely than White participants to endorse these items at lower levels of the trait. Effect sizes for two of these items were in the moderate to severe range. There was also one item with significant non-uniform DIF. Interestingly, despite group differences in mean levels for a number of scales, there was not substantial evidence of DIF for Asian-American or Hispanic participants.

Discussion

It has been argued that social stressors including racial discrimination increases risk for schizophrenia-spectrum disorders (Anglin et al., 2020; Selten et al., 2013). From this view, it might be expected that racial and ethnic minorities might have increased levels of schizophrenia-spectrum traits. Conversely, there is evidence of gender and racial bias in diagnosing schizophrenia-spectrum disorders and in measures of schizophrenia-spectrum traits, with measures of these traits developed in White samples. From this view, any evidence of elevated levels of schizophrenia-spectrum traits in females or in non-White participants could reflect biased assessment. The current research examined evidence of differential item functioning in some frequently used self-report measures of schizotypy. Overall, I found further evidence that some measures of schizotypy exhibit differential item functioning, especially between females

and males and between Black and White participants. But evidence of differential item functioning was not found in all group comparisons nor across all schizotypy measures.

For all schizotypy scales, there was evidence of differential item functioning across gender. For instance, for the Revised Social Anhedonia Scale (RSAS), 80% of items exhibited uniform DIF, non-uniform DIF, or both across gender. Although we did not have data for the entire RSAS, current results might generalize to the full scale as the number of items with DIF (n=12) reflects a high enough proportion to be considered evidence of DIF for the full scale (30%; 12 out of 40; with 20% a commonly used threshold; Cicero et al., 2019). Effect sizes for uniform DIF varied, but were sometimes in the medium to severe range (whereas effect sizes for non-uniform DIF in the current study were nearly all considered negligible in size). For social anhedonia items that exhibited uniform DIF, the pattern of current results is somewhat consistent with the view that mean gender differences in social anhedonia could at least in part reflect DIF rather than true gender differences in social anhedonia, with males being more likely to endorse social anhedonia at lower levels of the trait. However, a majority of the social anhedonia items exhibiting DIF exhibited non-uniform DIF, which makes interpreting how this affects scores by gender more complicated. These results appear generally similar to the two previous IRT studies on the RSAS involving American samples (Cicero et al., 2019; Winterstein et al., 2011), although the sample size in the current study is far larger than in these prior studies (e.g., current study with nearly seven times more male participants than in previous studies). On the whole, the current results suggest caution about attempting to compare social anhedonia scores between gender. Social anhedonia has been found to be an important measure in the prediction of schizophrenia-spectrum disorders (e.g., 10-year risk = 28% versus 1% in control participants; Kwapil, 1998; also Gooding et al., 2005). However, the presence of DIF by gender for the RSAS

should be an important factor when attempting to use this scale in predicting disorder risk. For instance, the current results suggest that it is important to consider norms by gender for the RSAS.

In the current research, I also examined for first time the possible presence of DIF in the Cognitive Slippage Scale (CSS), a measure of disorganized schizotypy. The current study found evidence of DIF by gender for the CSS, with seven out of fifteen items exhibiting significant DIF. Further, although we did not have data for all CSS items, the number of items exhibiting DIF would be considered a problematic level of DIF for the full scale (20%; or 7 out of 35). Effect sizes for some of the items exhibiting uniform DIF were in the moderate to severe range. Hence, the current results suggest that the presence of DIF for the CSS is an important factor to consider when using this scale.

There was also evidence of DIF across gender for the two positive schizotypy scales, Perceptual Aberration and Magical Ideation. For our fifteen positive schizotypy items, eight displayed significant DIF. For all of these items, males tended to be less likely to endorse items at lower levels of the trait, which might tend to decrease male scores relative to female scores. These positive schizotypy scales have been found to predict future psychotic disorders (Chapman et al., 1994). Possible DIF by gender for these scales could be an important factor in the use of these scales. However, it should also be noted that effect sizes for DIF for positive schizotypy were often in the negligible range. This seems generally consistent with the results of previous IRT research examining DIF by gender for these positive schizotypy scales (Cicero et al., 2019). Another caveat to the current results for positive schizotypy is that given that only used a subset of these scale items were used then it is not possible to generalize to the full scales. For instance, in the current study we found DIF by gender for 53% of positive schizotypy items (8 out of 15).

However, in the full context of these scales, this would not reach the threshold of at least 20% of items exhibiting DIF (13%; 8 out of 63).

In addition to evidence of DIF by gender, the current study also found evidence of DIF for Black versus White participants for three of the four schizotypy scales, the RSAS and the two positive schizotypy scales. For the RSAS, the number of items displaying DIF would be considered a problematic level of DIF for the full scale. It is also somewhat striking that for the five RSAS items exhibiting uniform DIF that the effect sizes tended to be all in the moderate to severe range, with these effect sizes being the largest amongst the current results. For these five items, the pattern of results suggests that White participants were less likely to endorse items at the lower levels of the trait which would tend to decrease their scores relative to Black participants. However, the presence of significant non-uniform DIF for several items would further complicate interpretations of scores between Black and White participants. The current results appear roughly similar to the one previous IRT RSAS study involving Black and White participants (Winterstein et al., 2011). Overall, the current study found evidence of potentially substantial DIF for the RSAS for Black versus White participants, suggesting that this is an important factor in using this scale.

There was also evidence of DIF for Black versus White participants for the majority of positive schizotypy items. Across the two positive schizotypy scales, there was significant evidence of DIF for 67% of the items. Interestingly, there was not a clear pattern of DIF for Black versus White participants across these scales so it is not the case that DIF would be expected to generally raise or lower mean scores for either group. However, the evidence of DIF for 67% of items potentially suggests that these scales are not directly comparable for Black and White participants and that this should inform use of these scales. It was also the case that effect

sizes for DIF were only in the moderate to severe range for three of the ten items. Further, given that only used a subset of these scale items were used then it is not possible to generalize to the full scales (i.e., DIF found for 10 items; so only DIF for 16% in the context of the full scales, 10 out of 63).

The only other seemingly notable example of DIF in the current results was for Multiracial versus White participants for the Magical Ideation scale. On this scale, there was DIF for 50% of items. When also considering the Perceptual Aberration scale, there was DIF for 40% of positive schizotypy items, with half of these items exhibiting non-uniform DIF. Hence, the current results provide novel evidence of potential DIF for Multiracial participants relative to White participants for positive schizotypy. However, given that only used a subset of these scale items were used then it is not possible to generalize to the full scales. The current research suggests that continuing to examine potential DIF for positive schizotypy for Multiracial versus White participants is warranted.

