June 2023

An Evaluation of Measurement Invariance of DSM-5 Borderline Personality Disorder Criteria Across Heterosexual, Lesbian, Gay, and Bisexual Adults

E. Elisa Carsten
University of South Florida

An Evaluation of Measurement Invariance of DSM-5 Borderline Personality Disorder Criteria Across Heterosexual, Lesbian, Gay, and Bisexual Adults

by

E. Elisa Carsten

A thesis submitted in partial fulfilment
of the requirements for the degree of
Master of Arts
Department of Psychology
College of Arts and Sciences
University of South Florida

Co-Major Professor: Marina Bornovalova, Ph.D.
Co-Major Professor: Jennifer Bosson, Ph.D.
Jonathan Rottenberg, Ph.D.
Brenton Wiernik, Ph.D.

Date of Approval:
June 9, 2023

Keywords: Differential item functioning, personality pathology, sexual orientation, NESARC-III

Table of Contents

List of Tables

List of Figures

Abstract

There is a robust, yet poorly understood relationship between non-heterosexual orientation and borderline personality disorder (BPD), with lesbian, gay, and bisexual individuals evidencing greater BPD symptoms compared to heterosexual individuals. Recent evidence suggests possible bias in BPD diagnostic criteria leading to greater endorsement among sexual minority individuals, which hinders researchers' ability to make valid group comparisons. The present study utilized an epidemiological sample of 35,995 men and women to evaluate the extent of differential item functioning (DIF) among BPD criteria across sexual orientation groups using a multiple indicators multiple causes (MIMIC) approach. All criteria except affective instability and emptiness indicated DIF for at least one sexual minority focal group, although those each demonstrated DIF in sensitivity analyses. DIF was largely nonuniform and although no clear pattern emerged across all models, DIF was most consistently indicated for suicidality, efforts to avoid abandonment, and impulsivity. Contrary to predictions, DIF was mostly nonuniform with greater item discrimination for focal groups compared to their heterosexual counterparts. Consistent with predictions, all DIF was explained by perceived minority stressors. Finally, all estimated effect sizes were small, suggesting that DIF was not practically meaningful and unlikely to impact the validity of group comparisons for BPD symptoms across heterosexual and non-heterosexual men and women.

Chapter One:

Introduction

Borderline personality disorder, or BPD, is a highly stigmatized and functionally impairing illness associated with substantial psychosocial dysfunction (Gunderson et al., 2011). Lesbian, gay, and bisexual (LGB) individuals are more likely than their heterosexual counterparts to be diagnosed with BPD (Grant et al., 2011; Rodriguez-Seijas et al., 2020). While some theories posit that these diagnostic disparities stem from true differences in risk for psychopathology (Bailey, 2020; Meyer, 2003; Meyer, 2021), others have raised concerns of provider bias, as well as bias among the diagnostic criteria upon which individuals are evaluated (Eubanks-Carter & Goldfried, 2006; Rodriguez-Seijas et al., 2020; Rodriguez-Seijas et al., 2021). One potential contributing factor that has yet to be investigated is the possibility of measurement bias. Valid comparison across groups hinges on the assumption that measurement tools assess the same construct in each group and lack of functional equivalence of indicators across groups precludes meaningful interpretation of mean-level differences (Borsboom, 2006). Unsubstantiated assumptions of measurement invariance of BPD diagnostic criteria across sexual orientation groups therefore hinder research progress and may also lead to misdiagnosis of BPD, which may in turn affect treatment outcomes (Campbell et al., 2020; Cochran et al., 2003). Thus, a better understanding of how BPD criteria function in sexual minority individuals may both bolster research progress and enhance treatment efficacy for this population. The present study

seeks to fill this gap by evaluating whether BPD diagnostic criteria are invariant across sexual orientation groups.

Borderline Personality Disorder

Borderline personality disorder (BPD) is a serious and long-lasting condition characterized by severe emotion dysregulation, intense fear of abandonment, turbulent interpersonal relationships, identity disturbances, and impulsivity (Grant et al., 2008). It is often severely functionally impairing and is linked with increased risk for suicide and substance use (Black et al., 2004). BPD is best understood as a moderately heritable diathesis-stress disorder, wherein biologically based temperamental vulnerabilities (e.g., affective instability, impulsivity, interpersonal instability) interact with environmental stressors, such as emotional abuse (Bornovalova et al., 2006; Kendler et al., 2011; Linehan, 1993). Under the predominant conceptualization of personality disorders as extreme, maladaptive variants of normative personality traits, BPD is understood as a constellation of high neuroticism, low conscientiousness, and low agreeableness (Samuel & Widiger, 2008). Lifetime prevalence of BPD is between 0.7% and 2.7% in the general population and BPD is present among 10-12% of outpatient and 22% of inpatient clinical populations (Ellison et al., 2018; Trull et al., 2010). In clinical populations, BPD occurs at higher rates among women than men (Zanarini et al., 2017), although gender differences are smaller or nonexistent in epidemiological samples (Trull et al., 2010). Due in part to its negative interpersonal effects and association with poor treatment outcomes (Gunderson et al., 2011), BPD is a highly stigmatized disorder often met with therapeutic nihilism among treatment providers, who may be less willing to treat an individual diagnosed with BPD relative to other forms of psychopathology (Campbell et al., 2020).

Links Between BPD and Non-heterosexuality

Research consistently links non-heterosexuality to BPD in community and clinical samples. Despite sexual minority individuals making up only a small percentage of the U.S. population (approximately 5.2%; Gallup, 2020), rates of non-heterosexuality range from 14% to 57% in BPD psychiatric samples, significantly higher than rates among other psychiatric disorders (Molina et al., 2002; Paris et al., 1995; Reich & Zanarini, 2008; Singh et al., 2011; Zanarini et al., 2020; Zubenko et al., 1987). Further, when compared to their heterosexual counterparts, sexual minority patients in psychiatric settings are more likely to be diagnosed with BPD (Grant et al., 2011; Rodriguez-Seijas et al., 2020). Sexual minority individuals in community samples also demonstrate greater BPD symptoms than their heterosexual peers (Marshal et al., 2013; Reuter et al., 2016; Rodriguez-Seijas et al., 2021). Within sexual minority subgroups, bisexual individuals are significantly more likely than gay men and lesbian women to receive a BPD diagnosis (Rodriguez-Seijas et al., 2020) and score significantly higher on BPD measures in community samples (Kerridge et al., 2017; Marshal et al., 2013; Reuter et al., 2016). This is consistent with the pattern found for other mental illnesses (King et al., 2008; Marshal et al., 2008; McCabe et al., 2009; Mustanski et al., 2014; Ross et al., 2018; Salway et al., 2019).

Proposed Explanations for Mental Health Disparities

Several possible explanations exist for the observed disparities. First, results from these studies may reflect true mean-level differences in latent personality pathology between heterosexual and sexual minority individuals due to sociocultural factors, such as exposure to stigma and discrimination (Meyer, 2003; 2021) or shared biological and genetic pathways (Bailey, 2020). It is also possible that factors beyond a sexual minority individual's true level of BPD may influence the scores they receive. Methodological issues, such as a diagnostic bias

among clinicians or bias among diagnostic criteria, may artificially inflate scores and exacerbate diagnostic disparities (Eubanks-Carter & Goldfried, 2006; Rodriguez-Seijas et al., 2020; 2021). Each of these possibilities will be explored, followed by an empirical test of measurement invariance of DSM-5 BPD criteria across sexual orientation.

Elevated Risk Due to Minority Stress Processes

Most research on sexual minority mental health disparities has followed the minority stress model (Meyer, 1995; 2003; Meyer et al., 2021), which theorizes that increased experiences of social stigma, prejudice, and discrimination lead to higher rates of "stress-sensitive disorders". This theory has considerable support, with greater mental health problems among sexual minorities linked to identity concealment and internalized homonegativity (Mereish et al., 2017; Pachankis, 2007), parental rejection due to sexual orientation (Kessler et al., 2012), and self-reported sexual orientation-based harassment and discrimination (Eaton, 2014). Quasi-experimental studies (Hatzenbuehler et al., 2009; Meyer, 2021; Raifman et al., 2017; Raifman et al., 2018) and longitudinal studies (Eldahan et al., 2016; Pachankis et al., 2018) have also linked changes in minority stressors with subsequent changes in sexual minority mental health. While most of this research has focused on more common mental health problems, such as depression and anxiety, sexual orientation-related parental rejection is linked to personality pathology specifically (Davis & Anderson, 2021).

In line with the minority stress model, the *dual marginalization* theory posits that bisexual individuals experience greater minority stress due to ostracization from both heterosexual and gay/lesbian individuals (Ochs, 1996). Many stereotypes about bisexuality persist, including beliefs that bisexuality is a temporary stepping stone to a gay or lesbian identity, and that bisexual individuals are sexually confused, sexually promiscuous, and more

likely than those of other sexual orientations to cheat in romantic relationships (Alarie & Gaudet, 2013; Bostwick & Hequembourg, 2014; Brewster & Moradi, 2010; Dodge et al., 2016; Hayfield et al., 2014; Mohr & Rochlen, 1999). Bisexual individuals report lower connectedness and feelings of exclusion from LGBTQ spaces (Balsam & Mohr, 2007; Hayfield et al., 2014). Scholars suggest these experiences may lead to greater levels of distress and risk for psychopathology broadly (Feinstein & Dyar, 2017; Ochs, 1996; Ross et al., 2010), as well as BPD specifically (Reuter et al., 2016).

However, the theorized mechanism through which minority stress increases symptoms of psychopathology (a chronic stress process) is inconsistent with research findings on the mechanism through which BPD, a putatively diathesis-stress disorder (Linehan, 2003; Crowell et al., 2009), develops. Importantly, environmental stressors alone are not sufficient for the development of BPD. Thus, increased exposure to environmental stressors via minority stress alone fails to explain the high rates of BPD among sexual minority individuals. It may therefore be more likely that minority stressors account for the presence of some BPD features among sexual minority individuals (e.g., suicidality; discussed further below), but not prototypical BPD itself.

Elevated Risk Due to Biological and Genetic (Common Cause) Factors

A separate line of sexual minority psychopathology research focuses on potential biological and genetic explanations (Bailey, 2020). Unlike the minority stress model, which is agnostic with respect to the etiology of non-heterosexuality (Meyer, 2021), biological theories postulate that a shared biological or genetic "common cause" of both non-heterosexuality and mental illness largely explains the observed mental health disparities between heterosexual and non-heterosexual groups (Bailey, 2020). Some (e.g., Bailey, 2020) have proposed direct or

indirect shared genetic influences across sexual orientation and latent transdiagnostic factors (i.e., traits that cut across most forms of psychopathology), citing genetic correlations of non-heterosexuality with neuroticism, depression and anxiety, and substance use (e.g., Ganna et al., 2019). According to these scholars, these factors may simultaneously increase risk for mental health problems and produce a disposition toward perceiving the environment as discriminatory, as well as a propensity for behaviors that elicit greater negative experiences (Bailey, 2020). Actual exposure to stigma is suggested to play a lesser, though not inconsequential role in the development of mental health problems and may constitute a consequence rather than a cause.

This starkly contrasts minority stress theories, which propose that actual experiences of rejection precede and contribute to development and maintenance of rejection sensitivity and other mental health problems (Feinstein, 2020). The notion that societal stigma and marginalization are largely imagined or elicited by the individual is also inconsistent with evidence of negative attitudes about non-heterosexuality in national probability studies (Glick et al., 2015), as well as myriad empirical evidence of structural stigma, which is unaffected by self-report bias (Meyer et al., 2021). Moreover, leading etiological theories of BPD suggest a core role of neuroticism as an underlying temperamental vulnerability (Samuel et al., 2013) and as such, under a common cause explanation, sexual minority individuals should demonstrate higher neuroticism than heterosexual individuals on average. While neuroticism appears to be higher among gay men compared to heterosexual men, meta-analytic results indicate no differences in neuroticism between heterosexual and lesbian women, nor heterosexual and bisexual individuals (Allen & Robson, 2020). Finally, diagnostic disparities persist even after controlling for transdiagnostic factors (Reuter et al., 2016; Rodriguez-Seijas et al., 2021). Biological and genetic factors are therefore unlikely to explain elevated rates of BPD among sexual minority

individuals. However, another possibility is that the observed patterns of BPD prevalence do not accurately reflect true elevations of BPD features among sexual minority individuals and may instead be artificially inflated by some form of methodological bias.

Elevated Risk Due to Provider Bias

Elevated BPD rates may be partially explained by diagnostic bias among health providers. Similar to racial diagnostic bias, which refers to clinicians' unsubstantiated judgments about patients based on their race (Feisthamel et al., 2009), provider bias based on sexual orientation involves clinicians making unwarranted judgments about the degree to which a person exhibits BPD symptoms on the basis of their sexual orientation. In one study, clinical psychologists were given one of several case vignettes of clients who had multiple, risky, sexual encounters, along with other behaviors that resembled BPD symptoms. Symptoms were identical across vignettes, but the gender of the client and their sexual partners differed. Male clients perceived by clinicians to be gay or bisexual (based on the gender of their sexual partners) were more likely to receive a BPD diagnosis (61%) compared to males perceived to be heterosexual (36%; Eubanks-Carter & Goldfried, 2006).

In another study, disparities between sexual minority and heterosexual patients in a partial hospitalization clinic were highest when BPD was assessed using unstructured interviews (which, of the three approaches tested, rely most on clinician impressions), followed by structured interviews, with lowest disparities for scores on a self-report trait-based measure of BPD (Rodriguez-Seijas et al., 2020). Using the self-report measure, disparities persisted only for bisexual individuals. The authors propose that this pattern may potentially stem from clinicians' lack of cultural familiarity with sexual minority populations, such as lack of awareness that

unprotected sex may occur at higher rates among sexual minority populations due to unassertiveness rather than prototypical impulsivity in BPD (Rodriguez-Seijas et al., 2020).

An inconsistent history of criteria for BPD in the Diagnostic and Statistical Manual of Mental Disorders (DSM) may contribute to possible provider bias, with previous editions listing uncertainty about sexual orientation as a target domain within which the BPD identity disturbance criterion may manifest (American Psychiatric Association [APA], 1987). It is plausible that a decades-long association of BPD with non-heterosexuality influences practitioners' perceptions of sexual minority individuals, leading them to expect and more readily perceive BPD pathology among sexual minority clients (Neacsiu et al., 2016).

Elevated Risk Due to Criterion Bias (Measurement Non-invariance)

Investigation into the causes of disparities in BPD is limited by unsubstantiated assumptions that diagnostic criteria are functionally equivalent across sexual orientation groups and free of systematic group-based bias. Criterion bias in the diagnosis of mental illness is defined as any systematic deviation from an expected value on the criteria listed in a construct's official nomenclature (i.e., the DSM-5; Widiger & Spitzer, 1991). Recently, researchers have questioned whether criterion bias may be present in DSM-5 BPD diagnostic criteria and suggested that minority-specific factors may partially account for the variance in BPD features among sexual and gender minorities, citing overlap in constructs strongly tied to each (Goldhammer et al., 2019; Rodriguez-Seijas et al., 2021).

Rodriguez-Seijas and colleagues (2021) evaluated rates of endorsement of BPD criterion items by sexual orientation in a U.S. epidemiological sample. They found that sexual minority individuals demonstrated significantly higher lifetime BPD prevalence rates compared to heterosexual individuals (19.8% vs. 10.88%), with an overall odds ratio (OR) of 1.93, and

endorsed 15 of the 18 criterion items at higher rates. Importantly, different diagnostic algorithms have been used in BPD research that vary with respect to the level of distress or functional impairment required for a criterion item to be coded as "present" or not. Under the strictest classification system, each criterion item is not marked "present" unless the individual endorses both a) the item itself and b) distress or functional impairment due to the item. Under more relaxed classification systems that require less distress and impairment, it is substantially easier for individuals to meet the diagnostic threshold (Trull et al., 2010). In Rodriguez-Seijas and colleagues' study, after the stricter distress/impairment requirement was added, lifetime prevalence was reduced to 1.56% and 1.1% among sexual minority and heterosexual individuals respectively, resulting in a nonsignificant OR.

At the criterion item level, 8 items demonstrate significantly higher odds of endorsement among sexual minority than among heterosexual individuals: efforts to avoid abandonment, identity disturbance, impulsivity, suicidality, difficulty controlling anger, and paranoid ideation. After controlling for differences in latent internalizing and externalizing psychopathology (e.g., anxiety, substance use disorders), higher likelihood of endorsement disappeared for all but four items indexing suicidality (items 23 and 24 in Table 1) and impulsivity (items 25 and 26), suggesting that most disparities can be explained by transdiagnostic factors. After applying the most stringent distress/impairment requirement, significant differences in endorsement across sexual orientation remained only for items 24, 25, and 26, suggesting that the suicidality and impulsivity criteria may constitute a bias in BPD diagnosis.

It is also plausible that, instead of constituting a causal pathway to BPD development as suggested in the minority stress model, perhaps chronic minority stressors may produce traits and behaviors that appear phenotypically similar to BPD features but do not actually reflect the

essence of BPD. For instance, suicidality, a BPD criterion characterized by suicidal behaviors and NSSI, may be more readily endorsed by sexual minority individuals as a result of minority stressors (and not because of BPD). Some sexual and gender minority youth also explicitly cite sexual orientation-related self-punishment, self-hatred, and coping with discrimination as motivations for engaging in NSSI (Alexander & Clare, 2004; Nickels et al., 2012; Scourfield et al., 2008) and minority stress concurrently predicts sexual minority NSSI in ecological momentary assessment research (Fehling, 2019). As Rodriguez-Seijas and colleagues note, some individuals engage in NSSI as an emotion regulation strategy (Klonsky & Muehlencamp, 2007) and this may be the case for sexual minorities in response to minority stressors. This may help explain higher rates of BPD diagnoses among bisexual individuals, who report the greatest levels of both NSSI (Batejan et al., 2015) and minority stressors (Fehling, 2019; Mereish et al., 2017; Puckett et al., 2016).

Efforts to avoid abandonment may constitute another BPD criterion that is relatable to sexual minority individuals due to chronic minority stressors rather than underlying personality pathology. Rejection sensitivity, defined as a tendency to expect, readily perceive, and exhibit intense negative reactions to cues that one may be rejected (Feinstein, 2017), is conceptually similar to fear of abandonment. It is common among both individuals with BPD (Cavicchioli & Maffei, 2020) and sexual minority individuals (Dyar et al., 2018; Pachankis et al., 2007). Indeed, links between rejection and mental health disparities among sexual minority individuals have been so robust that rejection sensitivity has been proposed as an extension of the minority stress framework (Feinstein, 2020). Rejection sensitivity among sexual minority individuals is theorized to result from a combination of repeated experiences of rejection and discrimination, as well as minority-specific internal processes such as sexual identity concealment, anxious

expectations of sexual orientation-related rejection, and internalized heterosexism. Even if minority stressors indeed contribute to a greater disposition to over-perceive rejection threat cues over time, this fundamentally represents a separate pathway through which fear of abandonment develops among sexual minority individuals. As such, items within this criterion domain may be susceptible to DIF.

Elevated endorsement of impulsivity-related BPD criteria may also constitute bias. This criterion is typically assessed based on endorsement of behaviors that are putatively impulsive in nature, which may in some cases fail to adequately capture the trait itself. Non-heterosexuality is linked to impulsive behaviors, such as risky sexual behavior (e.g., unprotected sex, sex with casual or unknown partners; Mojola & Everett, 2012; Mustanski et al., 2014; Salway et al., 2019). Regardless of health consequences, normalization of these types of sexual behaviors within LGBTQ communities may lead to greater engagement in behaviors associated with impulsivity independent of psychopathology. If a majority of sexual minority individuals endorse the item, "Have you gotten into sexual relationships quickly/without thinking about consequences?" then this item may capture a culture-specific behavioral norm rather than a unique predictor of impulsivity as a trait, thus exhibiting criterion bias.

Identity disturbance, a feature of BPD reflected by inconsistencies in one's goals, values, relationships, and behavior, and lack of a coherent life narrative (Erikson, 1968), may exhibit measurement non-invariance for several reasons. First, as Rodriguez-Seijas and colleagues (2020) suggest, the need for situational identity concealment to avoid discrimination may inflate identity disturbance scores. For some, this may mean deliberately shifting their vocal tone (e.g., among gay men, to sound more masculine) and concealing gender-atypical behavior (Pachankis, 2007). Consider the item "Have you been so different with different people or in different

situations that you sometimes don't know who you really are?". Such an item may be endorsed by sexual minority individuals who must conceal their sexual identity in certain settings (e.g., an unaccepting work environment). Such individuals may feel as though they are different people in different settings, leading to confusion and self-doubt. If this is the case, then identity disturbance criterion items may not capture the essence of BPD identity disturbance in sexual minority individuals and may instead reflect a response to pervasive stigma and nonacceptance.

Second, the process of navigating a stigmatized identity may mimic the appearance of pathologically distorted self-concept. Virtually all models of sexual minority identity development and integration include uncertainty or questioning of sexual orientation identity as a key developmental milestone (Cass, 1979; Yarhouse & Tan, 2004; see Eliason, 1996 for review), at which point individuals may experience their same-sex attracted identity as being inconsistent with their previous self-concept, leading to identity-related stress (Cass, 1979; Davis & Anderson, 2021; Eliason, 1996; Rosario et al., 2011). This shift in self-concept may lead sexual minority individuals even at low levels of BPD traits to endorse criterion items in this domain, such as the item asking, "Have you all of a sudden changed your sense of who you are and where you are headed?" Stigma may increase distress as one reaches developmental milestones and begins disclosing their sexual orientation to others, a consequence that may be especially pernicious among bisexual individuals who face additional bisexual-specific stigmas (e.g., that bisexuality is not real). Sexual orientation identity integration longitudinally predicts greater psychosocial adjustment (Rosario et al., 2011), further suggesting that identity problems may represent transient difficulties associated with minority-specific developmental milestones.

Identity disturbance may therefore reflect a temporary and non-pathological phase of identity development – precisely the type of situation-specific presentations that must be

distinguished from true BPD pathology, according to the DSM-5 (APA, 2013). While the DSM-5 explicitly states that individuals exhibiting conflicts about sexual orientation "may transiently display behaviors that misleadingly give the impression of [BPD]" (APA, 2013), the descriptive text also includes "changes in…sexual identity" as an example of a shift in self-image in the very same section, creating ambiguity and plausible confusion among practitioners surrounding the nature of the relationship between sexual identity and BPD pathology. In this case, rather than constituting a diagnostic bias due to stereotypical beliefs held by the clinician, diagnostic disparities may instead reflect issues with the characterization of the criterion itself.

If criteria are more readily endorsed by sexual minority individuals due to shared qualities unrelated to latent BPD, then these criteria function differently across sexual orientation groups and thus violate the assumption of measurement invariance, presenting validity issues (Borsboom, 2006). Variance in the likelihood of item endorsement based on group membership is referred to as *differential item functioning*, or DIF (Woods et al., 2009). In clinical settings, DIF may lead to inaccurate diagnoses (Cicero & Ruggero, 2021). Practically, unless BPD criteria are revised, the only way to account for DIF in clinical assessment would be to specify different diagnostic cutoffs for different groups (Cicero & Ruggero, 2021). Modifications to scoring algorithms of BPD measures used in research settings would also likely be necessary to progress the scientific understanding of sexual minority mental health and avoid misdiagnosis.

Few studies have tested measurement invariance of psychiatric disorder symptoms across sexual orientation groups, although a handful of studies have found evidence of non-invariance for a measure of depression (Birnholz & Young, 2012; Gomez & McLaren, 2015; Radusky et al., 2019). Heteronormative bias in item wording has also been documented in non-measurement invariance studies. In a prime example, the item stating *"I have difficulty talking to attractive*

persons of the opposite sex" on the Social Interaction Anxiety Scale (Mattick & Clarke, 1998) functions differently for non-heterosexual individuals, presumably because individuals who experience no attraction to the other sex do not face the same threat of rejection from such individuals relative to their heterosexual counterparts (Lindner et al., 2013). To date, no known study has evaluated whether DIF exists across sexual orientation groups for BPD criteria.

In summation, elevations on BPD features among sexual minority individuals may reflect true disparities in overall BPD resulting from the cumulative psychological toll of minority stress processes and/or dispositional tendencies. They may also stem in part from a lack of cultural sensitivity among practitioners. Alternatively, or—perhaps more likely, *in addition*—sexual minority individuals may more readily endorse BPD criteria independent of BPD pathology due to the unique experiences of this population, which may lead to traits and behaviors that appear similar to features of BPD (e.g., suicidality, efforts to avoid abandonment, impulsivity, and identity problems). This leads to concerns that criterion indicators may function differently across sexual minority and heterosexual individuals (i.e., DIF), constituting criterion bias. It is crucial that any potential differences in the functioning of BPD criteria across groups are identified and accounted for to make valid comparisons, improve diagnostic accuracy, and maximize treatment efficacy for sexual minority individuals. A necessary next step will therefore be to directly test the assumption of measurement invariance.

By fitting a series of multiple indicators multiple causes (MIMIC; Muthen, 1989) models testing for DIF while accounting for self-reported perceived minority stressors, the current study will test whether sexual minority identity systematically predicts lower thresholds for BPD criterion item endorsement regardless of overall BPD. Importantly, the current study's results will not be sufficient for concluding that minority stressors account for elevated BPD criterion

endorsement among sexual minority individuals. However, any presence of DIF regardless of its relationship to minority stress warrants further investigation. DIF must be accounted for in order to make valid group comparisons and continue to make progress in this field of research. This preliminary investigation that incorporates a potential role of minority stress in predicting DIF may inform future research on whether the role of minority stressors warrants additional investigation.

Research Questions and Hypotheses

Research Question 1: Are DSM-5 borderline personality disorder symptoms invariant across self-identified lesbian, gay, bisexual, and heterosexual men and women?

Hypothesis 1a: The efforts to avoid abandonment criterion (items 7-10) will show evidence of DIF, with lower item thresholds for sexual minority men and women.

Hypothesis 1b: The identity disturbance criterion (items 14-17) will show evidence of DIF, with lower item thresholds for sexual minority men and women.

Hypothesis 1c: The suicidality criterion (items 23-24) will show evidence of DIF, with lower item thresholds for sexual minority men and women.

Hypothesis 1d: The impulsivity criterion (items 25-30) will show evidence of DIF, with lower item thresholds for sexual minority men and women.

Analyses for DIF among items in all other criterion domains will be purely exploratory.

Research Question 2: Is DIF of criterion items indexing suicidality and efforts to avoid abandonment across self-identified lesbian, gay, bisexual, and heterosexual men and women partially explained by perceived minority stress?

Hypothesis 2a: DIF among criterion items indexing efforts to avoid abandonment will no longer be statistically significant after accounting for perceived minority stress.

Hypothesis 2b: DIF among criterion items indexing suicidality will no longer be statistically significant after accounting for perceived minority stress.

Hypotheses 1a through 1d are considered to be supported if DIF is indicated for suicidality, efforts to avoid abandonment, impulsivity, and/or identity disturbance. Such results indicate that BPD criterion items in the predicted domains perform differently across individuals as a function of sexual orientation. In order to establish overarching support for any given hypothesis, the corresponding criterion must demonstrate DIF in at least three of the five group comparisons. Hypotheses 2a and 2b are considered to be supported if DIF for efforts to avoid abandonment and suicidality is no longer statistically significant after perceived minority stress is statistically accounted for. Again, overarching support for a hypothesis will only be established if a criterion no longer demonstrates DIF after accounting for perceived minority stress in at least three of the five group comparisons. If any DIF across sexual orientation groups cannot be accounted for by perceived minority stress, then results will be taken to suggest that criterion bias in BPD diagnosis exists for sexual minority individuals regardless of the degree to which they report having previously experienced minority stressors. If this is found to be the case, then this indicates that the effects of sexual orientation on item endorsement are not limited to those who report having experienced minority stress. Analyses for the influence of perceived minority stress among DIF-flagged items in all other criterion domains were purely exploratory.

Table 1

NESARC-III Borderline Personality Disorder Criterion Item Endorsement

Criterion Item	**Total Endorsed (*N*)**	**Total Endorsed (%)**
Unstable Relationships	**5,203**	**14.50%**
1) Have you usually gotten attached to people very quickly?	1,422	3.97%
2) Have your relationships with people you really care about have had lots of extreme ups and downs?	3,170	8.83%
3) Have you often started out thinking that someone was a great person only to be disappointed when they didn't live up to your expectations?	3,620	10.09%
Affective Instability	**3,941**	**10.95%**
4) Have you often become very sad, anxious or angry over little things?	2,611	7.27%
5) Have other people have often wondered why you get so upset so easily?	2,604	7.27%
6) Have you had a lot of sudden mood changes?	2,178	6.07%
Efforts to Avoid Abandonment	**2,934**	**8.15%**
7) Have you needed people close to you to reassure you that they would never leave you?	1,509	4.20%
8) Have you put a lot of time and effort into doing things to keep someone from leaving you?	1,787	4.98%
9) Have you often become frantic when you thought that someone you really cared about was going to leave you?	1,895	5.28%
10) Have you gone to extremes to keep people from leaving you?	1,002	2.79%

Table 1

(Continued)

Anger	**3,704**	**10.29%**
11) Have you often had temper outbursts or gotten so angry that you lose control?	2,146	5.97%
12) Have you hit people or thrown things when you got angry?	2,338	6.50%
13) Have even little things have made you angry or have you had difficulty controlling your anger?	2,181	6.07%
Unstable Sense of Self	**2,366**	**6.57%**
14) Have there been lots of sudden changes in your personal goals, career plans, religious beliefs, or other important aspects of your life?	1,511	4.21%
15) Have you been so different with different people or in different situations that you sometimes don't know who you really are?	821	2.29%
16) Has your sense of who you are often changed depending on the situation or whom you are with?	882	2.46%
17) Have you all of a sudden have changed your sense of who you are and where you are headed?	950	2.65%
Chronic Emptiness	**2,183**	**6.07%**
18) Have you often felt like your life had no purpose or meaning?	1,533	4.27%
19) Have you often felt empty inside?	1,819	5.07%
Paranoid Ideation	**2,410**	**6.70%**
20) Have you often felt that you weren't real when under a lot of stress?	804	2.24%
21) Have you often felt like you were outside your body when under a lot of stress?	629	1.75%
22) Have you felt suspicious or distrustful of other people when under a lot of stress?	2,076	5.78%

Table 1

(Continued)

Suicidality	**1,165**	**3.24%**
23) Have you cut, burned, or scratched yourself on purpose when under a lot of stress?	396	1.10%
24) Have you tried to hurt or kill yourself, or threatened to do so?	1,052	2.93%
Impulsivity	**2,705**	**7.52%**
25) Have you gotten into sexual relationships quickly or without thinking about consequences?	1,653	4.61%
26) Have there been periods of your life when you often spent too much money while shopping or gambling?	2,277	6.33%
27) Have you had periods in your life when you drank a lot more or used a lot more drugs than you meant to?	2,569	7.15%
28) Have you had periods in your life when you often took too many risks when driving?	1,303	3.63%
29) Have you often done things impulsively, not caring about the consequences?	1,974	5.50%
30) Have you often engaged in reckless behavior without thinking about how dangerous it could be?	1,340	3.73%
BPD Diagnosis	**1,978**	**5.50%**

Chapter Two:

Method

Participants

Participants were 36,309 non-institutionalized civilian U.S. adults aged 18-90 years old who participated in Wave 3 of the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC-III), sponsored by the National Institute on Alcohol Abuse and Alcoholism (NIAAA). The total sample was reduced to 35,995 after removing participants for whom sexual orientation was unknown. Of this sample, 56.3% were female and 43.7% were male. Participants were 53% White, 21.3% Black, 4.9% Asian, Native Hawaiian, or other Pacific Islander, and 1.4% American Indian/Alaska Native. 19.4% of participants were Hispanic (any race). Participant ages ranged from 18 to 90 years old, with a mean age of 45.6 and a standard deviation (SD) of 17.5. With regard to self-identified sexual orientation, the majority (34,644, or 96.3%) reported being "heterosexual (straight)", 586 (1.6%) identified as "gay or lesbian", 566 (1.6%) identified as "bisexual", and 199 (0.6 %) selected "not sure". Of all male participants, 15,190 (96.6%) identified as "heterosexual (straight)", 321 (2%) identified as "gay or lesbian", 144 (0.9%) identified as "bisexual", and 69 (0.4%) selected "not sure". Among female participants, 19,454 (96%) identified as "heterosexual (straight)", 265 (1.3%) identified as "gay or lesbian", 422 (2.1%) identified as "bisexual", and 130 (0.6%) selected "not sure". Although not included in the present study proposal, an additional 314 (0.9 %) selected "unknown". Participants who selected gay/lesbian, bisexual, and not sure were classified as non-heterosexual, totaling 1351 (3.8%). A breakdown of sexual orientation frequencies is displayed in Table 2.

The NESARC-III sample was randomly selected from 50 states and the District of Columbia between April 2012 and June 2013 using multistage probability sampling. Primary sampling units consisted of individual counties or groups of contiguous counties, secondary sampling units consisted of census-defined blocks, third-stage sampling units involved selection of households within secondary sampling units, and final-stage sampling units consisted of randomly selected eligible adults within identified households. The screener had a household response rate of 72% and a personal-level response rate of 84% (total response rate = 60.1%). Participants provided full consent for their participation and were compensated $90 for their participation. Additional sample characteristics and recruitment information can be found in Grant et al. (2015).

Procedures

Face-to-face interviews were conducted by trained lay interviewers from the U.S. Bureau of the Census. Interviewers completed a 4 1/2-day training program and had relevant experience. Interviews lasted one hour on average. The NESARC research protocol was fully reviewed and approved by the U.S. Census Bureau and Office of Management and Budget (Grant et al., 2014).

Materials

Sexual Orientation Identity

Sexual orientation identity was measured using a single item asking participants to select the category that best describes their sexual orientation (“heterosexual/straight”, “gay or lesbian”, “bisexual”, “not sure”, or “unknown”). Natal sex was measured using a single item asking participants to report their sex (“male”, “female”).

Borderline Personality Disorder

Diagnostic criteria for BPD were measured using the Alcohol Use Disorder and Associated Disabilities Interview Schedule-5 (AUDADIS-5; Grant et al., 2015). The AUDADIS-5 is a structured, computer-assisted interview-style diagnostic instrument on which respondents are evaluated on 9 DSM-5 BPD criteria measured by 30 items on a lifetime basis. Criterion domains included unstable relationships (3 items), affective instability (3 items), efforts to avoid abandonment (4 items), difficulties controlling anger (3 items), identity disturbance (4 items), emptiness (2 items), suicidality (2 items), impulsivity (6 items), and stress-related paranoid ideation or dissociative symptoms (3 items; see Table 1). Following each item, distress or social-occupational dysfunction due to that criterion item was assessed using the question, "Did this ever trouble you or cause problems at work, school, or with your family or other people?". In order for a criterion to be considered present, participants must have endorsed at least one item within that criterion domain and must have also endorsed distress or social-occupational dysfunction related to that criterion item.[1] The impulsivity criterion required endorsement in two areas and distress/impairment for one of the two areas. In accordance with the DSM-5's approach to providing diagnoses based on criterion counts, a BPD diagnosis was coded for the participant if five or more criteria were endorsed. In primary analyses, total scores for each BPD criterion (number of items endorsed) with dysfunction/distress required were used. In sensitivity analyses, total scores (number of items endorsed) without the dysfunction/distress requirement were used.

Test-retest reliability for BPD diagnosis fell into the "good" range (.4-.6; Grant et al., 2015), with a kappa statistic of .54 (SE = .04). In the current sample, internal consistency for the

[1] In sensitivity analyses testing for DIF under less stringent requirements, dysfunction/distress was not required.

BPD scale was good across the 9 criteria (α = .86) and excellent across the 30 items (α = .93). Across criterion domains, internal consistency varied from poor to excellent. It was questionable for unstable relationships (α = .55), good for affective instability (α = .82) and efforts to avoid abandonment (α = .83), acceptable for anger (α = .79) and unstable sense of self (α = .75), good for empty (α = .81), questionable for paranoid ideation (α = .66), poor for suicidality (α = .55), and acceptable for impulsivity (α = .78).

Perceived Sexual Orientation-Related Minority Stressors

Perceived experiences of minority stress related to sexual orientation were measured using six items derived from the Experiences of Discrimination scales (Krieger & Sidney, 1997). Respondents who reported a non-heterosexual orientation were asked to rate the frequency with which they had experienced discrimination because they were assumed to be gay, lesbian, or bisexual in six different domains: (1) ability to obtain healthcare or health insurance coverage; (2) healthcare treatment; (3) in public settings (e.g., on the street, in stores, or in restaurants); (4) other situations (e.g., obtaining a job or on the job, getting admitted to a school or training program, in the courts, or by the police); (5) verbal harassment (e.g., called names); and (6) other victimization (e.g., made fun of, picked on, pushed, shoved, hit, or threatened with harm) within the past 12 months. Items were rated on a scale ranging from "0 – never" to "4 – often". A sum score across all six items was calculated, ranging from 0 to 24, with higher scores indicating greater frequency of discrimination across all six domains. The Experiences of Discrimination scales have demonstrated adequate validity and reliability in previous studies, with test-retest reliability coefficients averaging .70 (Krieger et al., 2005). In the current sample, internal consistency was good (α = .88).