Although the current study found some evidence of DIF for these scales, there were also some notable instances where DIF was not found. For instance, there was no evidence of substantial DIF for any racial/ethnic group comparisons for the disorganized schizotypy Cognitive Slippage Scale. Further, I did not find clear evidence of DIF for either Asian-American or Hispanic versus White non-Hispanic participants. A potentially important caveat here is that for the Revised Social Anhedonia Scale there was DIF for 27% of items for Asian-American versus White participants (4 out of 15 items). If extrapolated to the full scale, then this would be considered a problematic level of DIF. However, for all other scales and group comparisons, there was little evidence of DIF in the current study. The current results suggesting that DIF for schizotypy may not be substantial when comparing Asian-American or Hispanic

versus White non-Hispanic participants seems consistent with the one prior study to examine this issue (Cicero et al., 2019). However, there are some important caveats to the current results. Again, given that only a subset of these scale items were used, it is not possible to generalize to the full scales so results could have been different if we used the full scales. also, as a racial/ethnic category, "Asian-American" is a large panethnic category and is comprised of many different more specific ethnic groups (Amer & Bagasra, 2013; Hall, Yip, & Zárate, 2016). Hence, the current study results may not generalize to everyone who would endorse Asian-American as a racial/ethnic category. Also, it should be noted that there were fewer Asian-American and Hispanic participants than Black participants which could have influenced why we only found evidence of DIF for Black participants. Future research on potential DIF for schizotypy for Asian-American or Hispanic versus White non-Hispanic participants appears warranted, although again thus far strong evidence of DIF appears lacking.

The current study suggests several issues for future schizotypy research. Building off the results of this study, future investigations could examine the underlying factor structure of these constructs across groups. This may provide information regarding how the items on the scales are functioning for each group. For example, a recent investigation into the underlying factor structure of depression between racial groups using the Epidemiological Studies Depression Scale (CES-D) found strong evidence the original hypothesized structure may not be the best fit for all racial/ethnic groups (Kim et al., 2011). There is the potential for these traits to present differently across these groups and it would be important to understand the dimensionality of these traits across groups. Future investigations could also examine how stress could mediate or moderate group differences in schizotypy. As stated earlier, a theory behind group differences includes stress being a mechanism for high levels for minority groups. It would be important to

understand the role stress plays in the presentation of these traits. Along these lines, given evidence of substantial DIF for the RSAS for Black versus White participants, future research could also examine cultural factors specific to Black Americans that affect their social behavior. For example, Black individuals may be more likely to disengage from social behavior to protect themselves against potentially discriminatory racial events (DREs), especially in majority white spaces, which could be mistaken for social anhedonia. Finally, given that the current study involved college students, it would be important to examine DIF for these schizotypy scales in non-college student populations.

Overall, the current study found evidence of potentially problematic levels of DIF for a number of schizotypy scales and for a number of group comparisons. This suggests that in some contexts that these scales should be used with caution. However, there were other scales and some group comparisons where problematic levels of DIF were not found. Elevated levels of schizotypy symptoms have shown to predict significantly increased risk of developing schizophrenia-spectrum disorders (Chapman et al., 1994; Kwapil et al., 2007; Kwapil 1998; Gooding et al., 2005; Gooding et al., 2013). It has also been argued that schizotypy research can provide useful information regarding etiological mechanisms and developmental pathways as well as risk and protective factors for schizophrenia-spectrum conditions (Cicero, Martin, Krieg, 2019; Fonseca-Pedrero et al., 2018; Kerns, 2020; Kwapil, Barrantes-Vidal, 2015). It will be important for future research to continue to examine to what extent group differences in schizotypy levels potentially reflect etiologically meaningful differences versus biases in assessment instruments.

Appendix

Figure 1. 1PL and 2PL IRT Models. X-axis reflects level of the trait (theta). Y-axis reflects probability of response. See text for further description. (A) 1PL model. (B) 2PL model.

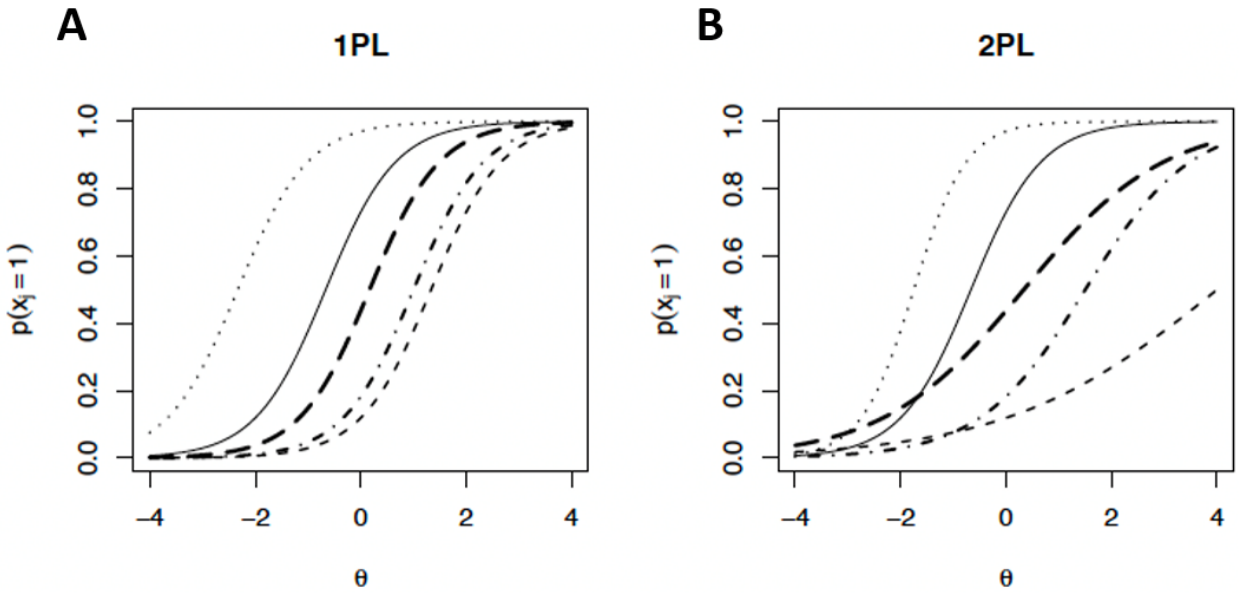

Table 1. One-way ANOVA for Z-Scores for Race/Ethnicity Groups

Measure	White		Black		Asian		Hispanic		Multiracial		*F*	*p*	n^2
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*			
Cognitive Slippage	-.02	1.00	.11	1.03	.11	1.03	-.09	.90	.18	1.10	5.53	<.001	.003
Social Anhedonia	-.05	.96	.30	1.08	.30	1.08	.09	1.13	.10	1.10	64.60	<.001	.015
Magical Ideation	-.04	.99	.15	1.00	.15	1.00	.13	1.02	.11	1.00	27.42	<.001	.006
Perceptual Aberration	-.044	.96	.12	1.07	.12	1.07	.15	1.06	.08	1.02	45.16	<.001	.010

Table 2. Independent Samples T-Test for White and Black Participants

Measures	White		Black		*t*	*p*	*Cohen's D*
	M	*SD*	*M*	*SD*			
Cognitive Slippage	-.02	1.00	.11	1.03	-2.45	.015	-.13
Social Anhedonia	-.05	.96	.30	1.08	-10.97	<.001	-.36
Magical Ideation	-.04	.99	.15	1.00	-6.839	<.001	-.20
Perceptual Aberration	-.044	.96	.12	1.07	-5.560	<.001	-.17