Data Analyses

All analyses were conducted in R (*R Core Team*, 2020) using the lavaan (Rosseel, 2012), semTools (Jorgensen et al., 2021), and psych (Revelle, 2020) packages. Participants who responded "unknown" to items measuring sexual orientation were excluded. Missing data were handled using full information maximum likelihood (FIML). To test measurement invariance of DSM-5 BPD criteria across sexual orientation groups, a series of MIMIC structural equation models were fit. A MIMIC approach was used because of its flexibility in modeling multiple grouping variables and ability to account for measurement error and accommodate smaller sample sizes compared to other approaches (Fleishman et al., 2002; Woods, 2009). MIMIC models sequentially test whether two or more groups differ in their likelihood of endorsing each item within a scale when scores on the latent construct, according to carefully chosen anchor items, are held constant (Woods et al., 2009). In a MIMIC model, a latent factor (e.g., BPD) is regressed onto a grouping variable (e.g., sexual orientation) and is then held constant as each item is sequentially regressed onto the grouping variable. If the grouping variable predicts item responses after the latent factor score is accounted for, the item is considered to demonstrate DIF.

DIF can be uniform, nonuniform, or both. Uniform DIF indicates that while the item is functioning differently across groups, the differential functioning occurs in the same way, or uniformly, at every level of the trait score (θ). In a MIMIC model, uniform DIF is indicated by differences in item difficulty, or the intercept or threshold for a given item *i*, represented by the regression coefficient (b_i). A negative regression coefficient (*b*) indicates that an item's threshold—i.e., the value of θ, or the BPD score, at which the probability of endorsing the item is 0.5—is lower for the focal group, while a positive *b* value indicates that the threshold is higher

for the focal group. For example, in the current study, a negative *b* value for a given item would indicate that non-heterosexual individuals were more likely to endorse an item regardless of their overall BPD score compared to their heterosexual counterparts.

Conversely, nonuniform DIF indicates that item discrimination, or the item's ability to discern between those scoring high and low on the latent trait, also varies depending on the level of the trait score (θ). Nonuniform DIF is indicated by differences in the slope (i.e., factor loading) for a given item *i*, represented by the statistic ω_i. In this case, the ω_i statistic is equivalent to a beta weight. In the current study, nonuniform DIF would indicate trait differentials for the two groups, such that non-heterosexual individuals were more (or less) likely to endorse the item at low levels of BPD severity, but not at high levels, or vice versa. It is also possible for an item to demonstrate both uniform and nonuniform DIF, as uniformity may exist for only a subset of scores (e.g., uniformity in the difference in item functioning at low levels of a trait but variation at higher levels). For this reason, items were first tested for any type of DIF generally, regardless of uniformity. If any DIF was detected for an item, then it was next tested for nonuniform DIF. If results did not indicate the presence of nonuniform DIF, the item was then tested for uniform DIF. However, if nonuniform DIF was present, then uniform DIF was not tested for because the presence of nonuniform DIF precludes valid interpretation of uniform DIF (Woods et al., 2009).

For models with items indicating DIF, additional analyses were run to determine whether perceived minority stress could account for any DIF observed across sexual orientation in the model. A perceived minority stress variable was added to the model with paths to the latent BPD factor and each criterion item. A latent trait-by-perceived minority stress interaction term (BPD x perceived minority stress) interaction term was also defined to account for moderation effects of

perceived minority stressors on item discrimination. If statistical significance for any DIF items identified in the previous step disappears after the path for perceived minority stress is added, then perceived minority stress may partially account for DIF across sexual orientation groups.

Model Fit Indices

Model fit indices suggested by (Hu & Bentler, 1999) were used to evaluate fit of the first model and to compare change in fit for subsequent models. This includes a chi-square test of exact fit, root mean square error of approximation (RMSEA), and comparative fit index (CFI). The Tucker-Lewis index (TLI) was also calculated as an indicator of model fit. Models are considered to have good approximate fit if RMSEA values are < .06 and CFI values are > .95. A non-significant chi-square value was not required, although it provides additional support for adequate model fit.

Tests of Significance

Likelihood ratio test statistics and changes in RMSEA and CFI were evaluated as indicators of change in model fit. Model fit indices were compared across models using Satorra-Bentler likelihood ratio tests (LRT; Satorra & Bentler, 2001). Difference in model fit were considered significant if the LRT p-value was less than .01, as well as if ΔRMSEA values were > .015 and ΔCFI values were > .01, as recommended by Chen (2007). If model fit was significantly different for a given item compared to the previous model fit statistics, then the item was flagged for DIF.

Tests of Impact

Effect sizes to quantify the magnitude of DIF were also calculated using the expected score standard difference (ESSD) statistic (Meade, 2010). The ESSD is a standardized index that estimates the difference between the focal and reference groups due to DIF and can be

interpreted using Gignac & Szodorai's effect size guidelines (2016), with .2, .4, and .6 reflecting small, medium, and large effect sizes, respectively. Effect sizes were also calculated at the overall BPD scale-level using the expected test score standard difference (ETSSD) statistic. The effect of DIF is only considered practically meaningful at the item-level if the ESSD value is ≥ .2 and at the test level if the ETSSD is ≥ .2.

Hypotheses 1a-1d: MIMIC models were fit to assess for uniform and nonuniform DIF among the 9 DSM-5 BPD diagnostic criterion items across lesbian, gay, bisexual, and heterosexual individuals using a two-step sequential-free baseline approach (see Figure 1; Chun et al., 2016). Five MIMIC models were fit to evaluate DIF across (1) heterosexual individuals and non-heterosexual individuals, (2) heterosexual and gay men, (3) heterosexual and bisexual men, (4) heterosexual and lesbian women, and (5) heterosexual and bisexual women. In each case, the heterosexual group served as the referent group and the non-heterosexual group served as the focal group. First, a constrained baseline model was specified in which sexual orientation was added as a grouping variable and all items were fixed to be equal across groups. Next, a latent trait-by-group interaction term (BPD x sexual orientation) was added. In this case, the latent BPD score represents the theta (θ).

In the first stage, a series of sequential-free baseline models were fit to identify items to be used as anchors in subsequent analyses. Due to non-convergence of models containing all 30 BPD items,[2] items were collapsed across their respective criterion domains, which were used in place of the items. Following Woods' (2009) recommendation, 10% of items—in this case one item—was used as the anchor. First, all item parameters were constrained to be equal across groups. Each item was then released individually, and the resulting model fit indices were

[2] Models failed to converge after the latent trait-by-group interaction term was added.

compared to the baseline model. Each item resulting in a significant LRT upon release was flagged for potential DIF. Of the items not flagged for DIF, the most discriminating item, or those with the highest factor loading, was chosen as an anchor. After the anchor item was chosen, the factor mean was fixed to 0 and the variance was fixed to 1 to identify the scale.

In the second stage, the item response for the anchor item was constrained to be equal across groups and all other items were allowed to vary freely. Then, each non-anchor item was iteratively tested for DIF using a 2*df* test in which both the factor loadings and thresholds were constrained to be equal across groups. After each item was constrained, model fit indices were compared against the fit of the baseline model. Each item that resulted in significantly worse fit ($p < .01$) according to LRTs after being constrained was flagged as demonstrating DIF. For each item flagged for DIF (whether due to uniform or nonuniform DIF), nonuniform DIF was tested using a 1*df* test in which the factor loading was constrained to be equal. If results suggested the presence of nonuniform DIF for a given item (as evidenced by a statistically significant ω statistic, representing the path between the item and trait-by-group interaction term ($p < .01$), then that item was not tested for uniform DIF. If nonuniform DIF was not detected, then a separate 1*df* test was conducted, in which the item threshold was constrained to be equal. If the resulting *b* value representing the regression path from the grouping variable to the item was significant ($p < .01$), then the item was deemed to exhibit uniform DIF.

Magnitude of DIF among flagged items was evaluated by calculating the ESSD and ETSSD statistics. If ESSD values for these items are ≥ .2, then the practical significance of DIF is considered meaningful. Additionally, an ETSSD statistic for the overall BPD scale ≥ .2 will indicate that DIF among criteria exerts a non-negligible impact on performance of the test

overall. If these values are not above their cutoff, then results are taken to suggest that differences in latent BPD are reflective of true elevations in BPD features.

Hypotheses 2a-2b: Next, tests were conducted to evaluate DIF of BPD criteria across sexual orientation while accounting for perceived minority stress (see Figure 2). First, a MIMIC model was specified using the same anchor item as used for Hypotheses 1a-1d. The latent BPD factor was regressed onto a perceived minority stress variable and an interaction term was specified between the latent BPD factor and perceived minority stress. Then, each DIF-flagged item was regressed onto the minority stress variable and the new interaction term and subsequently tested for DIF while controlling for the effect of minority stress on item endorsement.

Sensitivity Analyses

A series of sensitivity analyses were conducted to evaluate the robustness of findings. An additional set of MIMIC models were run that did not require dysfunction or distress for any item to be counted toward each BPD criterion. Investigation of DIF under these less stringent requirements was done to gain insight into the extent to which DIF among criteria is unique to clinical levels of BPD symptom endorsement compared to mere presence of BPD traits. As described previously, BPD symptoms can be thought of as extreme variants or tail ends of normally distributed personality traits. In other words, they represent the maladaptive presentations of these traits that have surpassed a designated threshold representing clinically significant impairment. Given that BPD features are present at lower levels of intensity among the non-clinical population, symptoms are commonly measured without the dysfunction/distress requirement in clinical and research settings, such as in screening tools and self-report research instruments, which typically yield much higher symptom endorsement rates (Trull et al., 2010).

These sensitivity analyses may be considered a preliminary investigation into the likelihood of DIF among a larger subset of individuals who endorse less clinically elevated levels of BPD on dimensional measures. The process for evaluating DIF in sensitivity analyses was identical to the process described above for primary analyses.

Table 2

Sexual Orientation Frequencies in Current Sample

	Heterosexual *N* (%)	Total Non-heterosexual *N* (%)	Gay or Lesbian *N* (%)	Bisexual *N* (%)	Not Sure *N* (%)
Overall	34,644 (96.3%)	1,351 (3.8%)	586 (1.6%)	566 (1.6%)	199 (0.6 %)
Male	15,190 (96.6%)	534 (3.4%)	321 (2%)	144 (0.9%)	69 (0.4%)
Female	19,454 (96%)	817 (4.0%)	265 (1.3%)	422 (2.1%)	130 (0.6%)

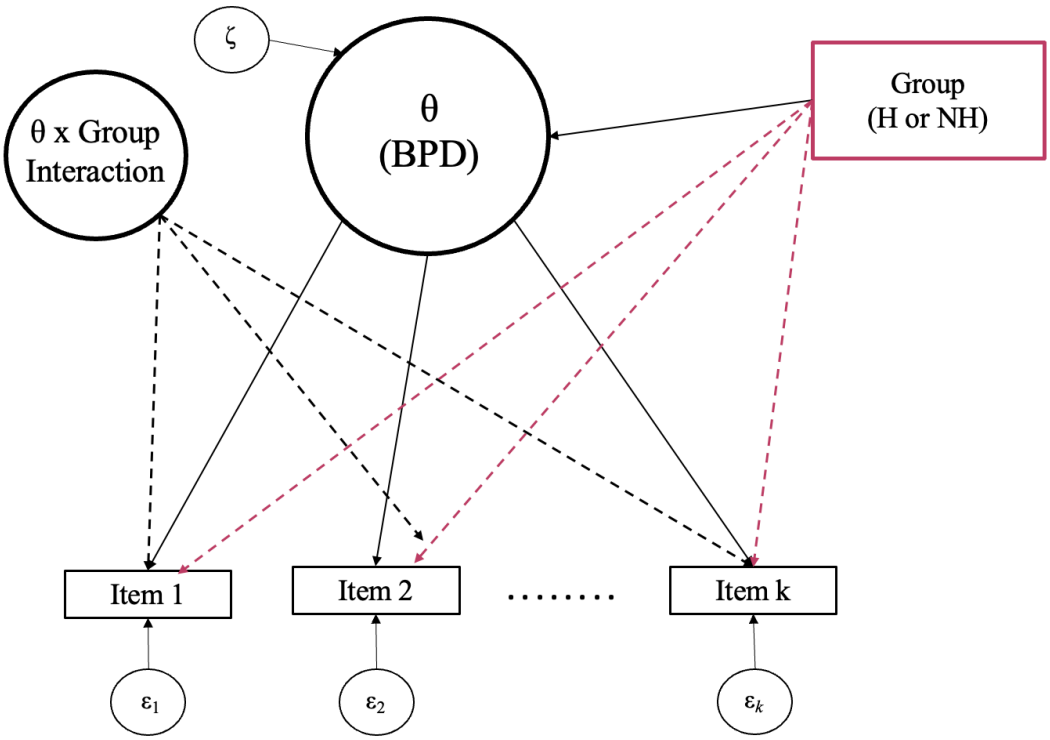

Figure 1

MIMIC Model 1

Note. H = heterosexual; NH = non-heterosexual; θ = latent BPD score; ζ = residual for latent BPD score; ε_i = item measurement error.

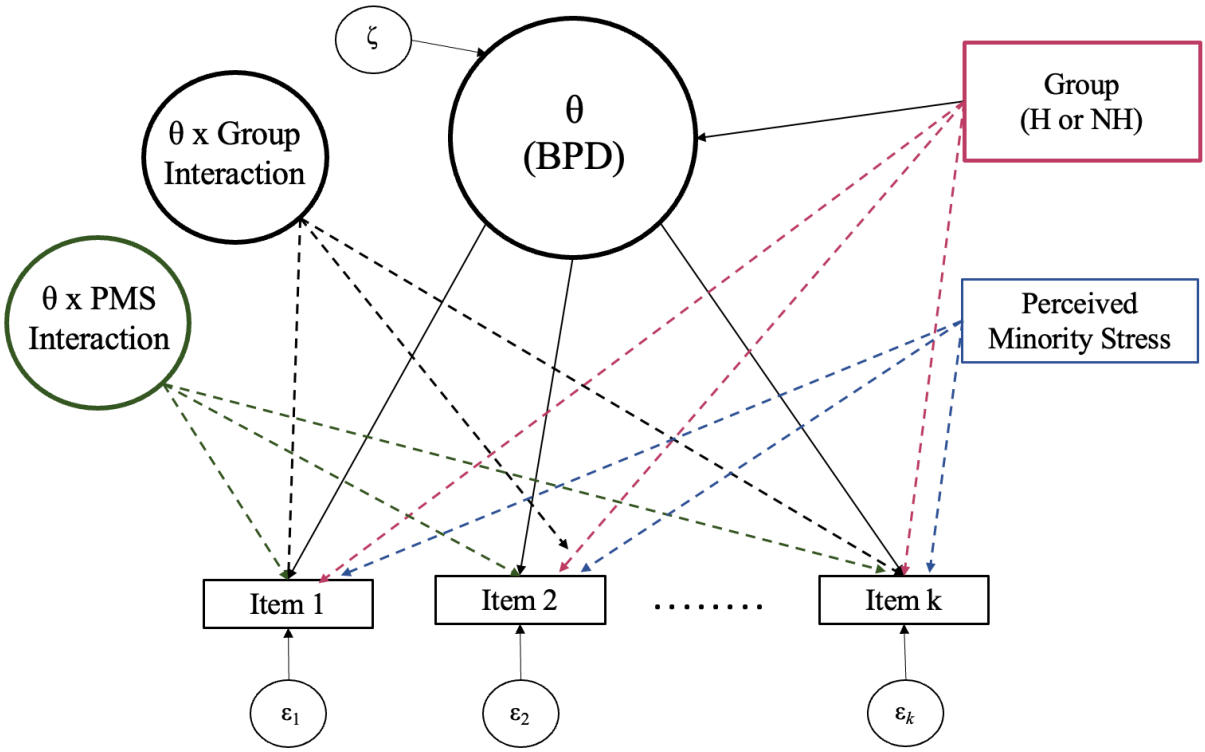

Figure 2

MIMIC Model 2

Note. H = heterosexual; NH = non-heterosexual; PMS = perceived minority stress; θ = latent BPD score; ζ = residual for latent BPD score; ε_i = item measurement error.

Chapter Three:

Results

Descriptive Results

In the current sample, 5.5% of total participants met diagnostic criteria for BPD, including 5.1% of heterosexual participants and 15.1% of sexual minority participants (9% of gay men, 12.5% of bisexual men, 11.3% of lesbian women, and 20.4% of bisexual women). By gender, BPD criteria were met by 4.9% of male and 6% of female participants. The average number of BPD symptoms endorsed (requiring distress/impairment) was 0.74 (*SD* = 1.67) of the 9 symptoms. The average number of BPD items endorsed with accompanied distress/impairment was 1.40 (*SD* = 3.61) of the 30 items. The average number of BPD items endorsed regardless of whether they were associated with distress/impairment was 3.49 (*SD* = 4.86) of the 30 items.

The average score for perceived minority stressors across all non-heterosexual participants was 7.60 (*SD* = 3.3). Among gay men, the average score was 8.28 (*SD* = 3.5). Among bisexual men, the average score was 7.47 (*SD* = 3.4). Lesbian women demonstrated an average score of 8.49 (*SD* = 3.7) and bisexual women had an average score of 6.93 (*SD* = 2.5).

Measurement Invariance Results

DIF Analysis for BPD Models

It was hypothesized that DIF would be indicated for four criteria including suicidality (1a), efforts to avoid abandonment (1b), impulsivity (1c), and identity disturbance (1d), with lower item thresholds for sexual minority men and women. It was additionally hypothesized that

after accounting for perceived minority stress, DIF would no longer be indicated for suicidality (2a) and efforts to avoid abandonment (2b).