Table 3. Independent Samples T-Test for White and Asian Participants

Measures	White		Asian		*t*	*p*	*Cohen's D*
	M	*SD*	*M*	*SD*			
Cognitive Slippage	-.02	1.00	.24	.97	-3.216	.001	-.26
Social Anhedonia	-.05	.96	.40	1.15	-9.348	<.001	-.47
Magical Ideation	-.04	.99	.24	1.08	-6.340	<.001	-.28
Perceptual Aberration	-.044	.96	.42	1.25	-9.261	<.001	-.47

Table 4. Independent Samples T-Test for White and Hispanic Participants

Measures	White		Hispanic		*t*	*p*	*Cohen's D*
	M	*SD*	*M*	*SD*			
Cognitive Slippage	-.02	1.00	-.09	.90	.682	.495	.07
Social Anhedonia	-.05	.96	.09	1.13	-2.076	.039	-.15
Magical Ideation	-.04	.99	.13	1.02	-2.809	.005	-.17
Perceptual Aberration	-.044	.96	.15	1.06	-3.174	.002	-.20

Table 5. Independent Samples T-Test for White and Multiracial Participants

Measures	White		Multi		*t*	*p*	*Cohen's D*
	M	*SD*	*M*	*SD*			
Cognitive Slippage	-.02	1.00	.18	1.10	-2.352	.020	-.20
Social Anhedonia	-.05	.96	.10	1.10	-4.225	<.001	-.16
Magical Ideation	-.04	.99	.11	1.00	-4.820	<.001	-.15
Perceptual Aberration	-.044	.96	.08	1.02	-3.859	<.001	-.13

Table 6. Independent Samples T-Test for Gender

Measures	Female		Male		*t*	*p*	*Cohen's D*
	M	*SD*	*M*	*SD*			
Cognitive Slippage	-.01	1.01	.02	.98	-1.093	.274	-.03
Social Anhedonia	-.06	.99	.06	1.00	-7.713	<.001	-.12
Magical Ideation	.01	1.00	-.02	1.00	1.722	.085	.03
Perceptual Aberration	-.05	.97	.04	1.02	-6.040	<.001	-.10

Table 7. Global Measurement Invariance Analysis

Model	X^2	Parameters	AIC	BIC	SABIC	X^2_{diff} *(df)*	*p*
Cognitive Slippage							
Freely estimated	2198.48	105	75035.84	76224.16	75668.06		
Fixed to zero	2408.76	35	75279.24	75516.91	75405.68	366.345 (70)	<.001
Social Anhedonia							
Freely estimated	2363.32	105	147838.91	149196.28	148640.14		
Fixed to zero	3007.38	35	148403.22	148674.70	148563.47	718.194(70)	<.001
Magical Ideation							
Freely estimated	215.57	56	111677.53	112391.11	112101.92		
Fixed to zero	430.48	21	111802.53	111967.21	111900.47	220.814(35)	<.001
Perceptual Aberration							
Freely estimated	136.74	49	87922.87	88542.41	88291.35		
Fixed to zero	327.95	19	88012.31	88161.31	88100.93	196.013(30)	<.001

Note. AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion; SABIC = Sample Size Adjusted Bayesian Information Criterion; *df* = degrees of freedom. Freely estimate models include parameter estimates for ethnicity and sex variables regression on the latent factor and all items. In the fixed to zero models, these parameters are constrained to zero.

Table 8. Number and Percentage of Schizotypy Scale Items Displaying DIF by Group

Group	Magical Ideation	Perceptual Aberration	Social Anhedonia	Cognitive Slippage
Gender	**4 (50%)**	**4 (57%)**	**12 (80%)**	**7 (47%)**
Race/Ethnicity				
Black	**5 (63%)**	**5 (71%)**	**8 (53%)**	1 (7%)
Asian	0	1 (14%)	4 (27%)	1 (7%)
Hispanic	1 (13%)	1 (14%)	0	1 (7%)
Multi-Racial	**4 (50%)**	2 (29%)	0	1 (7%)

Note. Bold figures represent at least 50% of items displayed DIF (either among the current set of used items or among the full scale of items)

Table 9. Uniform & Nonuniform Differential Item Functioning for Cognitive Slippage by Gender

Item	Uniform DIF			Non-Uniform DIF		
	β	*p*	ETS Δ	β	*p*	ETS Δ
1	-0.05	0.778	0.56	-0.16	0.011*	0.38
2	0.83	<.0.001*	0.12	-0.06	0.201	0.14
3	0.16	0.320	1.94	-0.11	0.020*	0.26
4	-0.04	0.836	0.38	-0.30	< 0.001*	0.71
5	0.63	< 0.001*	0.09	-0.12	0.001*	0.28
6	0.22	0.316	1.47	-0.14	0.074	0.33
7	0.69	0.004*	0.52	-0.70	0.486	1.65
8	0.38	0.072	1.62	0.31	0.758	0.73
9	-0.03	0.814	0.89	-0.04	0.370	0.09
10	0.42	0.011*	0.07	0.02	0.716	0.04
11	-0.08	0.629	0.98	0.04	0.515	0.09
12	-0.03	0.840	0.19	-0.01	0.931	0.02
13	0.32	0.057	0.07	-0.05	0.548	0.12
14	-0.17	0.437	0.75	0.15	0.118	0.35
15	0.09	0.430	0.40	-0.11	0.081	0.26

Table 10. Uniform & Nonuniform Differential Item Functioning for Social Anhedonia by Gender

Item	Uniform			Nonuniform		
	β	*p*	ETS Δ	β	*p*	ETS Δ
1	-0.70	0.000*	1.64	-0.04	0.549	0.09
2	-0.54	0.000*	1.26	-0.03	0.690	0.07
3	-0.55	0.019*	1.29	-0.06	0.543	0.14
4	-0.27	0.062	0.63	-0.09	0.148	0.21
5	0.12	0.197	0.28	-0.11	0.033*	0.25
6	0.04	0.812	0.09	-0.15	0.083	0.35
7	0.07	0.786	0.16	0.01	0.774	0.02
8	-0.25	0.008*	0.59	-0.15	0.001*	0.35
9	0.52	0.000*	1.23	-0.15	0.001*	0.35
10	-0.33	0.001*	0.79	-0.01	0.918	0.02
11	-0.73	0.000*	1.72	0.03	0.717	0.07
12	0.14	0.254	0.33	-0.20	0.008*	0.46
13	-0.44	0.017*	1.03	-0.19	0.050	0.12
14	0.15	0.071	0.35	-0.29	0.000*	0.69
15	0.21	0.006*	0.49	-0.18	0.003*	0.43

Table 11. Uniform & Nonuniform Differential Item Functioning for Magical Ideation by Gender

Item	Uniform			Nonuniform		
	β	*p*	ETS Δ	β	*p*	ETS Δ
1	0.18	0.024*	0.42	0.000	0.995	0
2	0.08	0.301	0.19	-0.032	0.582	0.08
3	0.09	0.475	0.21	0.067	0.392	0.16
4	0.35	0.008*	0.82	-0.008	0.930	0.02
5	0.15	0.104	0.35	-0.046	0.477	0.11
6	0.69	0.000*	1.62	0.071	0.483	0.17
7	0.39	0.011*	0.92	0.071	0.479	0.17
8	-0.10	0.226	0.24	0.028	0.770	0.07

Table 12. Uniform & Nonuniform Differential Item Functioning for Perceptual Aberration by Gender

Item	Uniform			Nonuniform		
	β	*p*	ETS Δ	β	*p*	ETS Δ
1	0.45	0.000*	1.06	-0.02	0.714	0.05
2	0.22	0.024*	0.52	-0.03	0.596	0.07
3	-0.16	0.139	0.38	0.13	0.056	0.31
4	0.15	0.381	0.35	0.02	0.849	0.05
5	0.40	0.000*	0.94	-0.05	0.263	0.12
6	0.63	0.000*	1.48	-0.07	0.454	0.16
7	-0.14	0.086	0.33	0.02	0.744	0.05

Table 13. Uniform Differential Item Functioning for Cognitive Slippage by Race

Item	Black			Asian			Hispanic			Multiracial		
	β	*p*	ETS Δ	β	*p*	ETS Δ	β	*p*	ETS Δ	β	*p*	ETS Δ
1	0.18	0.610	0.42	0.04	0.949	0.09	-0.23	0.648	0.59	0.03	0.949	0.07
2	0.50	0.067	1.18	0.65	0.949	1.53	-0.10	0.857	0.24	0.26	0.333	0.61
3	0.25	0.475	0.59	0.67	0.050	1.57	-1.25	0.336	2.94	-0.42	0.400	0.99
4	0.44	0.261	1.03	0.28	0.098	0.66	-0.57	0.403	1.34	-0.23	0.595	0.54
5	0.39	0.041*	0.92	-0.03	0.644	0.07	0.25	0.450	0.59	0.52	0.048*	1.22
6	0.16	0.772	0.38	-0.02	0.925	0.05	-0.47	0.629	1.10	-1.62	0.142	3.81
7	0.61	0.213	1.43	1.16	0.056	2.73	-0.78	0.495	1.83	1.39	0.207	3.27
8	0.08	0.874	0.19	0.78	0.056	1.83	-0.24	0.805	0.56	0.16	0.727	0.38
9	0.01	0.978	0.02	-0.28	0.476	0.61	-0.86	0.113	2.02	0.00	0.987	0.00
10	0.35	0.309	0.82	0.75	0.116	1.76	-0.56	0.463	1.32	-0.05	0.898	0.12
11	0.38	0.174	0.89	0.66	0.079	1.55	-0.27	0.645	0.63	-	-	-
12	-0.12	0.840	0.28	0.27	0.622	0.63	0.00	0.999	0.00	-	-	-
13	0.49	0.774	1.15	0.72	0.622	1.69	-0.17	0.999	0.40	-	-	-
14	-0.31	0.644	0.73	0.32	0.655	0.75	-1.29	0.401	3.03	-	-	-
15	0.00	0.985	0.00	0.11	0.743	0.26	0.52	0.208	1.22	-	-	-

Table 14. Nonuniform Differential Item Functioning for Cognitive Slippage by Race

Item	Black			Asian			Hispanic			Multiracial		
	β	*p*	ETS Δ	β	*p*	ETS Δ	β	*p*	ETS Δ	β	*p*	ETS Δ
1	-0.15	0.282	0.35	-0.09	0.745	0.21	-0.47	0.016	1.10	0.16	0.547	0.38
2	0.00	0.967	0	-0.20	0.140	0.47	0.01	0.958	0.02	-0.04	0.751	0.09
3	-0.04	0.749	0.09	-0.27	0.113	0.63	0.36	0.490	0.85	0.14	0.529	0.33
4	-0.25	0.126	0.59	-0.28	0.241	0.66	-0.42	0.768	0.99	-0.08	0.768	0.19
5	-0.01	0.912	0.02	-0.11	0.429	0.26	-0.13	0.377	0.31	0.15	0.292	0.35
6	0.01	0.666	0.02	-0.09	0.745	0.21	-0.11	0.758	0.26	0.67	0.156	1.57
7	0.07	0.713	0.16	-0.19	0.430	0.45	-0.04	0.929	0.09	0.46	0.298	1.08
8	0.08	0.627	0.19	-0.40	0.008*	0.94	-0.01	0.982	0.02	-0.06	0.75	0.14
9	-0.10	0.234	0.24	-0.07	0.639	0.16	0.16	0.504	0.38	0.04	0.811	0.09
10	0.04	0.769	0.09	-0.07	0.697	0.16	-0.01	0.975	0.02	-0.05	0.766	0.12
11	-0.14	0.332	0.33	-0.316	0.087	0.74	-0.27	0.361	0.63	-	-	-
12	0.14	0.461	0.33	0.01	0.971	0.02	-0.32	0.211	0.75	-	-	-
13	-0.15	0.258	0.35	0.08	0.812	0.19	-0.05	0.895	0.12	-	-	-
14	0.65	0.089	1.53	0.17	0.659	0.40	0.13	0.830	0.31	-	-	-
15	-0.05	0.716	0.12	-0.15	0.441	0.35	-0.01	0.967	0.02	-	-	-

Table 15. Uniform Differential Item Functioning for Social Anhedonia by Race

Item	Black			Asian			Hispanic			Multiracial		
	β	*p*	ETS Δ	β	*p*	ETS Δ	β	*p*	ETS Δ	β	*p*	ETS Δ
1	-0.82	0.049*	1.93	-0.27	0.572	0.63	-0.62	0.403	1.46	-0.51	0.299	1.20
2	-0.96	0.006*	2.26	-0.92	0.057	2.16	-1.28	0.136	3.01	-0.07	0.869	0.16
3	-0.13	0.737	0.31	0.02	0.966	0.05	-0.30	0.711	0.71	0.50	0.27	1.18
4	-0.85	0.015*	2.00	-0.87	0.12	2.04	-1.49	0.114	3.50	0.03	0.928	0.07
5	0.17	0.272	0.40	-0.78	0.073	1.83	-0.42	0.335	0.99	-0.35	0.217	0.82
6	-0.21	0.504	0.49	-0.46	0.417	1.08	-0.77	0.376	1.81	-0.41	0.344	0.96
7	-0.01	0.978	0.02	-0.26	0.762	0.61	-2.08	0.28	4.88	0.12	0.816	0.28
8	0.59	0.000*	1.39	0.57	0.018*	1.34	0.19	0.557	0.45	0.26	0.306	0.61
9	-0.08	0.63	0.18	-0.10	0.64	0.24	-0.51	0.194	1.20	0.34	0.069	0.80
10	-0.68	0.010*	1.60	-0.07	0.786	0.16	-1.16	0.128	2.73	-0.10	0.746	0.24
11	-1.49	0.000*	3.50	-0.51	0.244	1.20	-0.91	0.253	2.14	-0.10	0.847	0.24
12	0.05	0.845	0.18	-0.25	0.521	0.59	-0.43	0.517	1.01	0.45	0.309	1.06
13	-0.27	0.461	0.63	-1.88	0.027*	4.42	-1.52	0.196	3.57	0.09	0.861	0.21
14	-0.02	0.913	0.05	-0.40	0.251	0.94	-0.06	0.875	0.14	-0.10	0.74	0.23
15	0.29	0.056	0.68	0.46	0.037*	1.08	-0.06	0.901	0.14	0.20	0.366	0.47