Heterosexual Versus Non-Heterosexual Individuals. In the first set of MIMIC models comparing heterosexual men and women to non-heterosexual men and women, affective instability was used as the anchor item. Fit indices for the baseline model and stage 1 and 2 sequential models were generally acceptable but hovered around the cutoffs for good model fit, with RMSEA values between .058-.061 and CFI values ranging from .948-.952. ΔRMSEA did not exceed .015 and ΔCFI did not exceed .01 in any model comparison, although significant LRTs suggested the presence of DIF for several items. Model fit indices for stages 1 and 2 can be found in Tables 3.1 and 3.2, respectively. In the final model, nonuniform DIF was indicated for five criteria, including efforts to avoid abandonment ($\omega = .055$, $p < .001$), identity disturbance ($\omega = .056$, $p < .01$), paranoid ideation ($\omega = .049$, $p < .01$), suicidality ($\omega = .112$, $p < .001$), and impulsivity ($\omega = .044$, $p < .001$). For all criteria, ω statistics were positive, indicating that these five criteria are more discriminating for non-heterosexual individuals compared to heterosexual individuals. This pattern of results is contrary to the prediction of lower thresholds for non-heterosexual individuals. Thus, hypotheses 1a through 1d were not supported. ESSD indices were below the cutoff of ≥ .2 for practically meaningful DIF at the criterion level for all five criteria (efforts to avoid abandonment [.036], identity disturbance [.028], paranoid ideation [.022], suicidality [.099], impulsivity [.041]). The ETSSD was .02, below the cutoff of ≥ .2 for practically meaningful DIF at the overall test level.

All criteria lost statistical significance after accounting for perceived minority stressors (efforts to avoid abandonment [$\omega = .016$, $p = .864$], identity disturbance [$\omega = -.208$, $p = .743$], paranoid ideation [$\omega = -.200$, $p = .745$], suicidality [$\omega = .122$, $p = .109$], and impulsivity [$\omega =$

.153, p = .132]). As such, hypotheses 2a and 2b were supported. These results suggest that there are not practically meaningful differences in endorsement of BPD criteria across heterosexual and non-heterosexual individuals. The DIF displayed for efforts to avoid abandonment, identity disturbance, paranoid ideation, suicidality, and impulsivity, while not practically meaningful, can be accounted for by perceived minority stressors. Results from MIMIC models comparing heterosexual individuals to non-heterosexual individuals are displayed in Table 3, Table 4, and Table 5.

Heterosexual Versus Gay Men. Subsequent sets of models tested for DIF across genders and sexual minority subgroups. In the model comparing heterosexual men to gay men, affective instability was used as the anchor item. Fit indices for the baseline model and stage 1 and 2 sequential models suggested model fit was generally less than good, with RMSEA values between .055-.058 and CFI values ranging from .941-.942. ΔRMSEA did not exceed .015 and ΔCFI did not exceed .01 in any model comparison, although significant LRTs suggested the presence of DIF for some items. Model fit indices for stages 1 and 2 can be found in Tables 4.1 and 4.2, respectively. In the final model, nonuniform DIF was indicated for unstable relationships (ω = .052, p < .001) and identity disturbance (ω = .053, p < .01). The positive ω statistics for these criteria indicate greater discrimination for gay men. Though flagged for DIF in stage 2, 1df and 2df tests showed that anger, suicidality, and impulsivity were not significant for uniform or nonuniform DIF in the final model. Thus, hypotheses 1a through 1d were not supported. ESSD indices were below the cutoff for both unstable relationships (.079) and identity disturbance (.044), indicating that DIF demonstrated for these criteria was not practically meaningful. The ETSSD was -.001, below the cutoff for practically meaningful DIF at the overall test level.

Both criteria lost statistical significance after accounting for perceived minority stressors (unstable relationships [ω = .173, p = .687], identity disturbance [ω = .085, p = .757]). The DIF displayed for unstable relationships and identity disturbance, while not practically meaningful, can be accounted for by perceived minority stressors. As neither efforts to avoid abandonment nor suicidality indicated significant DIF prior to accounting for perceived minority stress, hypotheses 2a and 2b were not supported. These results suggest that there are not practically meaningful differences in endorsement of BPD criteria across heterosexual and gay men. Results from MIMIC models comparing heterosexual and gay men are displayed in Table 6, Table 7, and Table 8.

Heterosexual Versus Bisexual Men. In the model comparing heterosexual men to bisexual men, affective instability was used as the anchor item. Fit indices for the baseline model and stage 1 and 2 sequential models were below the cutoffs for good fit, with RMSEA values between .078-.082 and CFI values ranging from .894-.895. ΔRMSEA did not exceed .015 and ΔCFI did not exceed .01 in any model comparison, although LRTs were significant in stage 2 for several items. Model fit indices for stages 1 and 2 can be found in Tables 5.1 and 5.2, respectively. In the final model, no criteria evidenced DIF. Though flagged for DIF in stage 2, 1*df* and 2*df* tests showed that efforts to avoid abandonment and suicidality were not significant for uniform or nonuniform DIF in the final model. As such, hypotheses 1a through 1d were not supported, nor were hypotheses 2a or 2b. The ETSSD was -.001, below the cutoff for practically meaningful DIF at the overall test level. These results suggest that there are not practically meaningful differences in endorsement of BPD criteria across heterosexual and bisexual men. Results from MIMIC models comparing heterosexual and bisexual men are displayed in Table 9, Table 10, and Table 11.

Heterosexual Versus Lesbian Women. In the model comparing heterosexual women to lesbian women, affective instability was used as the anchor item. Fit indices for the baseline model and stage 1 and 2 sequential models suggested good model fit, with RMSEA values between .052-.053 and CFI values ranging from .956-.957. ΔRMSEA did not exceed .015 and ΔCFI did not exceed .01 in any model comparison, although LRTs were significant in stage 2 for several items. Model fit indices for stages 1 and 2 can be found in Tables 6.1 and 6.2, respectively. In the final model, no criteria evidenced DIF. Though flagged for DIF in stage 2, 1*df* and 2*df* tests showed that efforts to avoid abandonment, suicidality, and impulsivity were not significant for uniform or nonuniform DIF in the final model. As such, hypotheses 1a through 1d were not supported, nor were hypotheses 2a or 2b. The ETSSD was .014, below the cutoff for practically meaningful DIF at the overall test level. These results suggest that there are not practically meaningful differences in endorsement of BPD criteria across heterosexual and lesbian women. Results from MIMIC models comparing heterosexual and lesbian women are displayed in Table 12, Table 13, and Table 14.

Heterosexual Versus Bisexual Women. In the model comparing heterosexual women to bisexual women, affective instability was used as the anchor item. Fit indices for the baseline model and stage 1 and 2 sequential models fell below the cutoffs for good fit, with RMSEA values between .061-.065 and CFI values ranging from .948-.950. ΔRMSEA did not exceed .015 and ΔCFI did not exceed .01 in any model comparison, although LRTs were significant in stages 1 and 2 for several items. Model fit indices for stages 1 and 2 can be found in Tables 7.1 and 7.2, respectively. In the final model, three items evidenced nonuniform DIF, including anger (ω = .054, $p < .001$), suicidality (ω = .085, $p < .001$), and impulsivity (ω = .075, $p < .001$). Though flagged for DIF in stage 2, 1*df* tests showed that identity disturbance, emptiness, and paranoid

ideation were not significant for uniform or nonuniform DIF in the final model. The positive ω statistics for all three criteria indicate greater discrimination for bisexual women. Thus, hypotheses 1a through 1d were not supported. ESSD values were below the cutoff for anger (.058), suicidality (.106), and impulsivity (.101), suggesting that DIF was not practically meaningful for these criteria. The ETSSD was .028, below the cutoff for practically meaningful DIF at the overall test level.

All three criteria lost statistical significance after perceived minority stressors was accounted for (anger [$\omega = .001$, $p = .995$], suicidality [$\omega = .261$, $p = .292$], and impulsivity [$\omega = -.126$, $p = .466$]). As such, hypothesis 2b was supported. These results suggest that there are not practically meaningful differences in endorsement of BPD criteria across heterosexual and bisexual women. The DIF displayed for anger, suicidality, and impulsivity, while not practically meaningful, can be accounted for by perceived minority stressors. Results from MIMIC models comparing heterosexual individuals to non-heterosexual individuals are displayed in Table 15, Table 16, and Table 17.

Across all models, affective instability was chosen as the anchor item as it consistently yielded the highest factor loading among items not flagged for DIF. Overall, hypotheses pertaining to Research Question 1 (1a-1d), predicting lower thresholds for non-heterosexual individuals for efforts to avoid abandonment, identity disturbance, suicidality, and impulsivity, were not supported. While many criteria evidenced significant DIF, it was nonuniform, indicating greater discrimination for focal groups. Across all models, ESSD values for all criteria were below the cutoff of $\geq .2$, suggesting DIF was not practically meaningful for any criterion. Hypotheses pertaining to Research Question 2 (2a-2b) received considerable support across models, with perceived minority stressors explaining DIF for both efforts to avoid abandonment

and suicidality in all cases in which they displayed DIF. Although not hypothesized, several additional items displayed DIF in at least one model, including unstable relationships, anger, and paranoid ideation. In all cases, perceived minority stress appeared to account for the observed patterns of DIF. In sum, a negligible degree of DIF was demonstrated for efforts to avoid abandonment, identity disturbance, suicidality, and impulsivity, as well as unstable relationships, anger, and paranoid ideation, which was accounted for by perceived minority stress.

DIF Analyses for Models Not Requiring Dysfunction

Sensitivity analyses were conducted to assess the extent of DIF without the requirement of accompanying dysfunction or distress for each symptom. Affective instability was chosen as the anchor item in nearly all cases as it demonstrated the highest factor loading among all non-DIF items, with the exception of the first set of models comparing heterosexual and non-heterosexual individuals, in which case identity disturbance was used.

Across sensitivity analyses, results pertaining to hypotheses generally followed the same pattern as results for primary analyses, with some exceptions. In models comparing heterosexual and non-heterosexual men and women DIF was not indicated for identity disturbance or impulsivity. In models comparing heterosexual and gay men, uniform DIF was indicated for suicidality ($b = .017$, $p < .01$), with a positive b value indicating higher thresholds for gay men, while identity disturbance did not demonstrate DIF. After accounting for perceived minority stressors, DIF for suicidality was no longer significant ($b = .009$, $p = .139$) and hypothesis 2b was therefore supported. In models comparing heterosexual women to lesbian women, uniform DIF was indicated for suicidality ($b = .016$, $p < .01$) and impulsivity ($b = .020$, $p < .001$) with positive b value indicating higher thresholds for lesbian women. Hypotheses 1a through 1d were therefore not supported. After accounting for perceived minority stressors, DIF was no longer

significant for suicidality (b = .009, p = .226). Thus, hypothesis 2b was supported. In comparisons of heterosexual and bisexual women, nonuniform DIF was evidenced for efforts to avoid abandonment (ω = .033, p < .01), suicidality (ω = .091, p < .001), and impulsivity (ω = .047, p < .001), with positive ω values indicating greater discrimination for bisexual women. Thus, hypotheses 1a through 1d were not supported. After accounting for perceived minority stressors, DIF was no longer significant for efforts to avoid abandonment (ω = .391, p = .285), suicidality (ω = -.017, p = .886), and impulsivity (ω = .060, p = .422). As such, hypotheses 2a and 2b were supported. Some differences were additionally present for other non-hypothesized items in some models, as shown in Tables 19 through 33 in Appendix A displaying results from sensitivity MIMIC models.

Throughout sensitivity analyses, hypotheses pertaining to Research Question 1 (1a-1d), predicting lower thresholds for non-heterosexual individuals for efforts to avoid abandonment, identity disturbance, suicidality, and impulsivity, did not receive support. While many of these criteria demonstrated DIF, it was largely nonuniform, with greater discrimination for focal groups. Across all models, ESSD values for all criteria were below the cutoff of ≥ .2, suggesting DIF was not practically meaningful for any criterion. Hypotheses pertaining to Research Question 2 (2a-2b) received some support across models, with perceived minority stressors explaining DIF for efforts to avoid abandonment in two out of five models and suicidality in four out of five models. DIF was also indicated for several additional items in at least one model, including affective instability, emptiness, paranoid ideation, and anger. In all cases, DIF was fully accounted for by perceived minority stress. In sum, a negligible degree of DIF was demonstrated for efforts to avoid abandonment and suicidality, as well as affective instability,

emptiness, paranoid ideation, and anger in sensitivity models, which was accounted for by perceived minority stress.

Overall, across primary analyses and sensitivity analyses, DIF was evidenced most frequently in broad models comparing all heterosexual and non-heterosexual men and women, in which DIF was apparent for five criteria, and for bisexual women, in which DIF was indicated for three criteria. All nine criteria demonstrated DIF in at least one model. Evidence of DIF was most robust for suicidality (6 out of 10 models), impulsivity (4 out of 10 models), and efforts to avoid abandonment (3 out of 10 models), although the nature of DIF was not as predicted in hypotheses 1a through 1d. Instead of revealing lower thresholds for non-heterosexual individuals, DIF was almost entirely nonuniform, with greater discrimination for non-heterosexual individuals. DIF was overwhelmingly explained by perceived minority stressors, thus providing general support for hypotheses 2a and 2b. ESSD values were all below the cutoff of $\geq .2$, suggesting DIF was not practically meaningful for any criterion. Model fit was acceptable for most models but hovered around cutoffs for good fit in many cases, and consistently fell outside the range of good fit for bisexual men. Results of all hypothesis testing can be found in Table 18.

Table 3

Fit Indices for Stage 1 Sequential MIMIC Models Comparing Heterosexual and Non-Heterosexual Individuals

Model	BL Std. Ld.	RMSEA	CFI	TLI	AIC	X^2	*df*	ΔRMSEA	ΔCFI	ΔTLI	ΔAIC	ΔX^2	Δ*df*
		Baseline Model											
Baseline		.058	.948	.938	-111114.1	1206.89	142	–	–	–	–	–	–
		Stage 1 Sequential Models											
Unstable Relationships	.656	.059	.948	.937	-111112.4	1192.36	140	.001	.000	.001	-1.78	0.88	2
Affective Instability	.749	.059	.948	.937	-111122.2	1191.60	140	.001	.000	.001	8.12	5.09	2
Efforts to Avoid Abandonment	.711	.059	.948	.937	-111169.5	1190.25	140	.001	.000	.001	55.35	15.48**	2
Anger	.693	.059	.948	.937	-111114.7	1192.99	140	.001	.000	.001	0.58	1.39	2
Identity Disturbance	.721	.059	.948	.937	-111172.8	1192.01	140	.001	.000	.001	58.67	10.92*	2
Emptiness	.727	.059	.948	.937	-111118.7	1192.52	140	.001	.000	.001	4.58	2.80	2
Paranoid Ideation	.727	.059	.948	.937	-111145.4	1193.52	140	.001	.000	.001	31.30	6.30	2
Suicidality	.481	.058	.949	.938	-111485.6	1174.24	140	.000	.001	.000	371.49	54.64**	2
Impulsivity	.652	.059	.948	.937	-111127.5	1191.73	140	.001	.000	.001	13.34	6.17	2

Note. Model fit indices for baseline MIMIC model and stage 1 sequential models. In baseline model, all item parameters were constrained to be equal across groups. In stage 1 sequential models, each item was released individually and model fit indices were compared to the baseline model. Each item resulting in a significant Likelihood Ratio Test upon release was flagged for potential DIF.
BL = Baseline; RMSEA = Root Square Mean Error of Approximation; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; AIC = Akaike Information Criterion; *df* = degrees of freedom.
* $p < .01$
** $p < .001$

Table 4

Fit Indices for Stage 2 Sequential MIMIC Models Comparing Heterosexual and Non-Heterosexual Individuals

Model	RMSEA	CFI	TLI	AIC	X^2	*df*	ΔRMSEA	ΔCFI	ΔTLI	ΔAIC	ΔX^2	Δ*df*
Baseline Model												
Baseline	.060	.952	.934	-112250.3	1206.89	126	–	–	–	–	–	–
Stage 2 Sequential Models												
Unstable Relationships	.060	.951	.935	-111112.4	340.95	128	.000	.001	.001	-1.78	2.15	2
Affective Instability	*Anchor item*											
Efforts to Avoid Abandonment	.060	.950	.934	-111169.5	342.47	128	.000	.002	.000	55.35	57.16**	2
Anger	.060	.951	.934	-111114.7	1202.44	128	.000	.001	.000	0.58	4.46	2
Identity Disturbance	.060	.950	.934	-111172.8	1146.47	128	.000	.002	.000	58.67	60.43**	2
Emptiness	.060	.951	.934	-111118.7	1198.30	128	.000	.001	.000	4.58	8.60	2
Paranoid Ideation	.060	.950	.934	-111145.4	1172.75	128	.000	.002	.000	31.30	34.14**	2
Suicidality	.061	.950	.933	-111485.6	847.97	128	.001	.002	.001	371.49	358.92**	2
Impulsivity	.060	.951	.934	-111127.5	1190.28	128	.000	.001	.000	13.34	16.61**	2

Note. Model fit indices for baseline MIMIC model and stage 2 sequential models. Anchor item response was constrained to be equal across groups and all other items varied freely. Then, each non-anchor item was iteratively tested for DIF using a 2*df* test in which both the slope and threshold were constrained to be equal across groups. Model fit indices were then compared against the fit of the baseline model. Each item that resulted in significantly worse fit ($p < .01$) according to Likelihood Ratio Tests after being constrained was flagged for DIF.

RMSEA = Root Square Mean Error of Approximation; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; AIC = Akaike Information Criterion; *df* = degrees of freedom.

* $p < .01$

** $p < .001$

Table 5

DIF and ESSD Results for MIMIC Models Comparing Heterosexual and Non-Heterosexual Individuals

Criterion	Factor Std. Ld.	Uniform *b*	Non-Uniform ω	ESSD	Uniform *b*	Non-Uniform ω	PMS *b*	PMS* BPD *b*	ESSD
	Final Model				*Perceived Minority Stress Model*				
Unstable Relationships	.656	–	–	–	–	–	–	–	–
Affective Instability	.749	–	–	–	–	–	–	–	–
Efforts to Avoid Abandonment	–	–	.055**	.036	–	.016	-.009*	.043	.029
Anger	.693	–	–	–	–	–	–	–	–
Identity Disturbance	–	–	.056**	.028	. –	-.208	-.007	.268	.047
Emptiness	.727	–	–	–	–	–	–	–	–
Paranoid Ideation	–	–	.049*	.022	–	-.200	-.007	.254	.050
Suicidality	–	–	.112**	.099	–	.122	-.001	-.011	.043
Impulsivity	–	–	.042**	.041	–	.153	-.008	-.115	.049

Note. For each item flagged for DIF, nonuniform DIF was tested using a 1*df* test in which the slope was constrained to be equal. If results suggested the presence of nonuniform DIF for a given item, then that item was not tested for uniform DIF. If nonuniform DIF was not detected, then a separate 1*df* test was conducted, in which the threshold was constrained to be equal. If the resulting *b* value was significant ($p <$ .01), then the item was deemed to exhibit uniform DIF. Magnitude of DIF among flagged items is evaluated by the ESSD statistic, with .2, .4, and .6 reflecting small, medium, and large effect sizes, respectively. The effect of DIF is only considered practically meaningful at the item-level if the ESSD value is ≥ .2.

b = standardized regression coefficient from item to latent factor; ω = standardized regression coefficient from item to latent trait-by-group interaction term; ESSD = Expected score standardized difference; PMS = Perceived minority stress.