Table 16. Nonuniform Differential Item Functioning for Social Anhedonia by Race

Item	Black			Asian			Hispanic			Multiracial		
	β	*p*	ETS Δ	β	*p*	ETS Δ	β	*p*	ETS Δ	β	*p*	ETS Δ
1	0.03	0.870	0.07	-0.17	0.353	0.40	0.02	0.937	0.05	0.24	0.223	0.56
2	0.87	0.386	2.04	0.08	0.666	0.19	0.26	0.462	0.61	0.26	0.126	0.61
3	-0.30	0.047*	0.70	-0.36	0.096	0.85	-0.30	0.315	0.71	-0.24	0.19	0.56
4	0.13	0.307	0.31	0.05	0.781	0.12	0.3	0.41	0.71	-0.02	0.908	0.05
5	-0.29	0.001	0.68	0.06	0.758	0.14	-0.06	0.791	0.14	0.14	0.224	0.33
6	-0.22	0.085	0.517	-0.14	0.512	0.33	0.07	0.836	0.16	-0.07	0.693	0.16
7	-0.30	0.079	0.705	0.00	0.994	0	0.77	0.3	1.81	0.00	0.994	0
8	-0.19	0.018*	0.44	-0.21	0.089	0.50	-0.14	0.467	0.33	0.10	0.367	0.23
9	-0.14	0.092	0.33	-0.34	0.003*	0.80	0.11	0.609	0.26	-0.06	0.602	0.14
10	0.17	0.169	0.40	-0.27	0.054	0.63	0.44	0.309	1.03	0.20	0.226	0.47
11	0.10	0.541	0.24	0.02	0.91	0.05	0.35	0.319	0.82	-0.04	0.815	0.09
12	-0.19	0.165	0.45	-0.21	0.243	0.49	0.12	0.748	0.28	0.15	0.577	0.35
13	-0.23	0.197	0.54	0.33	0.304	0.78	0.39	0.451	0.92	-0.01	0.967	0.02
14	-0.03	0.778	0.07	-0.05	0.789	0.12	-0.10	0.697	0.24	0.14	0.491	0.33
15	-0.17	0.121	0.40	-0.08	0.666	0.19	0.48	0.23	1.13	0.04	0.823	0.09

Table 17. Uniform Differential Item Functioning for Magical Ideation by Race

Item	Black			Asian			Hispanic			Multiracial		
	β	*p*	ETS Δ	β	*p*	ETS Δ	β	*p*	ETS Δ	β	*p*	ETS Δ
1	0.20	0.203	0.47	0.05	0.822	0.12	0.19	0.530	0.45	0.46	0.021*	1.08
2	0.97	0.000*	2.28	0.23	0.254	0.54	0.51	0.201	1.20	0.74	0.001*	1.74
3	0.35	0.095	0.82	0.21	0.522	0.49	0.74	0.032*	1.74	0.67	0.011*	1.57
4	0.90	0.000*	2.14	-0.37	0.394	0.87	0.63	0.166	1.48	0.89	0.000*	2.09
5	0.60	0.000*	1.41	0.16	0.56	0.38	0.42	0.156	0.99	0.40	0.075	0.94
6	0.76	0.001*	1.79	0.05	0.926	0.12	0.81	0.077	1.90	0.61	0.080	1.43
7	0.34	0.286	0.80	0.04	0.927	0.09	-0.36	0.704	0.85	-0.20	0.662	0.47
8	-0.57	0.023*	1.34	0.06	0.822	0.14	-0.83	0.336	1.95	-0.09	0.692	0.21

Table 18. Nonuniform Differential Item Functioning for Magical Ideation by Race

Item	Black			Asian			Hispanic			Multiracial		
	β	*p*	ETS Δ	β	*p*	ETS Δ	β	*p*	ETS Δ	β	*p*	ETS Δ
1	0.01	0.942	0.02	-0.13	0.344	0.31	-0.16	0.464	0.38	-0.10	0.396	0.23
2	-0.17	0.107	0.40	-0.17	0.226	0.40	0.18	0.555	0.42	0.04	0.737	0.09
3	-0.15	0.279	0.35	-0.00	0.998	0	-0.41	0.084	0.96	-0.12	0.427	0.28
4	-0.31	0.042*	0.73	0.08	0.767	0.19	-0.31	0.357	0.73	-0.42	0.005*	0.99
5	-0.12	0.296	0.28	0.02	0.892	0.05	-0.29	0.217	0.68	-0.06	0.655	0.14
6	-0.37	0.012*	0.87	0.49	0.136	1.15	-0.49	0.118	1.15	-0.09	0.651	0.21
7	0.30	0.172	0.71	0.10	0.694	0.24	0.73	0.275	1.72	0.39	0.13	0.92
8	0.19	0.381	0.45	-0.18	0.370	0.42	0.81	0.356	1.90	-0.07	0.784	0.16

Table 19. Uniform Differential Item Functioning for Perceptual Aberration by Race

Item	Black			Asian			Hispanic			Multiracial		
	β	*p*	ETS Δ	β	*p*	ETS Δ	β	*p*	ETS Δ	β	*p*	ETS Δ
1	-0.08	0.712	0.19	0.55	0.022*	1.29	0.50	0.112	1.18	0.17	0.379	0.40
2	-0.45	0.048*	1.06	-0.49	0.142	1.15	-0.57	0.201	1.34	0.04	0.826	0.09
3	-0.15	0.512	0.35	-0.03	0.937	0.07	0.04	0.924	-.09	-0.38	0.169	0.90
4	-1.00	0.036*	2.35	-0.59	0.361	1.39	-0.72	0.41	1.69	0.09	0.792	0.21
5	0.33	0.007*	0.78	0.15	0.496	0.35	-0.23	0.521	0.54	-0.30	0.118	0.71
6	-0.96	0.048*	2.26	-0.11	0.832	0.26	-0.87	0.323	2.04	-0.26	0.525	0.61
7	0.27	0.039*	0.63	0.06	0.784	0.14	0.07	0.836	0.16	0.17	0.333	0.40

Table 20. Nonuniform Differential Item Functioning for Perceptual Aberration by Race