* $p < .01$

** $p < .001$

Table 6

Fit Indices for Stage 1 Sequential MIMIC Models Comparing Heterosexual and Gay Men

Model	BL Std. Ld.	RMSEA	CFI	TLI	AIC	X^2	*df*	ΔRMSEA	ΔCFI	ΔTLI	ΔAIC	ΔX^2	Δ*df*
		Baseline Model											
Baseline		.055	.941	.929	-189040.1	346.96	142	–	–	–	–	–	–
		Stage 1 Sequential Models											
Unstable Relationships	.638	.055	.941	.928	-189084.4	340.95	140	.000	.000	.001	44.38	20.83**	2
Affective Instability	.730	.056	.941	.928	-189037.3	342.47	140	.001	.000	.001	-2.77	0.42	2
Efforts to Avoid Abandonment	.673	.056	.941	.928	-189045.0	342.44	140	.001	.000	.001	4.99	2.06	2
Anger	.663	.056	.941	.928	-189045.8	342.23	140	.001	.000	.001	5.73	3.19	2
Identity Disturbance	.704	.055	.941	.928	-189091.5	341.27	140	.000	.000	.001	51.45	9.38*	2
Emptiness	.703	.056	.941	.928	-189042.1	342.42	140	.001	.000	.001	2.10	1.69	2
Paranoid Ideation	.707	.056	.941	.928	-189039.6	342.88	140	.001	.000	.001	-0.50	0.57	2
Suicidality	.467	.056	.941	.928	-189048.5	342.95	140	.001	.000	.001	8.44	1.46	2
Impulsivity	.636	.055	.941	.928	-189058.3	341.77	140	.000	.000	.001	44.38	8.89	2

Note. Model fit indices for baseline MIMIC model and stage 1 sequential models. In baseline model, all item parameters were constrained to be equal across groups. In stage 1 sequential models, each item was released individually and model fit indices were compared to the baseline model. Each item resulting in a significant Likelihood Ratio Test upon release was flagged for potential DIF.
BL = Baseline; RMSEA = Root Square Mean Error of Approximation; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; AIC = Akaike Information Criterion; *df* = degrees of freedom.
* $p < .01$
** $p < .001$

Table 7

Fit Indices for Stage 2 Sequential MIMIC Models Comparing Heterosexual and Gay Men

Model	RMSEA	CFI	TLI	AIC	X^2	*df*	ΔRMSEA	ΔCFI	ΔTLI	ΔAIC	ΔX^2	Δ*df*
Baseline Model												
Baseline	.058	.942	.921	-189220.0	346.96	126	–	–	–	–	–	–
Stage 2 Sequential Models												
Unstable Relationships	.058	.941	.921	-189084.4	299.31	128	.000	.001	.000	44.38	47.65**	2
Affective Instability	*Anchor item*											
Efforts to Avoid Abandonment	.058	.942	.922	-189037.3	338.05	128	.000	.000	.001	4.99	8.91	2
Anger	.058	.942	.922	-189045.0	337.30	128	.000	.000	.001	5.73	9.66*	2
Identity Disturbance	.058	.941	.921	-189045.8	292.09	128	.000	.001	.000	51.45	54.88**	2
Emptiness	.058	.942	.922	-189091.5	340.93	128	.000	.000	.001	2.10	6.03	2
Paranoid Ideation	.057	.942	.922	-189042.1	343.49	128	.001	.000	.001	-0.50	3.47	2
Suicidality	.058	.942	.922	-189039.6	334.64	128	.000	.000	.001	8.44	12.32*	2
Impulsivity	.057	.942	.922	-189048.5	324.85	128	.001	.000	.001	44.38	22.11**	2

Note. Model fit indices for baseline MIMIC model and stage 2 sequential models. Anchor item response was constrained to be equal across groups and all other items varied freely. Then, each non-anchor item was iteratively tested for DIF using a 2*df* test in which both the slope and threshold were constrained to be equal across groups. Model fit indices were then compared against the fit of the baseline model. Each item that resulted in significantly worse fit ($p < .01$) according to Likelihood Ratio Tests after being constrained was flagged for DIF.
RMSEA = Root Square Mean Error of Approximation; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; AIC = Akaike Information Criterion; *df* = degrees of freedom.
* $p < .01$
** $p < .001$

Table 8

DIF and ESSD Results for MIMIC Models Comparing Heterosexual and Gay Men

Criterion	Factor Std. Ld.	Uniform *b*	Non-Uniform ω	ESSD	Uniform *b*	Non-Uniform ω	PMS *b*	PMS*BPD *b*	ESSD
	Final Model				*Perceived Minority Stress Model*				
Unstable Relationships	–	–	.052**	.079	–	.173	.008	-.126	-.130
Affective Instability	.730	–	–	–	–	–	–	–	–
Efforts to Avoid Abandonment	.673	–	–	–	–	–	–	–	–
Anger	–	-.010	-.009	-.059	-.013	-.042	.003	.034	-.014
Identity Disturbance	–	–	.053*	.044	–	.085	-.001	-.039	-.015
Emptiness	.703	–	–	–	–	–	–	–	–
Paranoid Ideation	.707	–	–	–	–	–	–	–	–
Suicidality	–	.010	.032	.037	.013	.131	-.008	-.098	-.058
Impulsivity	–	-.012	-.020	-.073	.018	.051	-.025	-.073	.108

Note. For each item flagged for DIF, nonuniform DIF was tested using a 1*df* test in which the slope was constrained to be equal. If results suggested the presence of nonuniform DIF for a given item, then that item was not tested for uniform DIF. If nonuniform DIF was not detected, then a separate 1*df* test was conducted, in which the threshold was constrained to be equal. If the resulting *b* value was significant ($p < .01$), then the item was deemed to exhibit uniform DIF. Magnitude of DIF among flagged items is evaluated by the ESSD statistic, with .2, .4, and .6 reflecting small, medium, and large effect sizes, respectively. The effect of DIF is only considered practically meaningful at the item-level if the ESSD value is ≥ .2.

b = standardized regression coefficient from item to latent factor; ω = standardized regression coefficient from item to latent trait-by-group interaction term; ESSD = Expected score standardized difference; PMS = Perceived minority stress.

* $p < .01$

** $p < .001$

Table 9

Fit Indices for Stage 1 Sequential MIMIC Models Comparing Heterosexual and Bisexual Men

Model	BL Std. Ld.	RMSEA	CFI	TLI	AIC	X^2	*df*	ΔRMSEA	ΔCFI	ΔTLI	ΔAIC	ΔX^2	Δ*df*
		Baseline Model											
Baseline		.078	.894	.872	-263667.7	375.32	142	–	–	–	–	–	–
		Stage 1 Sequential Models											
Unstable Relationships	.636	.078	.894	.871	-263664.7	370.47	140	.000	.000	.001	-3.04	0.20	2
Affective Instability	.731	.078	.894	.871	-263667.3	370.39	140	.000	.000	.001	-0.34	0.80	2
Efforts to Avoid Abandonment	.672	.078	.894	.871	-263676.5	370.43	140	.000	.000	.001	8.76	1.90	2
Anger	.664	.078	.894	.871	-263665.0	370.31	140	.000	.000	.001	-2.73	0.39	2
Identity Disturbance	.703	.078	.894	.871	-263665.2	370.54	140	.000	.000	.001	-2.49	0.27	2
Emptiness	.703	.078	.894	.871	-263671.9	370.19	140	.000	.000	.001	4.22	2.51	2
Paranoid Ideation	.709	.078	.894	.871	-263664.2	371.05	140	.000	.000	.001	-3.46	0.05	2
Suicidality	.473	.078	.895	.871	-263732.9	369.83	140	0.000	.001	.001	65.19	6.32	2
Impulsivity	.637	.078	.894	.871	-263666.9	370.29	140	.000	.000	.001	-0.78	0.95	2

Note. Model fit indices for baseline MIMIC model and stage 1 sequential models. In baseline model, all item parameters were constrained to be equal across groups. In stage 1 sequential models, each item was released individually and model fit indices were compared to the baseline model. Each item resulting in a significant Likelihood Ratio Test upon release was flagged for potential DIF.
BL = Baseline; RMSEA = Root Square Mean Error of Approximation; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; AIC = Akaike Information Criterion; *df* = degrees of freedom.
* $p < .01$
** $p < .001$

Table 10

Fit Indices for Stage 2 Sequential MIMIC Models Comparing Heterosexual and Bisexual Men

Model	RMSEA	CFI	TLI	AIC	X^2	*df*	ΔRMSEA	ΔCFI	ΔTLI	ΔAIC	ΔX^2	Δ*df*
Baseline Model												
Baseline	.082	.895	.857	-263792.8	375.32	126	–	–	–	–	–	–
Stage 2 Sequential Models												
Unstable Relationships	.082	.894	.859	-263664.7	374.37	128	.000	.001	.002	-3.04	0.95	2
Affective Instability	*Anchor item*											
Efforts to Avoid Abandonment	.082	.894	.859	-263676.5	362.68	128	.000	.001	.002	8.76	12.64	2
Anger	.082	.894	.859	-263665.0	374.06	128	.000	.001	.002	-2.73	1.26*	2
Identity Disturbance	.082	.894	.859	-263665.2	373.82	128	.000	.001	.002	-2.49	1.50	2
Emptiness	.082	.894	.859	-263671.9	367.19	128	.000	.001	.002	4.22	8.13	2
Paranoid Ideation	.081	.895	.859	-263664.2	374.78	128	.001	.000	.002	-3.46	0.54	2
Suicidality	.082	.894	.858	-263732.9	306.95	128	.000	.001	.001	65.19	68.37**	2
Impulsivity	.081	.895	.859	-263666.9	372.14	128	.001	.000	.002	-0.78	3.18	2

Note. Model fit indices for baseline MIMIC model and stage 2 sequential models. Anchor item response was constrained to be equal across groups and all other items varied freely. Then, each non-anchor item was iteratively tested for DIF using a 2*df* test in which both the slope and threshold were constrained to be equal across groups. Model fit indices were then compared against the fit of the baseline model. Each item that resulted in significantly worse fit ($p < .01$) according to Likelihood Ratio Tests after being constrained was flagged for DIF.
RMSEA = Root Square Mean Error of Approximation; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; AIC = Akaike Information Criterion; *df* = degrees of freedom.
* $p < .01$
** $p < .001$

Table 11

DIF and ESSD Results for MIMIC Models Comparing Heterosexual and Bisexual Men

Criterion	Factor Std. Ld.	Uniform *b*	Non-Uniform ω	ESSD	Uniform *b*	Non-Uniform ω	PMS *b*	PMS*BPD *b*	ESSD
	Final Model				*Perceived Minority Stress Model*				
Unstable Relationships	.636	–	–	–	–	–	–	–	–
Affective Instability	.731	–	–	–	–	–	–	–	–
Efforts to Avoid Abandonment	–	.008	.027	.044	.010	.007	-.011	.012	.064
Anger	.664	–	–	–	–	–	–	–	–
Identity Disturbance	.703	–	–	–	–	–	–	–	–
Emptiness	.702	–	–	–	–	–	–	–	–
Paranoid Ideation	.709	–	–	–	–	–	–	–	–
Suicidality	–	.021	.066	.094	.017	-.096	-.010	.176	.061
Impulsivity	.637	–	–	–	–	–	–	–	–

Note. For each item flagged for DIF, nonuniform DIF was tested using a 1*df* test in which the slope was constrained to be equal. If results suggested the presence of nonuniform DIF for a given item, then that item was not tested for uniform DIF. If nonuniform DIF was not detected, then a separate 1*df* test was conducted, in which the threshold was constrained to be equal. If the resulting *b* value was significant ($p < .01$), then the item was deemed to exhibit uniform DIF. Magnitude of DIF among flagged items is evaluated by the ESSD statistic, with .2, .4, and .6 reflecting small, medium, and large effect sizes, respectively. The effect of DIF is only considered practically meaningful at the item-level if the ESSD value is ≥ .2.

b = standardized regression coefficient from item to latent factor; ω = standardized regression coefficient from item to latent trait-by-group interaction term; ESSD = Expected score standardized difference; PMS = Perceived minority stress.

* $p < .01$

** $p < .001$

Table 12

Fit Indices for Stage 1 Sequential MIMIC Models Comparing Heterosexual and Lesbian Women

Model	BL Std. Ld.	RMSEA	CFI	TLI	AIC	X^2	*df*	ΔRMSEA	ΔCFI	ΔTLI	ΔAIC	ΔX^2	Δ*df*
		Baseline Model											
Baseline		.052	.956	.947	-303137.0	300.91	142	–	–	–	–	–	–
		Stage 1 Sequential Models											
Unstable Relationships	.660	.053	.956	.947	-303134.8	296.80	140	.001	.000	.000	-2.21	0.95	2
Affective Instability	.753	.053	.956	.947	-303145.8	296.60	140	.001	.000	.000	8.78	6.02	2
Efforts to Avoid Abandonment	.728	.053	.956	.947	-303146.3	296.69	140	.001	.000	.000	9.27	3.90	2
Anger	.711	.053	.956	.947	-303135.3	296.93	140	.001	.000	.000	-1.66	0.66	2
Identity Disturbance	.725	.053	.956	.947	-303134.1	297.18	140	.001	.000	.000	-2.85	0.19	2
Emptiness	.736	.053	.956	.947	-303138.9	296.85	140	.001	.000	.000	1.94	1.70	2
Paranoid Ideation	.734	.053	.956	.947	-303139.4	296.96	140	.001	.000	.000	2.38	1.32	2
Suicidality	.481	.053	.957	.947	-303188.2	296.12	140	.001	.001	.000	51.18	8.35	2
Impulsivity	.678	.053	.956	.947	-303147.0	296.70	140	.001	.000	.000	10.04	3.85	2

Note. Model fit indices for baseline MIMIC model and stage 1 sequential models. In baseline model, all item parameters were constrained to be equal across groups. In stage 1 sequential models, each item was released individually and model fit indices were compared to the baseline model. Each item resulting in a significant Likelihood Ratio Test upon release was flagged for potential DIF.
BL = Baseline; RMSEA = Root Square Mean Error of Approximation; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; AIC = Akaike Information Criterion; *df* = degrees of freedom.
* $p < .01$
** $p < .001$

Table 13

Fit Indices for Stage 2 Sequential MIMIC Models Comparing Heterosexual and Lesbian Women

Model	RMSEA	CFI	TLI	AIC	X^2	*df*	ΔRMSEA	ΔCFI	ΔTLI	ΔAIC	ΔX^2	Δ*df*
Baseline Model												
Baseline	.055	.957	.942	-303284.7	300.91	126	–	–	–	–	–	–
Stage 2 Sequential Models												
Unstable Relationships	.055	.957	.942	-303134.8	299.14	128	.000	.000	.000	-2.21	1.773	2
Affective Instability	*Anchor item*											
Efforts to Avoid Abandonment	.055	.957	.942	-303146.3	287.73	128	.000	.000	.000	9.27	13.18*	2
Anger	.055	.957	.942	-303135.3	298.60	128	.000	.000	.000	-1.66	2.32	2
Identity Disturbance	.055	.957	.942	-303134.1	299.77	128	.000	.000	.000	-2.85	1.14	2
Emptiness	.055	.957	.942	-303138.9	295.01	128	.000	.000	.000	1.94	5.90	2
Paranoid Ideation	.055	.957	.942	-303139.4	294.58	128	.000	.000	.000	2.38	6.33	2
Suicidality	.055	.956	.942	-303188.2	246.29	128	.000	.001	.000	51.18	54.62**	2
Impulsivity	.055	.957	.942	-303147.0	287.03	128	.000	.000	.000	10.04	13.88*	2

Note. Model fit indices for baseline MIMIC model and stage 2 sequential models. Anchor item response was constrained to be equal across groups and all other items varied freely. Then, each non-anchor item was iteratively tested for DIF using a 2*df* test in which both the slope and threshold were constrained to be equal across groups. Model fit indices were then compared against the fit of the baseline model. Each item that resulted in significantly worse fit ($p < .01$) according to Likelihood Ratio Tests after being constrained was flagged for DIF.
RMSEA = Root Square Mean Error of Approximation; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; AIC = Akaike Information Criterion; *df* = degrees of freedom.
* $p < .01$
** $p < .001$

Table 14

DIF and ESSD Results for MIMIC Models Comparing Heterosexual and Lesbian Women

Criterion	Factor Std. Ld.	Uniform *b*	Non-Uniform ω	ESSD	Uniform *b*	Non-Uniform ω	PMS *b*	PMS*BPD *b*	ESSD
	Final Model				*Perceived Minority Stress Model*				
Unstable Relationships	.660	–	–	–	–	–	–	–	–
Affective Instability	.753	–	–	–	–	–	–	–	–
Efforts to Avoid Abandonment	–	-.004	.025	.024	.004	-.037	-.006	.067	.108
Anger	.711	–	–	–	–	–	–	–	–
Identity Disturbance	.725	–	–	–	–	–	–	–	–
Emptiness	.736	–	–	–	–	–	–	–	–
Paranoid Ideation	.734	–	–	–	–	–	–	–	–
Suicidality	–	.016	.153	.079	-.001	.204	.006	-.153	-.387
Impulsivity	–	.010	.027	.060	.007	.268	-.001	-.250	-.536

Note. For each item flagged for DIF, nonuniform DIF was tested using a 1*df* test in which the slope was constrained to be equal. If results suggested the presence of nonuniform DIF for a given item, then that item was not tested for uniform DIF. If nonuniform DIF was not detected, then a separate 1*df* test was conducted, in which the threshold was constrained to be equal. If the resulting *b* value was significant ($p < .01$), then the item was deemed to exhibit uniform DIF. Magnitude of DIF among flagged items is evaluated by the ESSD statistic, with .2, .4, and .6 reflecting small, medium, and large effect sizes, respectively. The effect of DIF is only considered practically meaningful at the item-level if the ESSD value is ≥ .2.

b = standardized regression coefficient from item to latent factor; ω = standardized regression coefficient from item to latent trait-by-group interaction term; ESSD = Expected score standardized difference; PMS = Perceived minority stress.