Item	Black			Asian			Hispanic			Multiracial		
	β	*p*	ETS Δ	β	*p*	ETS Δ	β	*p*	ETS Δ	β	*p*	ETS Δ
1	0.08	0.498	0.19	-0.00	0.992	0	-0.34	0.047*	0.80	-0.25	0.015*	0.59
2	0.06	0.614	-0.14	0.20	0.23	0.47	-0.05	0.823	0.12	-0.26	0.016*	0.61
3	-0.00	0.983	0	0.12	0.455	0.28	-0.10	0.66	0.24	0.09	0.515	0.21
4	0.38	0.085	0.89	0.38	0.158	0.89	0.26	0.525	0.61	-0.24	0.137	0.56
5	-0.06	0.476	0.14	0.19	0.148	0.45	0.21	0.331	0.49	0.20	0.092	0.47
6	0.34	0.124	0.80	0.18	0.477	0.42	0.16	0.782	0.38	0.01	0.98	0.02
7	-0.16	0.084	0.38	-0.17	0.166	0.40	0.03	0.894	0.07	0.05	0.706	0.12

ITEMS

Magical Ideation Items

1. I have wondered whether the spirits of the dead can influence the living.
2. I have felt that I might cause something to happen just by thinking too much about it.
3. I have sometimes felt that strangers were reading my mind.
4. I have felt that there were messages for me in the way things were arranged, like in a store window
5. Things sometimes seem to be in different places when I get home, even though no one has been there.
6. I have worried that people on other planets may be influencing what happens on earth.
7. I have occasionally had the silly feeling that a TV or radio broadcaster knew I was listening to him.
8. I have sometimes been fearful of stepping on sidewalk cracks.

Perceptual Aberration Items

1. Sometimes I have had the feeling that I am united with another near me.
2. Sometimes when I look at things like tables and chairs, they seem strange.
3. Sometimes I feel like everything around me is tilting.
4. I have sometimes felt that some part of my body no longer belonged to me.
5. Sometimes part of my body seems smaller than it really is.
6. I have sometimes had the feeling that my body is decaying inside.
7. Now and then when I look in the mirror, my face seems quite different than usual.

Social Anhedonia Items

1. Although there are things, I can enjoy doing myself, I usually seem to have more fun when I do things with other people.
2. I prefer hobbies and leisure activities that do not involve other people.
3. I'm much too independent to really get involved with other people.
4. I never had really close friends in high school.
5. I feel pleased and gratified as I learn more and more about the emotional life of my friends.
6. When others try to tell me about their problems and hang-ups, I usually listen with interest and attention.
7. Just being with friends can make me feel really good.
8. There are few things more tiring than to have a long, personal discussion with someone.
9. When things are bothering me, I like to talk to other people about it.
10. When I am home alone, I often resent people telephoning me or knocking on my door.
11. In many ways, I prefer the company of pets to the company of people.
12. My relationships with other people never get very intense.
13. I don't really feel very close to my friends.
14. It made me sad to see all my high school friends go their separate ways when high school was over.
15. I find that people too often assume that their daily activities and opinions will be interesting to me.

Cognitive Slippage Items

1. My thoughts are orderly most of the time.
2. I almost always feel as though my thoughts are on a different wavelength from 98% of the population.
3. Often when I am talking, I feel that I am not making any sense.
4. My thoughts are more random than orderly.
5. The way I perceive things is much the same as the way in which others perceive them. –
6. I can usually keep my thoughts going straight.
7. My thoughts are so vague and hazy that I wish I could just reach up and pull them into place.
8. I usually feel that people understand what I say.
9. I have no difficulty controlling my thoughts.
10. My thinking often gets "cloudy" for no apparent reason.
11. I often find myself saying something that comes out completely backwards.
12. My thoughts often jump from topic to topic without any logical connection.
13. My thoughts seem to come and go so quickly that I can't keep up.
14. I often feel confused when I try to explain my ideas.
15. Usually, my thoughts aren't difficult to keep track of.

References

Anglin D. M. (2020). Advancing the study of local ethnic density and psychosis. *The lancet. Psychiatry*, *7*(4), 293–294. https://doi.org/10.1016/S2215-0366(20)30101-2

Anglin, D. M., Galea, S., & Bachman, P. (2020). Going Upstream to Advance Psychosis Prevention and Improve Public Health. *JAMA psychiatry*, *77*(7), 665–666. https://doi.org/10.1001/jamapsychiatry.2020.0142

Bao, Y., Fisher, J., & Studnicki, J. (2008). Racial differences in behavioral inpatient diagnosis: examining the mechanisms using the 2004 Florida Inpatient Discharge Data. *The journal of behavioral health services & research*, *35*(3), 347–357. https://doi.org/10.1007/s11414-008-9116-4

Barnes, A. (2004). Race, schizophrenia, and admission to state psychiatric hospitals. *Administration and Policy in Mental Health*, *31*(3), 241–252. https://doi-org.proxy.mul.missouri.edu/10.1023/B:APIH.0000018832.73673.54

Barrantes-Vidal, N., Grant, P., & Kwapil, T. R. (2015). The role of schizotypy in the study of the etiology of schizophrenia spectrum disorders. *Schizophrenia bulletin*, *41 Suppl 2*(Suppl 2), S408–S416. https://doi.org/10.1093/schbul/sbu191

Blanchard, J. J., Collins, L. M., Aghevli, M., Leung, W. W., & Cohen, A. S. (2011). Social anhedonia and schizotypy in a community sample: the Maryland longitudinal study of schizotypy. *Schizophrenia bulletin*, *37*(3), 587–602. https://doi.org/10.1093/schbul/sbp107

Blow, F. C., Zeber, J. E., McCarthy, J. F., Valenstein, M., Gillon, L., & Bingham, C. R. (2004). Ethnicity and diagnostic patterns in veterans with psychoses. *Social psychiatry and psychiatric epidemiology*, *39*(10), 841–851. https://doi.org/10.1007/s00127-004-0824-7

Cappelleri, J. C., Jason Lundy, J., & Hays, R. D. (2014). Overview of classical test theory and item response theory for the quantitative assessment of items in developing patient-reported outcomes measures. *Clinical therapeutics*, *36*(5), 648–662. https://doi.org/10.1016/j.clinthera.2014.04.006

Chapman, L. J., Chapman, J. P., Kwapil, T. R., Eckblad, M., & Zinser, M. C. (1994). Putatively psychosis-prone subjects 10 years later. *Journal of abnormal psychology*, *103*(2), 171–183. https://doi.org/10.1037//0021-843x.103.2.171

Chapman, L. J., Chapman, J. P., & Raulin, M. L. (1976). Scales for physical and social anhedonia. *Journal of abnormal psychology*, *85*(4), 374–382. https://doi.org/10.1037//0021-843x.85.4.374

Chapman, L. J., Chapman, J. P., & Raulin, M. L. (1978). Body-image aberration in Schizophrenia. *Journal of abnormal psychology*, *87*(4), 399–407. https://doi.org/10.1037//0021-843x.87.4.399