* $p < .01$

** $p < .001$

Table 15

Fit Indices for Stage 1 Sequential MIMIC Models Comparing Heterosexual and Bisexual Women

Model	BL Std. Ld.	RMSEA	CFI	TLI	AIC	X^2	*df*	ΔRMSEA	ΔCFI	ΔTLI	ΔAIC	ΔX^2	Δ*df*
		Baseline Model											
Baseline		.065	.948	.937	-125058.9	604.32	142	–	–	–	–	–	–
		Stage 1 Sequential Models											
Unstable Relationships	.663	.061	.948	.936	-125055.6	596.69	140	.004	.000	.001	-3.32	0.28	2
Affective Instability	.757	.061	.948	.936	-125056.4	596.67	140	.004	.000	.001	-2.52	0.60	2
Efforts to Avoid Abandonment	.728	.061	.948	.936	-125058.5	597.19	140	.004	.000	.001	-0.41	0.87	2
Anger	.716	.061	.948	.936	-125083.1	595.87	140	.004	.000	.001	24.20	8.16	2
Identity Disturbance	.727	.061	.948	.936	-125077.9	597.10	140	.004	.000	.001	18.97	3.73	2
Emptiness	.737	.061	.948	.936	-125075.1	596.13	140	.004	.000	.001	16.15	6.28	2
Paranoid Ideation	.736	.061	.948	.936	-125064.6	597.11	140	.004	.000	.001	5.65	2.08	2
Suicidality	.465	.061	.948	.937	-125157.9	593.52	140	.004	.000	.000	99.00	17.39**	2
Impulsivity	.684	.061	.948	.937	-125135.1	593.78	140	.004	.000	.000	76.17	20.34**	2

Note. Model fit indices for baseline MIMIC model and stage 1 sequential models. In baseline model, all item parameters were constrained to be equal across groups. In stage 1 sequential models, each item was released individually and model fit indices were compared to the baseline model. Each item resulting in a significant Likelihood Ratio Test upon release was flagged for potential DIF.
BL = Baseline; RMSEA = Root Square Mean Error of Approximation; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; AIC = Akaike Information Criterion; *df* = degrees of freedom.
* $p < .01$
** $p < .001$

Table 16

Fit Indices for Stage 2 Sequential MIMIC Models Comparing Heterosexual and Bisexual Women

Model	RMSEA	CFI	TLI	AIC	X^2	*df*	ΔRMSEA	ΔCFI	ΔTLI	ΔAIC	ΔX^2	Δ*df*
Baseline Model												
Baseline	.063	.950	.932	-125524.9	604.32	126	–	–	–	–	–	–
Stage 2 Sequential Models												
Unstable Relationships	.063	.950	.933	-125055.6	603.65	128	.000	.000	.001	-3.32	0.68	2
Affective Instability	*Anchor item*											
Efforts to Avoid Abandonment	.063	.949	.932	-125058.5	600.80	128	.000	.001	.000	-0.41	3.52	2
Anger	.063	.949	.932	-125083.1	577.13	128	.000	.001	.000	24.20	27.19**	2
Identity Disturbance	.063	.949	.932	-125077.9	582.11	128	.000	.001	.000	18.97	22.22**	2
Emptiness	.063	.950	.933	-125075.1	584.70	128	.000	.000	.000	16.15	19.63**	2
Paranoid Ideation	.063	.949	.932	-125064.6	594.89	128	.000	.001	.000	5.65	9.43*	2
Suicidality	.063	.949	.932	-125157.9	505.26	128	.000	.001	.000	99.00	99.07**	2
Impulsivity	.064	.949	.932	-125135.1	527.19	128	.001	.001	.000	76.17	77.13**	2

Note. Model fit indices for baseline MIMIC model and stage 2 sequential models. Anchor item response was constrained to be equal across groups and all other items varied freely. Then, each non-anchor item was iteratively tested for DIF using a 2*df* test in which both the slope and threshold were constrained to be equal across groups. Model fit indices were then compared against the fit of the baseline model. Each item that resulted in significantly worse fit ($p < .01$) according to Likelihood Ratio Tests after being constrained was flagged for DIF.
RMSEA = Root Square Mean Error of Approximation; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; AIC = Akaike Information Criterion; *df* = degrees of freedom.
* $p < .01$
** $p < .001$

Table 17

DIF and ESSD Results for MIMIC Models Comparing Heterosexual and Bisexual Women

Criterion	Factor Std. Ld.	Uniform *b*	Non-Uniform ω	ESSD	Uniform *b*	Non-Uniform ω	PMS *b*	PMS* BPD *b*	ESSD
	Final Model				*Perceived Minority Stress Model*				
Unstable Relationships	.663	–	–	–	–	–	–	–	–
Affective Instability	.756	–	–	–	–	–	–	–	–
Efforts to Avoid Abandonment	.727	–	–	–	–	–	–	–	–
Anger	–	–	.054**	.058	–	.001	-.010	.055	.001
Identity Disturbance	–	.003	.049	.020	.019	1.154	-.073	-1.072	1.175
Emptiness	–	-.013	.012	-.056	-.015	-.017	-.008	.039	-.080
Paranoid Ideation	–	-.002	.038	-.001	-.052	-.800	.058	.839	-.965
Suicidality	–	–	.085**	.106	–	.261	-.013	-.174	.205
Impulsivity	–	–	.075**	.101	–	-.126	-.001	.195	-.114

Note. For each item flagged for DIF, nonuniform DIF was tested using a 1*df* test in which the slope was constrained to be equal. If results suggested the presence of nonuniform DIF for a given item, then that item was not tested for uniform DIF. If nonuniform DIF was not detected, then a separate 1*df* test was conducted, in which the threshold was constrained to be equal. If the resulting *b* value was significant ($p < .01$), then the item was deemed to exhibit uniform DIF. Magnitude of DIF among flagged items is evaluated by the ESSD statistic, with .2, .4, and .6 reflecting small, medium, and large effect sizes, respectively. The effect of DIF is only considered practically meaningful at the item-level if the ESSD value is ≥ .2.

b = standardized regression coefficient from item to latent factor; ω = standardized regression coefficient from item to latent trait-by-group interaction term; ESSD = Expected score standardized difference; PMS = Perceived minority stress.

* $p < .01$

** $p < .001$

Table 18

Hypothesis Testing in Primary and Sensitivity Analyses

	H x NH	**H x G/L**	**H x B**	**H x G/L**	**H x B**
		Men		Women	
Efforts to Avoid Abandonment (H1a)	Not Supported	Not Supported	Not Supported	Not Supported	Not Supported
Identity Disturbance (H1b)	Not Supported	Not Supported	Not Supported	Not Supported	Not Supported
Suicidality (H1c)	Not Supported	Not Supported	Not Supported	Not Supported	Not Supported
Impulsivity (H1d)	Not Supported	Not Supported	Not Supported	Not Supported	Not Supported
Efforts to Avoid Abandonment with PMS (H2a)	***Supported***	Not Supported	Not Supported	Not Supported	***Mixed Support****
Suicidality with PMS (H2b)	***Supported***	***Mixed Support****	Not Supported	***Mixed Support****	***Supported***

Note. H = heterosexual; NH = non-heterosexual; G/L = gay/lesbian, B = bisexual; PMS = perceived minority stressors.
* Supported in sensitivity analyses only.

Chapter Four:

Discussion

In this study, a MIMIC model approach was used to identify DIF among BPD criteria across heterosexual and non-heterosexual men and women, as well as to evaluate whether any observed DIF was explained by perceived minority stressors. Primary and sensitivity analyses showed that DIF was evident for some criteria in some sexual minority group comparisons but was not practically meaningful in any case, nor at the overall scale level. The observed DIF was most robust for suicidality, impulsivity, and efforts to avoid abandonment. This is in line with the notion that suicidal and NSSI behaviors, impulsive behaviors, and rejection sensitivity play an important role in non-heterosexual individuals' elevated BPD rates. However, the positive values of regression coefficients representing criterion slopes indicates that these criteria exhibit greater discrimination for some types of non-heterosexual individuals compared to heterosexual individuals. A difference in item discrimination (i.e., nonuniform DIF) occurs when one group has an advantage at some levels of BPD severity but is disadvantaged at other levels of BPD severity. In other words, there is an interaction between sexual orientation and BPD level, such that as the level of BPD severity changes the magnitude of change in endorsement for each item may differ across groups. Thus, the direction of DIF observed here is not consistent with hypotheses pertaining to Research Question 1. Instead of lower thresholds for sexual minority focal groups, most models displayed evidence that differentially-functioning criteria were biased in the opposite direction, such that they were more predictive of overall BPD level for non-heterosexual individuals compared to heterosexual individuals.

Effect size estimates below the .2 threshold indicated that the observed DIF is not practically meaningful for any criterion or at the overall BPD scale level. Consequently, group comparisons regarding BPD symptoms across sexual orientation groups without adjustment for DIF are unlikely to have large-scale impacts research and clinical practice. Nonetheless, researchers and clinicians interested in maximizing measurement precision may consider adjusting for these small amounts of DIF. Clinicians may choose to weight items differently depending on their differential factor loadings across groups to account for variation in the representativeness of the overall BPD construct across items. For instance, they may allow items pertaining to suicidality to account for a greater proportion of a client's total score if they endorse a non-heterosexual orientation for which factor loadings for these criteria are higher (e.g., in the case of bisexual women). This can be accomplished by adjusting factor scores based on case-specific measurement error as outlined in Lai et al. (2022), for example. Such coding schemes could be programmed into apps or computer software to reduce burden on clinicians.

One notable finding is that contrary to hypotheses, identity disturbance evidenced little DIF across models. Identity disturbance was even chosen as the anchor item in one set of sensitivity analysis models due to its high factor loading and non-significance for DIF. This suggests that despite the unique processes associated with non-heterosexual identity development, this aspect of self-conceptualization does not appear to influence the manner in which non-heterosexual men and women respond to BPD items measuring this criterion. One notable exception may be gay men, although as with other cases of DIF, it is not meaningfully impactful.

Another notable finding was that across primary and sensitivity analyses, the greatest amount of DIF was revealed for bisexual women, while the least amount of DIF was found for

bisexual men. This pattern of heterogeneity across genders underscores the diversity of experiences among bisexual men and women. The finding that perceived minority stress explained DIF among bisexual women suggests that minority stressors may differentially impact bisexual men and women in some ways. More broadly, the diversity of findings across comparisons of different groups highlights the heterogeneity across sexual minority groups and underscores importance of evaluating gay, lesbian, bisexual, and heterosexual men and women individually in sexual minority mental health research.

Given that differential functioning of BPD criteria appears to be an insufficient explanation for the discrepancies in BPD prevalence across sexual orientation groups—and as such, mean-level differences can generally be taken to accurately reflect these differences—next steps in this research area may continue to identify factors contributing to the disparities in BPD prevalence across groups.

In all cases, perceived minority stress appeared to account for the observed patterns of DIF. This finding is notable in light of ongoing efforts to disentangle biological "common cause" and culture-specific influences on sexual minority mental health difficulties (Bailey, 2020; Meyer et al., 2021; Pachankis et al., 2018), as it suggests that perceived minority stressors may have some degree, albeit slight, of influence over how non-heterosexual individuals rate their symptoms of personality pathology. Although this influence is not impactful enough to negatively impact group comparisons, future research into this topic may further inform researchers' understanding of the impact on minority stress on sexual minority mental health.

Strengths and Limitations

This study has several notable strengths. First, it is the first known study to test the assumption of measurement invariance of BPD features across sexual orientation groups.

Second, the large, nationally representative sample is a rare and valuable strength given the low population prevalence rate of non-heterosexuality. Third, the recognition of heterogeneity across sexual minority groups and partitioning of lesbian, gay, and bisexual men and women in analyses is an important advantage over other studies that have historically treated sexual minorities as homogenous. Finally, the inclusion of sensitivity analyses to evaluate DIF under less stringent requirements helped elucidate the robustness of DIF findings and may help inform future research investigating measurement invariance of dimensional BPD measures. While the interview-style AUDADIS-5 may not fully represent typical response patterns on dimensional BPD measures, results may provide a starting point for future research evaluating measurement invariance in dimensional, trait-based self-report measures of BPD features. Future research in this area is highly encouraged to increase confidence in the validity of observed group differences in studies using dimensional measurement approaches.

The present study also has several limitations. Among several models, fit indices fell below cutoffs for good fit (e.g., RMSEA values > .06 and CFI values < .95). This was especially true for models involving bisexual men, with CFI values of .894-.895 in primary analyses. Fit indices for models that included pathways for minority stress were extremely poor (e.g., CFI values < .400, RMSEA values ≥ .200), which may be due to the large number of parameters. Results from these models should be interpreted cautiously.

Despite the large sample size, some sexual orientation focal groups were small (e.g., 144 bisexual men). As described by Woods et al. (2009), parameter estimates may be less accurate when focal group sample sizes are under approximately 100. This is particularly true for models with many parameters, as was true for the current study. Thus, while all group sizes were above

100 in this study, parameter estimates for models involving some groups of relatively small size (e.g., bisexual men) may have a higher probability of inaccuracy.

In addition, NESARC-III participants were not queried about their gender identity. As a result, it was not possible to identify which participants are transgender or otherwise non-cisgender. Sexual minorities are significantly more likely than heterosexual individuals to identify with a gender that differs from their sex assigned at birth (James et al., 2016) and as many as 77% of transgender individuals report being non-heterosexual (James et al., 2016), therefore it is highly likely that at least some sexual minority participants in the NESARC-III sample were transgender. Transgender individuals are also diagnosed with BPD at higher rates and may share the same or similar characteristics and minority-specific processes as sexual minority individuals (e.g., complex identity development, greater suicidality; Anzani et al., 2020; Goldhammer et al., 2019; Russell et al., 2017). It is therefore possible that some observed DIF may be partially explained by transgender-specific processes, representing a potential confound. While examination of measurement invariance among transgender and gender-expansive individuals was beyond the scope of the present study and not possible using the NESARC-III dataset, this will be an important area for future investigation.

The study is also limited by its use of a novel interview-style measure of BPD that was created for the NESARC-III study and is thus not widely used in research or clinical settings. However, items included in the AUDADIS-5 are based on BPD criteria as described in the DSM-5 and thus follow very similar structure to the Structural Clinical Interview for DSM (SCID-II; First, 2014), a widely used diagnostic tool in both research and clinical contexts. Items in the AUDADIS-5 are nearly identical to the nine items in the commonly used SCID-II screener. Relatedly, interviews were conducted by lay persons rather than licensed mental health

professionals. Although interviewers underwent extensive training, it is possible that their lack of expertise in mental health conditions led them to more readily overlook information suggesting that participants' endorsement of BPD criteria may reflect culture-specific factors rather than true elevations in personality pathology.

NESARC-III participants were not given the option to endorse other sexual minority orientations, such as asexual, pansexual, and queer. These identities are under-researched, but some extant research suggests asexual individuals may score higher than non-asexual individuals on diagnostic measures of personality pathology (Brotto et al., 2010). It is possible that criterion bias may also exist for members of these groups. This will be an important area for future research.

Finally, and importantly, this study utilized a cross-sectional design and was thus unable to establish directionality or account for other possible third variables, precluding causal conclusions. These results should be considered in the context of this limitation.

Chapter Five:

Conclusion

In sum, extant research has found consistent evidence of elevated BPD features among non-heterosexual individuals across both clinical and community samples. Measurement invariance testing is an important prerequisite to interpreting observed disparities across sexual orientation groups. The current study used a MIMIC approach to show that some BPD criteria may function differently across some sexual orientation groups, although these differences are small and not impactful enough to jeopardize the validity of group comparisons. DIF was most consistently indicated for suicidality, efforts to avoid abandonment, and impulsivity and was nonuniform in most cases. Contrary to predictions, these items were better able to discriminate among non-heterosexual groups compared to heterosexual men and women. All DIF was explained by perceived minority stressors, although this finding should be interpreted cautiously due to poor model fit when perceived minority stressors was included. Results suggest that researchers and clinicians can therefore proceed with making group comparisons using the AUDADIS-5 and likely other similarly structured measures without concern about practically meaningful measurement non-invariance. In all, this study fills a critical gap in extant research sexual minority mental health. Results provide important context to help researchers and clinicians better understand the nature of problems unique to sexual minorities and their relevance to diagnostic assessment of BPD symptoms.

References

Alarie, M., & Gaudet, S. (2013). “I don't know if she is bisexual or if she just wants to get attention”: Analyzing the various mechanisms through which emerging adults invisibilize bisexuality. *Journal of Bisexuality, 13*(2), 191-214. https://doi-org/10.1080/15299716.2013.780004

Alexander, N., & Clare, L. (2004). You still feel different: The experience and meaning of women’s self-injury in the context of a lesbian or bisexual identity. *Journal of Community & Applied Social Psychology, 14*(2), 70–84. https://doi.org/10.1002/casp.764

Allen, M. S., & Robson, D. A. (2020). Personality and sexual orientation: New data and meta-analysis. *The Journal of Sex Research, 57*(8), 953–965. https://doi.org/10.1080/00224499.2020.1768204

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders: DSM-5*. 5th ed. Washington, D.C.: American Psychiatric Association.

American Psychiatric Association (1980). *Diagnostic and statistical manual of mental disorders, DSM-III*. 3rd ed. Washington, DC: American Psychiatric Association.

American Psychiatric Association (1987). *Diagnostic and statistical manual of mental disorders, DSM-III-R*. 3rd ed. revised. Washington, DC: American Psychiatric Association.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders: DSM-IV*. 4th ed. Washington, DC: American Psychiatric Association.

Anzani, A., Panfilis, C. D., Scandurra, C., & Prunas, A. (2020). Personality disorders and personality profiles in a sample of transgender individuals requesting gender-affirming treatments. *International Journal of Environmental Research and Public Health, 17*(5), 1521. https://doi.org/10.3390/ijerph17051521

Bailey, J. M. (2020). The minority stress model deserves reconsideration, not just extension. *Archives of Sexual Behavior, 49*(7), 1-4. https://doi-org/10.1007/s10508-019-01606-9

Balsam, K. F., & Mohr, J. J. (2007). Adaptation to sexual orientation stigma: A comparison of bisexual and lesbian/gay adults. *Journal of Counseling Psychology*, *54*(3), 306–319. https://doi.org/10.1037/0022-0167.54.3.306

Batejan, K. L., Jarvi, S. M., & Swenson, L. P. (2015). Sexual orientation and non-suicidal self-injury: A meta-analytic review. *Archives of Suicide Research, 19*(2), 131-150. https://doi.org/10.1080/13811118.2014.957450

Black, D. W., Blum, N., Pfohl, B., & Hale, N. (2004). Suicidal behavior in borderline personality disorder: Prevalence, risk factors, prediction, and prevention. *Journal of Personality Disorders*, *18*(3), 226–239. https://doi.org/10.1521/pedi.18.3.226.35445

Blum, A. W., Lust, K., Christenson, G., & Grant, J. E. (2020). Links between sexuality, impulsivity, compulsivity, and addiction in a large sample of university students. *CNS Spectrums, 25*(1), 9-15. https://doi.org/10.1017/S1092852918001591

Birnholz, J. L., & Young, M. A. (2012). Differential item functioning for lesbians, bisexual, and heterosexual women in the center for epidemiological studies depression scale. *Assessment*, 19(4), 502-505. https://doi.org/10.1177/1073191112440032

Borsboom, D. (2006). When does measurement invariance matter?. *Medical Care, 44*(11), S176-S181. https://doi.org/10.1080/13811118.2014.957450

Bostwick, W., & Hequembourg, A. (2014). 'Just a little hint': Bisexual-specific microaggressions and their connection to epistemic injustices. *Culture, Health & Sexuality, 16*(5), 488-503. https://doi.org/10.1080/13691058.2014.889754

Boyle, E. A., Li, Y. I., & Pritchard, J. K. (2017). An expanded view of complex traits: From polygenic to omnigenic. *Cell, 169*, 1177–1186. https://doi.org/10.1016/j.cell.2017.05.038

Brewster, M. E., & Moradi, B. (2010). Perceived experiences of anti-bisexual prejudice: Instrument development and evaluation. *Journal of Counseling Psychology*, *57*(4), 451–468. https://doi.org/10.1037/a0021116

Brotto, L. A., Knudson, G., Inskip, J., Rhodes, K., & Erskine, Y. (2010). Asexuality: A mixed-methods approach. *Archives of Sexual Behavior*, *39*(3), 599–618. https://doi.org/10.1007/s10508-008-9434-x

Campbell, K., Clarke, K. A., Massey, D., & Lakeman, R. (2020). Borderline personality disorder: To diagnose or not to diagnose? That is the question. *International Journal of Mental Health Nursing, 29*(5), 972-981. https://doi.org/10.1111/inm.12737

Cass, V. C. (1979). Homosexual identity formation: A theoretical model. *Journal of Homosexuality, 4*(3), 219-235. https://doi.org/10.1300/J082v04n03_01

Cavicchioli, M., & Maffei, C. (2020). Rejection sensitivity in borderline personality disorder and the cognitive–affective personality system: A meta-analytic review. *Personality Disorders: Theory, Research, and Treatment, 11*(1), 1–12. https://doi.org/10.1037/per0000359

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

Cicero, D. C., & Ruggero, C. J. (2021). Commentary– Opening a can of worms: The importance of testing the measurement invariance of hierarchical models of psychopathology – a commentary on He and Li (2020). *Journal of Child Psychology and Psychiatry*, *62*(3), 299–302. https://doi.org/10.1111/jcpp.13353

Cochran, S. D., Sullivan, J. G., & Mays, V. M. (2003). Prevalence of mental disorders, psychological distress, and mental health services use among lesbian, gay, and bisexual adults in the United States. *Journal of Consulting and Clinical Psychology, 71*(1), 53–61. https://doi.org/10.1037/0022-006X.71.1.53

Crowell, S. E., Beauchaine, T. P., & Linehan, M. M. (2009). A biosocial developmental model of borderline personality: Elaborating and extending Linehan's theory. *Psychological Bulletin, 135*(3), 495. https://doi.org/10.1037/a0015616

Davis, K. C., & Anderson, J. L. (2021). The role of perceived parental acceptance-rejection on personality psychopathology in sexual orientation minority young adults. *Journal of Child and Family Studies*, *30*(1), 276–291. https://doi.org/10.1007/s10826-020-01858-3

Dodge, B., Herbenick, D., Friedman, M. R., Schick, V., Fu, T.-C., Bostwick, W., Bartelt, E., Muñoz-Laboy, M., Pletta, D., Reece, M., & Sandfort, T. G. M. (2016). Attitudes toward bisexual men and women among a nationally representative probability sample of adults in the United States. *PLOS ONE, 11*(10), e0164430. https://doi.org/10.1371/journal.pone.0164430

Dyar, C., Feinstein, B. A., Eaton, N. R., & London, B. (2018). The mediating roles of rejection sensitivity and proximal stress in the association between discrimination and internalizing symptoms among sexual minority women. *Archives of Sexual Behavior, 47*(1), 205-218. https://doi.org/10.1007/s10508-016-0869-1

Eaton, N. R. (2014). Trans-diagnostic psychopathology factors and sexual minority mental health: evidence of disparities and associations with minority stressors. *Psychology of Sexual Orientation and Gender Diversity, 1*(3), 244–254. https://doi.org/10.1037/sgd0000048

Eldahan, A. I., Pachankis, J. E., Rendina, H. J., Ventuneac, A., Grov, C., & Parsons, J. T. (2016). Daily minority stress and affect among gay and bisexual men: A 30-day diary study. *Journal of Affective Disorders, 190*, 828-835. https://doi.org/10.1016/j.jad.2015.10.066

Eliason, M. J. (1996). Identity formation for lesbian, bisexual, and gay persons: Beyond a minoritizing view. *Journal of Homosexuality, 30*(3), 31-58. https://doi.org/10.1300/J082v30n03_03

Ellison, W. D., Rosenstein, L. K., Morgan, T. A., & Zimmerman, M. (2018). Community and clinical epidemiology of borderline personality disorder. *Psychiatric Clinics, 41*(4), 561-573. https://doi.org/10.1016/j.psc.2018.07.008

Erikson, E. H. (1968). *Identity, youth and crisis*. W. W. Norton & Company.

Fehling, K. B. (2019). *An ecological momentary assessment study of sexual minority stress and nonsuicidal self-injury in sexual minority adults* (Unpublished doctoral dissertation, Rutgers University).

Feinstein, B. A. (2020). The rejection sensitivity model as a framework for understanding sexual minority mental health. *Archives of Sexual Behavior*, *49*(7), 2247–2258. https://doi.org/10.1007/s10508-019-1428-3

Feinstein, B. A., & Dyar, C. (2017). Bisexuality, minority stress, and health. *Current Sexual Health Reports, 9*(1), 42–49. https://doi.org/10.1007/s11930-017-0096-3

Feisthamel, K., & Schwartz, R. (2009). Differences in mental health counselors' diagnoses based on client race: An investigation of adjustment, childhood, and substance-related disorders. *Journal of Mental Health Counseling, 31*(1), 47-59.

First, M. B. (2014). Structured clinical interview for the DSM (SCID). *The Encyclopedia of Clinical Psychology*, 1-6.

Fleishman, J., Spector, W., & Altman, B. (2002). Impact of differential item functioning on age and gender differences in functional disability. *Journal of Gerontology Series B: Psychological Sciences and Social Sciences, 57*, S275-S284. https://doi.org/10.1093/geronb/57.5.S275

Ganna, A., Verweij, K. J. H., Nivard, M. G., Maier, R., Wedow, R., Busch, A. S., Abdellaoui, A., Guo, S., Sathirapongsasuti, J. F., Team16, 23andMe Research, Lichtenstein, P., Lundström, S., Långström, N., Auton, A., Harris, K. M., Beecham, G. W., Martin, E. R., Sanders, A. R., Perry, J. R. B., … Zietsch, B. P. (2019). Large-scale GWAS reveals insights into the genetic architecture of same-sex sexual behavior. *Science*, *365*(6456). https://doi.org/10.1126/science.aat7693

Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences, 102*, 74-78. https://doi.org/10.1016/j.paid.2016.06.069

Glick, S. N., Cleary, S. D., & Golden, M. R. (2015). Brief report: increasing acceptance of homosexuality in the United States across racial and ethnic subgroups. *Journal of Acquired Immune Deficiency Syndromes, 70*(3), 319-322. https://doi.org/10.1097/QAI.0000000000000740

Goldhammer, H., Crall, C., & Keuroghlian, A. S. (2019). Distinguishing and addressing gender minority stress and borderline personality symptoms. *Harvard Review of Psychiatry, 27*(5), 317-325. https://doi.org/10.1097/HRP.0000000000000234

Gomez, R., & McLaren, S. (2017). The Center for Epidemiologic Studies Depression Scale: Invariance across heterosexual men, heterosexual women, gay men, and lesbians. *Psychological Assessment, 29*(4), 361. http://dx.doi.org/10.1037/pas0000352

Grant, B. F., Chu, A., Sigman, R., Amsbary, M., Kali, J., Sugawara, Y., Jiao, R., Ren, W., & Goldstein, R. (2014). Source and accuracy statement: national epidemiologic survey on alcohol and related conditions-III (NESARC-III). *Rockville, MD: National Institute on Alcohol Abuse and Alcoholism*, 1-125.

Grant, B. F., Goldstein, R. B., Smith, S. M., Jung, J., Zhang, H., Chou, S. P., ... & Aivadyan, C. (2015). The Alcohol Use Disorder and Associated Disabilities Interview Schedule-5 (AUDADIS-5): Reliability of substance use and psychiatric disorder modules in a general population sample. *Drug and Alcohol Dependence, 148*, 27-33. https://doi.org/10.1016/j.drugalcdep.2014.11.026

Grant, J. E., Flynn, M., Odlaug, B. L., & Schreiber, L. R. N. (2011). Personality disorders in gay, lesbian, bisexual and transgender chemically dependent patients. *The American Journal on Addictions / American Academy of Psychiatrists in Alcoholism and Addictions, 20*(5), 405–411. https://doi.org/10.1111/j.1521-0391.2011.00155.x

Gunderson, J. G., Stout, R. L., McGlashan, T. H., Shea, M. T., Morey, L. C., Grilo, C. M., Zanarini, M. C., Yen, S., Markowitz, J. C., Sanislow, C., Ansell, E., Pinto, A., & Skodol, A. E. (2011). Ten-year course of borderline personality disorder: psychopathology and function from the Collaborative Longitudinal Personality Disorders study. *Archives of General Psychiatry, 68*(8), 827-837. https://doi.org/10.1001/archgenpsychiatry.2011.37

Hatzenbuehler, M. L. (2009). How does sexual minority stigma "get under the skin"? A psychological mediation framework. *Psychological Bulletin, 135*(5), 707. https://doi.org/10.1037/a0016441

Hatzenbuehler, M. L. (2016). Structural stigma and health inequalities: Research evidence and implications for psychological science. *The American Psychologist*, *71*(8), 742–751. https://doi.org/10.1037/amp0000068

Hayfield, N., Clarke, V., & Halliwell, E. (2014). Bisexual women's understandings of social marginalisation: 'The heterosexuals don't understand us but nor do the lesbians.' *Feminism & Psychology, 24*(3), 352–372. https://doi.org/10.1177/0959353514539651

James, S. E., Herman, J. L., Rankin, S., Keisling, M., Mottet, L., & Anafi, M. (2016). *The Report of the 2015 U.S. Transgender Survey*. Washington, DC: National Center for Transgender Equality.

Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., & Rosseel, Y. (2021*). semTools: Useful tools for structural equation modeling.* R package version 0.5-4. Retrieved from https://CRAN.R-project.org/package=semTools

Kerridge, B. T., Pickering, R. P., Saha, T. D., Ruan, W. J., Chou, S. P., Zhang, H., Jung, J., & Hasin, D. S. (2017). Prevalence, sociodemographic correlates and DSM-5 substance use disorders and other psychiatric disorders among sexual minorities in the United States. *Drug and Alcohol Dependence, 170*, 82-92. https://doi.org/10.1016/j.drugalcdep.2016.10.038

King, M., Semlyen, J., Tai, S. S., Killaspy, H., Osborn, D., Popelyuk, D., & Nazareth, I. (2008). A systematic review of mental disorder, suicide, and deliberate self harm in lesbian, gay and bisexual people. *BMC Psychiatry*, *8*(1). https://doi.org/10.1186/1471-244X-8-70

Klonsky, E. D., & Muehlenkamp, J. J. (2007). Self-injury: A research review for the practitioner. *Journal of Clinical Psychology, 63*(11), 1045-1056. https://doi.org/10.1002/jclp.20412

Krieger, N., & Sidney, S. (1997). Prevalence and health implications of anti-gay discrimination: A study of black and white women and men in the CARDIA cohort. *International Journal of Health Services, 27*(1), 157-176. https://doi.org/10.2190/HPB8-5M2N-VK6X-0FWN

Krieger, N., Smith, K., Naishadham, D., Hartman, C., & Barbeau, E. M. (2005). Experiences of discrimination: Validity and reliability of a self-report measure for population health research on racism and health. *Social Science & Medicine, 61*(7), 1576-1596. https://doi.org/10.1016/j.socscimed.2005.03.006

Lai, M. H., Tse, W. W. Y., Zhang, G., Li, Y., & Hsiao, Y. Y. (2022). Correcting for Unreliability and Partial Invariance: A Two-Stage Path Analysis Approach. *Structural Equation Modeling: A Multidisciplinary Journal*, 1-14.

Lindner, P., Martell, C., Bergström, J., Andersson, G., & Carlbring, P. (2013). Clinical validation of a non-heteronormative version of the Social Interaction Anxiety Scale (SIAS). *Health and Quality of Life Outcomes, 11*(1), 209. https://doi-org/10.1186/1477-7525-11-209

Linehan, M. M. (1993). *Cognitive-behavioral treatment of borderline personality disorder*. New York: Guilford Press.

Marshal, M. P., Dermody, S. S., Shultz, M. L., Sucato, G. S., Stepp, S. D., Chung, T., Burton, C. M., Markovic, N., & Hipwell, A. E. (2013). Mental health and substance use disparities among urban adolescent lesbian and bisexual girls. *Journal of the American Psychiatric Nurses Association, 19*(5), 271–279. https://doi.org/10.1177/1078390313503552

Marshal, M. P., Friedman, M. S., Stall, R., King, K. M., Miles, J., Gold, M. A., Bukstein, O. G., & Morse, J. Q. (2008). Sexual orientation and adolescent substance use: A meta-analysis and methodological review. *Addiction, 103*(4), 546–556. https://doi.org/10.1111/j.1360-0443.2008.02149.x

Matsick, J. L., & Rubin, J. D. (2018). Bisexual prejudice among lesbian and gay people: Examining the roles of gender and perceived sexual orientation. *Psychology of Sexual Orientation and Gender Diversity*, *5*(2), 143–155. https://doi.org/10.1037/sgd0000283

Mattick, R. P., & Clarke, J. C. (1998). Development and validation of measures of social phobia scrutiny fear and social interaction anxiety. *Behaviour Research and Therapy, 36*(4), 455-470. https://doi.org/10.1016/S0005-7967(97)10031-6

McCabe, S. E., Hughes, T. L., Bostwick, W. B., West, B. T., & Boyd, C. J. (2009). Sexual orientation, substance use behaviors and substance dependence in the United States. *Addiction, 104*(8), 1333-1345. https://doi.org/10.1111/j.1360-0443.2009.02596.x

Meade, A. W. (2005). *Sample Size and Tests of Measurement Invariance*. Conference of the Society for Industrial and Organizational Psychology, Los Angeles, CA.

Meade, A. W. (2010). A taxonomy of effect size measures for the differential functioning of items and scales. *Journal of Applied Psychology, 95*(4), 728–743. https://doi.org/10.1037/a0018966

Mereish, E. H., Katz-Wise, S. L., & Woulfe, J. (2017). Bisexual-specific minority stressors, psychological distress, and suicidality in bisexual individuals: The mediating role of loneliness. *Prevention Science, 18*(6), 716–725. https://doi.org/10.1007/s11121-017-0804-2

Meyer, I. H. (2007). Prejudice, social stress, and mental health in lesbian, gay, and bisexual populations: Conceptual issues and research evidence. *Psychological Bulletin, 129*(5), 674-697. https://doi.org/10.1037/0033-2909.129.5.674

Meyer, I., Pachankis, J., & Klein, D. (2021). Do genes explain sexual minority mental health disparities? *Archives of Sexual Behavior. 50*(3), 731-737. https://doi-org/10.1007/s10508-020-01909-2

Mohr, J., & Rochlen, A. (1999). Measuring attitudes regarding bisexuality in lesbian, gay male, and heterosexual populations. *Journal of Counseling Psychology, 46*, 353–369. https://doi.org/10.1037/0022-0167.46.3.353

Mojola, S. A., & Everett, B. (2012). STD and HIV risk factors among US young adults: Variations by gender, race, ethnicity and sexual orientation. *Perspectives on Sexual and Reproductive Health, 44*(2), 125-133. https://doi.org/10.1363/4412512

Molina, Y., Marquez, J. H., Logan, D. E., Leeson, C. J., Balsam, K. F., & Kaysen, D. L. (2015). Current intimate relationship status, depression, and alcohol use among bisexual women: The mediating roles of bisexual-specific minority stressors. *Sex Roles, 73*(1), 43–57. https://doi.org/10.1007/s11199-015-0483-z

Mustanski, B., Andrews, R., Herrick, A., Stall, R., & Schnarrs, P. W. (2014). A syndemic of psychosocial health disparities and associations with risk for attempting suicide among young sexual minority men. *American Journal of Public Health*, *104*(2), 287–294. https://doi.org/10.2105/AJPH.2013.301744

Muthen, B. O. (1989). Latent variable modeling in heterogeneous populations. *Psychometrika, 54*, 557-585. https://doi.org/10.1007/BF02296397

Neacsiu, A. D., Eberle, J. W., Keng, S. L., Fang, C. M., & Rosenthal, M. Z. (2017). Understanding borderline personality disorder across sociocultural groups: Findings, issues, and future directions. *Current Psychiatry Reviews, 13*(3), 188-223. https://doi.org/10.2174/1573400513666170612122034

Nickels, S. J., Walls, N. E., Laser, J. A., & Wisneski, H. (2012). Differences in motivations of cutting behavior among sexual minority youth. *Child and Adolescent Social Work Journal, 29*(1), 41–59. https://doi.org/10.1007/s10560-011-0245-x

Ochs, R. (1996). Biphobia: It goes more than two ways. In B. A. Firestein (Ed.), *Bisexuality: The psychology and politics of an invisible minority* (pp. 217–239). Thousand Oaks, CA: Sage.