Chapman, L. J., Edell, W. S., & Chapman, J. P. (1980). Physical anhedonia, perceptual aberration, and psychosis proneness. *Schizophrenia bulletin*, *6*(4), 639–653. https://doi.org/10.1093/schbul/6.4.639

Cicero, D. C., & Kerns, J. G. (2010). Multidimensional factor structure of positive schizotypy. *Journal of personality disorders*, *24*(3), 327–343. https://doi.org/10.1521/pedi.2010.24.3.327

Cicero, D. C., Martin, E. A., & Krieg, A. (2019). Differential item functioning of the full and brief Wisconsin Schizotypy Scales in Asian, White, Hispanic, and multiethnic samples and between sexes. *Assessment*, *26*(6), 1001–1013. https://doi.org/10.1177/1073191117719509

Eack, S. M., Bahorik, A. L., Newhill, C. E., Neighbors, H. W., & Davis, L. E. (2012). Interviewer-perceived honesty as a mediator of racial disparities in the diagnosis of schizophrenia. *Psychiatric Services*, *63*(9), 875–880. https://doi-org.proxy.mul.missouri.edu/10.1176/appi.ps.201100388

Eckblad, M., Chapman, L. J., Chapman, J. P., & Mishlove, M. (1982). The Revised Social Anhedonia Scale. Department of Psychology, University of Wisconsin

Eckblad, M., & Chapman, L. J. (1983). Magical ideation as an indicator of schizotypy. *Journal of consulting and clinical psychology*, *51*(2), 215–225. https://doi.org/10.1037//0022-006x.51.2.215

Fonseca-Pedrero, E., Debbané, M., Ortuño-Sierra, J., Chan, R., Cicero, D. C., Zhang, L. C., Brenner, C., Barkus, E., Linscott, R. J., Kwapil, T., Barrantes-Vidal, N., Cohen, A., Raine, A., Compton, M. T., Tone, E. B., Suhr, J., Muñiz, J., Fumero, A., Giakoumaki, S., Tsaousis, I., … Jablensky, A. (2018). The structure of schizotypal personality traits: a cross-national study. *Psychological medicine*, *48*(3), 451–462. https://doi.org/10.1017/S0033291717001829

Fornasari, L., Picardi, A., Garzitto, M., Gigantesco, A., Sala, M., Romanò, M., Fabbro, F., & Brambilla, P. (2015). Reliability and normative data of the Perceptual Aberration Scale in an Italian juvenile general population sample. *Psychiatry research*, *228*(3), 495–500. https://doi.org/10.1016/j.psychres.2015.05.058

Giordano, G. M., Bucci, P., Mucci, A., Pezzella, P., & Galderisi, S. (2021). Gender Differences in Clinical and Psychosocial Features Among Persons With Schizophrenia: A Mini Review. *Frontiers in psychiatry*, *12*, 789179. https://doi.org/10.3389/fpsyt.2021.789179

Gooding, D. C. (2014). Cognitive slippage, psychosis-proneness, and schizotypy: A comment on Loas, et al (2013). *Psychological Reports*, *115*(2), 537–540. https://doi.org/10.2466/02.19.PR0.115c21z3

Gooding, D. C., & Iacono, W. G. (1995). Schizophrenia through the lens of a developmental psychopathology perspective. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology, Vol. 2: Risk, disorder, and adaptation.* (pp. 535–580). John Wiley & Sons.

Gooding, D. C., Ott, S. L., Roberts, S. A., & Erlenmeyer-Kimling, L. (2013). Thought disorder in mid-childhood as a predictor of adulthood diagnostic outcome: Findings from the New York High-Risk Project. *Psychological Medicine*, *43*(5), 1003–1112. https://doi.org/10.1017/S0033291712001791

Gooding, D. C., Tallant, K. A., & Hegyi, J. V. (2001). Cognitive slippage in schizotypic individuals. *Journal of Nervous and Mental Disease*, *189*(11), 750–756. https://doi.org/10.1097/00005053-200111000-00004

Gooding, D. C., Tallent, K. A., & Matts, C. W. (2005). Clinical status of at-risk individuals 5 years later: further validation of the psychometric high-risk strategy. *Journal of Abnormal Psychology*, *114*, 170–175. https://doi.org/10.1037/0021-843X.114.1.170

Goulding, S. M., McClure-Tone, E., & Compton, M. T. (2009). Associations between multiple dimensions of schizotypy and sociodemographic variables in a nonpsychiatric sample of young adults. *The Journal of nervous and mental disease*, *197*(10), 786–789. https://doi.org/10.1097/NMD.0b013e3181b975e6

Hewitt, J. K., & Claridge, G. (1989). The factor structure of schizotypy in a normal population. *Personality and Individual Differences*, *10*(3), 323–329. https://doi.org/10.1016/0191-8869(89)90105-0

Horan, W. P., Blanchard, J. J., Clark, L. A., & Green, M. F. (2008). Affective traits in schizophrenia and schizotypy. *Schizophrenia bulletin*, *34*(5), 856–874. https://doi.org/10.1093/schbul/sbn083

Kerns, J. G. (2020). Cluster A Personality Disorders. In C. W. Lejuez & K. L. Gratz (Eds.), *Handbook on Personality Disorders* (pp. 195-211). Cambridge University Press.

Kerns J. G. (2006). Schizotypy facets, cognitive control, and emotion. *Journal of abnormal psychology*, *115*(3), 418–427. https://doi.org/10.1037/0021-843X.115.3.418

Kerns, J. G., & Becker, T. M. (2008). Communication disturbances, working memory, and emotion in people with elevated disorganized schizotypy. *Schizophrenia research*, *100*(1-3), 172–180. https://doi.org/10.1016/j.schres.2007.11.005

Kim, G., DeCoster, J., Huang, C.-H., & Chiriboga, D. A. (2011). Race/ethnicity and the factor structure of the Center for Epidemiologic Studies Depression Scale: A meta-analysis. *Cultural Diversity and Ethnic Minority Psychology, 17*(4), 381–396

Kwapil T. R. (1998). Social anhedonia as a predictor of the development of schizophrenia-spectrum disorders. *Journal of Abnormal Psychology*, *107*, 558–565. https://doi.org/10.1037//0021-843x.107.4.558

Kwapil, T. R., & Barrantes-Vidal, N. (2015). Schizotypy: looking back and moving forward. *Schizophrenia bulletin*, *41 Suppl 2*(Suppl 2), S366–S373. https://doi.org/10.1093/schbul/sbu186

Kwapil, T. R., Barrantes-Vidal, N., & Silvia, P. J. (2008). The dimensional structure of the Wisconsin Schizotypy Scales: Factor identification and construct validity. *Schizophrenia Bulletin*, *34*(3), 444–457. https://doi.org/10.1093/schbul/sbm098

Kwapil, T. R., Brown, L. H., Silvia, P. J., Myin-Germeys, I., & Barrantes-Vidal, N. (2012). The expression of positive and negative schizotypy in daily life: An experience sampling study. *Psychological Medicine*, *42*(12), 2555–2566. https://doi.org/10.1017/S0033291712000827

Kwapil, T. R., Gross, G. M., Silvia, P. J., & Barrantes-Vidal, N. (2013). Prediction of psychopathology and functional impairment by positive and negative schizotypy in the Chapmans' ten-year longitudinal study. *Journal of abnormal psychology*, *122*(3), 807–815. https://doi.org/10.1037/a0033759

Kwapil, T. R., Gross, G. M., Silvia, P. J., Raulin, M. L., & Barrantes-Vidal, N. (2018). Development and psychometric properties of the Multidimensional Schizotypy Scale: A new measure for assessing positive, negative, and disorganized schizotypy. *Schizophrenia research*, *193*, 209–217. https://doi.org/10.1016/j.schres.2017.07.001

Lenzenweger M. F. (2018). Schizotypy, schizotypic psychopathology and schizophrenia. *World psychiatry : official journal of the World Psychiatric Association (WPA)*, *17*(1), 25–26. https://doi.org/10.1002/wps.20479

Li, L. Y., Cicero, D. C., Dodell-Feder, D., Germine, L., & Martin, E. A. (2021). Comparability of social anhedonia across epidemiological dimensions: A multinational study of measurement invariance of the Revised Social Anhedonia Scale. *Psychological assessment*, *33*(2), 171–179. https://doi.org/10.1037/pas0000972

Li, L. Y., Meyer, M. S., Martin, E. A., Gross, G. M., Kwapil, T. R., & Cicero, D. C. (2020). Differential item functioning of the Multidimensional Schizotypy Scale and Multidimensional Scale-Brief across ethnicity. *Psychological assessment*, *32*(4), 383–393. https://doi.org/10.1037/pas0000798

Lincoln, S., Johnson, T., Laquidara, J., Wilt, J. and Obeid, R., 2022. Increased rates of social defeat and schizotypy in racial minorities. *Personality and Individual Differences*, 186, p.111324.

Mellenbergh, G. J. (1982). Contingency Table Models for Assessing Item Bias. *Journal of Educational Statistics*, *7*(2), 105–118. https://doi.org/10.2307/1164960

Miers, T.C. & Raulin, M.L. (1987) Cognitive Slippage Scale . In K. Corcoran & J. Fishers (Eds.), Measures for clinical practice: a source book . New York : Free Press . Pp . 125 – 127 .

Minsky, S., Vega, W., Miskimen, T., Gara, M., & Escobar, J. (2003). Diagnostic patterns in Latino, African American, and European American psychiatric patients. *Archives of general psychiatry*, *60*(6), 637–644. https://doi.org/10.1001/archpsyc.60.6.637

Neal, D. J., Corbin, W. R., & Fromme, K. (2006). Measurement of alcohol-related consequences among high school and college students: application of item response models to the Rutgers Alcohol Problem Index. *Psychological assessment*, *18*(4), 402–414. https://doi.org/10.1037/1040-3590.18.4.402

Neighbors, H. W., Trierweiler, S. J., Ford, B. C., & Muroff, J. R. (2003). Racial Differences in DSM Diagnosis Using a Semi-Structured Instrument: The Importance of Clinical Judgment in the Diagnosis of African Americans. *Journal of Health and Social Behavior*, *44*(3), 237–256. https://doi-org.proxy.mul.missouri.edu/10.2307/1519777

Novick, D., Montgomery, W., Treuer, T., Moneta, M. V., & Haro, J. M. (2016). Sex differences in the course of schizophrenia across diverse regions of the world. *Neuropsychiatric disease and treatment*, *12*, 2927–2939. https://doi.org/10.2147/NDT.S101151

Nunnally, J.C. (1978). Psychometric theory (2nd ed.). New York: McGraw-Hill.

Olbert, C. M., Nagendra, A., & Buck, B. (2018). Meta-analysis of Black vs. White racial disparity in schizophrenia diagnosis in the United States: Do structured assessments attenuate racial disparities?. *Journal of abnormal psychology*, *127*(1), 104–115. https://doi.org/10.1037/abn0000309

Perry, B. L., Neltner, M., & Allen, T. (2013). A paradox of bias: Racial differences in forensic psychiatric diagnosis and determinations of criminal responsibility. *Race and Social Problems*, *5*(4), 239–249. https://doi.org/10.1007/s12552-013-9100-3

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., & Waller, N. G. (2009). Item response theory and clinical measurement. *Annual review of clinical psychology*, *5*, 27–48. https://doi.org/10.1146/annurev.clinpsy.032408.153553

Schwartz, R. C., & Blankenship, D. M. (2014). Racial disparities in psychotic disorder diagnosis: A review of empirical literature. *World journal of psychiatry*, *4*(4), 133–140. https://doi.org/10.5498/wjp.v4.i4.133

Sharpley, M. S., & Peters, E. R. (1999). Ethnicity, class and schizotypy. *Social psychiatry and psychiatric epidemiology*, *34*(10), 507–512. https://doi.org/10.1007/s001270050168

Selten, J. P., van der Ven, E., Rutten, B. P., & Cantor-Graae, E. (2013). The social defeat hypothesis of schizophrenia: an update. *Schizophrenia bulletin*, *39*(6), 1180–1186. https://doi.org/10.1093/schbul/sbt134

Straub, K. T., & Kerns, J. G. (2021). Positive schizotypy, maladaptive openness, and openness facets. *Personality disorders*, *12*(1), 51–58. https://doi.org/10.1037/per0000407

Swaminathan, H., & Rogers, H. J. (1990). Detecting differential item functioning using logistic regression procedures. *Journal of Educational Measurement*, *27*(4), 361–370. https://doi-org.proxy.mul.missouri.edu/10.1111/j.1745-3984.1990.tb00754.x

Tandon, R., Nasrallah, H. A., & Keshavan, M. S. (2009). Schizophrenia, "just the facts" 4. Clinical features and conceptualization. *Schizophrenia research*, *110*(1-3), 1–23. https://doi.org/10.1016/j.schres.2009.03.005

Winterstein, B. P., Ackerman, T. A., Silvia, P. J., & Kwapil, T. R. (2011). Psychometric properties of the Wisconsin Schizotypy Scales in an undergraduate sample: Classical test theory, item response theory, and differential item functioning. *Journal of Psychopathology and Behavioral Assessment*, *33*(4), 480–490. https://doi.org/10.1007/s10862-011-9242-9

Winterstein, B. P., Silvia, P. J., Kwapil, T. R., Kaufman, J. C., Reiter-Palmon, R., & Wigert, B. (2011). Brief assessment of schizotypy: Developing short forms of the Wisconsin Schizotypy Scales. *Personality and Individual Differences*, *51*(8), 920–924. https://doi.org/10.1016/j.paid.2011.07.027

Winterstein, B. P., Willse, J. T., Kwapil, T. R., & Silvia, P. J. (2010). Assessment of score dependability of the Wisconsin Schizotypy Scales using generalizability analysis. *Journal*

of Psychopathology and Behavioral Assessment, *32*(4), 575–585.

https://doi.org/10.1007/s10862-010-9181-x