Ock, J., McAbee, S. T., Mulfinger, E., & Oswald, F. L. (2020). The practical effects of measurement invariance: Gender invariance in two big five personality measures. *Assessment, 27*(4), 657–674. https://doi.org/10.1177/1073191119885018

Pachankis, J. E. (2007). The psychological implications of concealing a stigma: A cognitive-affective-behavioral model. *Psychological Bulletin*, *133*(2), 328–345. https://doi.org/10.1037/0033-2909.133.2.328

Pachankis, J. E., Rendina, H. J., Restar, A., Ventuneac, A., Grov, C., & Parsons, J. T. (2015). A minority stress—emotion regulation model of sexual compulsivity among highly sexually active gay and bisexual men. *Health Psychology, 34*(8), 829. https://doi.org/10.1037/hea0000180

Pachankis, J., Sullivan, T., Feinstein, B., & Newcomb, M. (2018). Young adult gay and bisexual men's stigma experiences and mental health: An 8-year longitudinal study. *Developmental Psychology*, 54. https://doi.org/10.1037/dev0000518

Paris, J., Zweig-Frank, H., & Guzder, J. (1995). Psychological factors associated with homosexuality in males with borderline personality disorder. *Journal of Personality Disorders, 9*(1), 56-61. https://doi.org/10.1521/pedi.1995.9.1.56

Puckett, J. A., Surace, F. I., Levitt, H. M., & Horne, S. G. (2016). Sexual orientation identity in relation to minority stress and mental health in sexual minority women. *LGBT Health, 3*(5), 350-356. https://doi-org/1089/lgbt.2015.0088

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, 41*, 71–90. https://doi.org/10.1016/j.dr.2016.06.004

R Core Team. (2020). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Radloff, L. S. (1977). The CES-D Scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement, 1*, 385-401. https://doi.org/10.1177/014662167700100306

Radusky, P. D., Rodriguez, V. J., Kumar, M., & Jones, D. L. (2019). Differential item functioning by HIV status and sexual orientation of the Center for Epidemiological Studies–Depression Scale: An item response theory analysis. *Assessment, 28*(4) 1173-1185. https://doi.org/10.1177/1073191119887445

Raifman, J., Moscoe, E., Austin, S. B., Hatzenbuehler, M. L., & Galea, S. (2018). Association of state laws permitting denial of services to same-sex couples with mental distress in sexual minority adults: A difference-in-difference-in-differences analysis. *JAMA Psychiatry, 75*(7), 671-677. https://doi.org/10.1521/pedi.1995.9.1.56

Raifman, J., Moscoe, E., Austin, S. B., & McConnell, M. (2017). Difference-in-differences analysis of the association between state same-sex marriage policies and adolescent suicide attempts. *JAMA Pediatrics, 171*(4), 350-356. https://doi.org/10.1001/jamapediatrics.2016.4529

Reich, D. B., & Zanarini, M. C. (2008). Sexual orientation and relationship choice in borderline personality disorder over ten years of prospective follow-up. *Journal of Personality Disorders, 22*(6), 564-572. https://doi.org/10.1521/pedi.2008.22.6.564

Reuter, T. R., Sharp, C., Kalpakci, A. H., Choi, H. J., & Temple, J. R. (2016). Sexual orientation and borderline personality disorder features in a community sample of adolescents. *Journal of Personality Disorders, 30*(5), 694–707. https://doi.org/10.1521/pedi_2015_29_224

Revelle, W. (2020). *psych: Procedures for psychological, psychometric, and personality research*. R package version, 2.0.12

Rodriguez-Seijas, C., Morgan, T. A., & Zimmerman, M. (2020). Is there a bias in the diagnosis of borderline personality disorder among lesbian, gay, and bisexual patients? *Assessment, 28*(3), 724-738. 1073191120961833. https://doi.org/10.1177/1073191120961833

Rodriguez-Seijas, C., Morgan, T. A., & Zimmerman, M. (2021). A population-based examination of criterion-level disparities in the diagnosis of borderline personality disorder among sexual minority adults. *Assessment, 28*(4), 1097-1109. https://doi.org/10.1177/1073191121991922

Rosario, M., Schrimshaw, E. W., & Hunter, J. (2011). Different patterns of sexual identity development over time: Implications for the psychological adjustment of lesbian, gay, and bisexual youths. *Journal of Sex Research, 48*(1), 3-15. https://doi.org/10.1080/00224490903331067

Ross, L. E., Dobinson, C., & Eady, A. (2010). Perceived determinants of mental health for bisexual people: A qualitative examination. *American Journal of Public Health, 100*(3), 496-502. https://doi.org/10.2105/AJPH.2008.156307

Ross, L. E., Salway, T., Tarasoff, L. A., MacKay, J. M., Hawkins, B. W., & Fehr, C. P. (2018). Prevalence of depression and anxiety among bisexual people compared to gay, lesbian, and heterosexual individuals: A systematic review and meta-analysis. *The Journal of Sex Research, 55*(4-5), 435-456. https://doi.org/10.1080/00224499.2017.1387755

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software, 48*(2), 1-36. https://doi.org/10.18637/jss.v048.i02

Sakaluk, J. K. (2019). Expanding statistical frontiers in sexual science: Taxometric, invariance, and equivalence testing. *Journal of Sex Research, 56*(4-5), 475-510. https://doi.org/10.1080/00224499.2019.1568377

Salway, T., Ross, L. E., Fehr, C. P., Burley, J., Asadi, S., Hawkins, B., & Tarasoff, L. A. (2019). A systematic review and meta-analysis of disparities in the prevalence of suicide ideation and attempt among bisexual populations. *Archives of Sexual Behavior*, *48*(1), 89–111. http://dx.doi.org/10.1007/s10508-018-1150-6

Samuel, D. B., Carroll, K. M., Rounsaville, B. J., & Ball, S. A. (2013). Personality disorders as maladaptive, extreme variants of normal personality: Borderline personality disorder and neuroticism in a substance using sample. *Journal of Personality Disorders, 27*(5), 625–635. https://doi.org/10.1521/pedi.2013.27.5.625

Samuel, D. B., & Widiger, T. A. (2008). A meta-analytic review of the relationships between the five-factor model and DSM-IV-TR personality disorders: A facet level analysis. *Clinical Psychology Review, 28*(8), 1326-1342. https://doi.org/10.1016/j.cpr.2008.07.002

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika, 66*(4), 507–514. https://doi.org/10.1007/BF02296192

Scourfield, J., Roen, K., & McDermott, L. (2008). Lesbian, gay, bisexual and transgender young people's experiences of distress: Resilience, ambivalence and self-destructive behaviour. *Health & Social Care in the Community, 16*(3), 329–336. https://doi.org/10.1111/j.1365-2524.2008.00769.x

Shearer, A., Herres, J., Kodish, T., Squitieri, H., James, K., Russon, J., Atte, T., & Diamond, G. S. (2016). Differences in mental health symptoms across lesbian, gay, bisexual, and questioning youth in primary care settings. *Journal of Adolescent Health*, *59*(1), 38–43. https://doi.org/10.1016/j.jadohealth.2016.02.005

Singh, D., McMain, S., & Zucker, K. J. (2011). Gender identity and sexual orientation in women with borderline personality disorder. *The Journal of Sexual Medicine, 8*(2), 447–454. https://doi.org/10.1111/j.1743-6109.2010.02086.x

Slater, M. E., Godette, D., Huang, B., Ruan, W. J., & Kerridge, B. T. (2017). Sexual orientation-based discrimination, excessive alcohol use, and substance use disorders among sexual minority adults. *LGBT Health*, *4*(5), 337–344. https://doi.org/10.1089/lgbt.2016.0117

Taggart, T. C., Rodriguez-Seijas, C., Dyar, C., Elliott, J. C., Thompson, R. G., Hasin, D. S., & Eaton, N. R. (2019). Sexual orientation and sex-related substance use: The unexplored role of bisexuality. *Behaviour Research and Therapy*, *115*, 55–63. https://doi.org/10.1016/j.brat.2018.12.012

Trull, T. J., Jahng, S., Tomko, R. L., Wood, P. K., & Sher, K. J. (2010). Revised NESARC personality disorder diagnoses: Gender, prevalence, and comorbidity with substance dependence disorders. *Journal of Personality Disorders*, *24*(4), 412–426. https://doi.org/10.1521/pedi.2010.24.4.412

Widiger, T. A., & Spitzer, R. L. (1991). Sex bias in the diagnosis of personality disorders: Conceptual and methodological issues. *Clinical Psychology Review, 11*(1), 1-22. https://doi.org/10.1016/0272-7358(91)90135-H

Woods, C. M. (2009). Evaluation of MIMIC-model methods for DIF testing with comparison to two-group analysis. *Multivariate Behavioral Research, 44*(1), 1-27. https://doi.org/10.1080/00273170802620121

Woods, C. M., & Grimm, K. J. (2011). Testing for nonuniform differential item functioning with multiple indicator multiple cause models. *Applied Psychological Measurement, 35*(5), 339-361.

Woods, C. M., Oltmanns, T. F., & Turkheimer, E. (2009). Illustration of MIMIC-model DIF testing with the Schedule for Nonadaptive and Adaptive Personality. *Journal of Psychopathology and Behavioral Assessment*, *31*(4), 320–330. https://doi.org/10.1007/s10862-008-9118-9

Yarhouse, M. A., & Tan, E. S. (2004). *Sexual identity synthesis: Attributions, meaning-making, and the search for congruence*. University Press of America.

Yost, M. R., & Thomas, G. D. (2012). Gender and binegativity: Men's and women's attitudes toward male and female bisexuals. *Archives of Sexual Behavior, 41*, 691–702. https://doi.org/10.1007/s1050 8-011-9767-8

Zanarini, M. C., Magni, L. R., Temes, C. M., Hein, K. E., Aguirre, B. A., & Goodman, M. (2020). Sexual orientation and gender of intimate relationship partners among adolescents with BPD and psychiatrically healthy adolescents. *Journal of Personality Disorders*, 1-7. https://doi.org/10.1521/pedi_2020_34_470

Zanarini, M. C., Temes, C. M., Magni, L. R., Fitzmaurice, G. M., Aguirre, B. A., & Goodman, M. (2017). Prevalence rates of borderline symptoms reported by adolescent inpatients with BPD, psychiatrically healthy adolescents and adult inpatients with BPD. *Personality and Mental Health*, *11*(3), 150-156. https://doi.org/10.1002/pmh.1378

Zimmermann, J., Kerber, A., Rek, K., Hopwood, C. J., & Krueger, R. F. (2019). A brief but comprehensive review of research on the alternative DSM-5 model for personality disorders. *Current Psychiatry Reports, 21*, 1-19. https://doi.org/10.1007/s11920-019-1079-z

Zubenko, G. S., George, A. W., Soloff, P. H., & Schulz, P. (1987). Sexual practices among patients with borderline personality disorder. *The American Journal of Psychiatry, 144*(6), 748-752. https://doi-org/10.1176/ajp.144.6.748

Appendix A: Supplemental Results Tables

Table 19

Fit Indices for Stage 1 Sequential MIMIC Models Comparing Heterosexual and Non-Heterosexual Individuals (Supplemental)

Model	BL Std. Ld.	RMSEA	CFI	TLI	AIC	X^2	*df*	ΔRMSEA	ΔCFI	ΔTLI	ΔAIC	ΔX^2	Δ*df*
		Baseline Model											
Baseline		.050	.955	.946	290991.6	1457.46	142	–	–	–	–	–	–
		Stage 1 Sequential Models											
Unstable Relationships	.570	.050	.955	.945	290990.4	1438.45	140	.000	.000	.000	1.18	19.01	2
Affective Instability	.731	.050	.955	.945	290982.3	1438.41	140	.000	.000	.000	9.28	19.05*	2
Efforts to Avoid Abandonment	.658	.050	.955	.945	290953.7	1436.78	140	.000	.000	.000	37.93	20.68**	2
Anger	.680	.050	.955	.945	290987.6	1439.94	140	.000	.000	.000	4.05	17.52	2
Identity Disturbance	.689	.050	.955	.945	290975.3	1439.52	140	.000	.000	.000	16.31	17.94	2
Emptiness	.681	.050	.955	.945	290985.8	1439.35	140	.000	.000	.000	5.85	18.11	2
Paranoid Ideation	.690	.050	.955	.945	290970.4	1439.50	140	.000	.000	.000	21.18	17.96*	2
Suicidality	.436	.050	.956	.946	290613.3	1406.24	140	.000	.001	.001	378.27	51.22**	2
Impulsivity	.593	.050	.955	.945	290987.6	1438.97	140	.000	.000	.000	4.04	18.49	2

Note. Model fit indices for baseline MIMIC model and stage 1 sequential models. In baseline model, all item parameters were constrained to be equal across groups. In stage 1 sequential models, each item was released individually and model fit indices were compared to the baseline model. Each item resulting in a significant Likelihood Ratio Test upon release was flagged for potential DIF.
BL = Baseline; RMSEA = Root Square Mean Error of Approximation; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; AIC = Akaike Information Criterion; *df* = degrees of freedom.
* $p < .01$
** $p < .001$

Table 20

Fit Indices for Stage 2 Sequential MIMIC Models Comparing Heterosexual and Non-Heterosexual Individuals (Supplemental)

Model	RMSEA	CFI	TLI	AIC	X^2	*df*	ΔRMSEA	ΔCFI	ΔTLI	ΔAIC	ΔX^2	Δ*df*
Baseline Model												
Baseline	.052	.957	.942	290991.6	1457.46	126	–	–	–	–	–	–
Stage 2 Sequential Models												
Unstable Relationships	.051	.957	.943	290990.4	1452.44	128	.001	.000	.001	1.18	5.02	2
Affective Instability	.051	.957	.942	290982.3	1444.19	128	.001	.000	.000	9.28	13.27*	2
Efforts to Avoid Abandonment	.052	.957	.942	290953.7	1416.55	128	.000	.000	.000	37.93	40.91**	2
Anger	.051	.957	.942	290987.6	1449.38	128	.001	.000	.000	4.05	8.08	2
Identity Disturbance	*Anchor item*											
Emptiness	.051	.957	.942	290985.8	1447.6	128	.001	.000	.000	5.85	9.86*	2
Paranoid Ideation	.051	.957	.942	290970.4	1432.8	128	.001	.000	.000	21.18	24.66**	2
Suicidality	.052	.955	.940	290613.3	1087.12	128	.000	.002	.002	378.27	370.34**	2
Impulsivity	.051	.957	.943	290987.6	1449.55	128	.001	.000	.001	4.04	7.91	2

Note. Model fit indices for baseline MIMIC model and stage 2 sequential models. Anchor item response was constrained to be equal across groups and all other items varied freely. Then, each non-anchor item was iteratively tested for DIF using a 2*df* test in which both the slope and threshold were constrained to be equal across groups. Model fit indices were then compared against the fit of the baseline model. Each item that resulted in significantly worse fit ($p < .01$) according to Likelihood Ratio Tests after being constrained was flagged for DIF.
RMSEA = Root Square Mean Error of Approximation; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; AIC = Akaike Information Criterion; *df* = degrees of freedom.
* $p < .01$
** $p < .001$

Table 21

DIF and ESSD Results for MIMIC Models Comparing Heterosexual and Non-Heterosexual Individuals (Supplemental)

Criterion	Factor Std. Ld.	Uniform *b*	Non-Uniform ω	ESSD	Uniform *b*	Non-Uniform ω	PMS *b*	PMS*BPD *b*	ESSD
	Final Model				*Perceived Minority Stress Model*				
Unstable Relationships	.570	–	–	–	–	–	–	–	–
Affective Instability	–	–	.023**	-.028	–	.137	-.030*	-.106	.076
Efforts to Avoid Abandonment	–	–	.045**	.042	–	.060	-.016*	-.009	.051
Anger	.680	–	–	–	–	–	–	–	–
Identity Disturbance	.688	–	–	–	–	–	–	–	–
Emptiness	–	–	.022*	-.027	–	-.611	-.003	.633	-.363
Paranoid Ideation	–	–	.039**	.021	–	-.560	-.001	.600	-.272
Suicidality	–	–	.110**	.145	–	-.257	.005	.362	-.029
Impulsivity	.593	–	–	–	–	–	–	–	–

Note. For each item flagged for DIF, nonuniform DIF was tested using a 1*df* test in which the slope was constrained to be equal. If results suggested the presence of nonuniform DIF for a given item, then that item was not tested for uniform DIF. If nonuniform DIF was not detected, then a separate 1*df* test was conducted, in which the threshold was constrained to be equal. If the resulting *b* value was significant ($p <$ .01), then the item was deemed to exhibit uniform DIF. Magnitude of DIF among flagged items is evaluated by the ESSD statistic, with .2, .4, and .6 reflecting small, medium, and large effect sizes, respectively. The effect of DIF is only considered practically meaningful at the item-level if the ESSD value is ≥ .2.

b = standardized regression coefficient from item to latent factor; ω = standardized regression coefficient from item to latent trait-by-group interaction term; ESSD = Expected score standardized difference; PMS = Perceived minority stress.

* $p < .01$

** $p < .001$

Table 22

Fit Indices for Stage 1 Sequential MIMIC Models Comparing Heterosexual and Gay Men (Supplemental)

Model	BL Std. Ld.	RMSEA	CFI	TLI	AIC	X^2	*df*	ΔRMSEA	ΔCFI	ΔTLI	ΔAIC	ΔX^2	Δ*df*
		Baseline Model											
Baseline		.052	.938	.925	3493.3	510.44	142	–	–	–	–	–	–
		Stage 1 Sequential Models											
Unstable Relationships	.552	.053	.938	.924	3492.7	503.39	140	.000	.000	.001	0.59	4.81	2
Affective Instability	.718	.053	.938	.924	3492.7	503.62	140	.000	.000	.001	0.56	3.08	2
Efforts to Avoid Abandonment	.632	.053	.938	.924	3487.7	503.52	140	.000	.000	.001	5.6	4.91	2
Anger	.666	.053	.938	.924	3484.4	503.30	140	.000	.000	.001	8.87	6.75	2
Identity Disturbance	.678	.053	.938	.924	3489.8	503.76	140	.000	.000	.001	3.49	3.39	2
Emptiness	.659	.053	.938	.924	3497.0	504.06	140	.000	.000	.001	-3.73	0.14	2
Paranoid Ideation	.683	.053	.938	.924	3497.1	504.47	140	.000	.000	.001	-3.81	0.07	2
Suicidality	.410	.053	.938	.924	3475.5	503.76	140	.000	.000	.001	17.84	5.19	2
Impulsivity	.610	.053	.938	.924	3487.9	503.21	140	.000	.000	.001	5.39	7.78	2

Note. Model fit indices for baseline MIMIC model and stage 1 sequential models. In baseline model, all item parameters were constrained to be equal across groups. In stage 1 sequential models, each item was released individually and model fit indices were compared to the baseline model. Each item resulting in a significant Likelihood Ratio Test upon release was flagged for potential DIF.
BL = Baseline; RMSEA = Root Square Mean Error of Approximation; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; AIC = Akaike Information Criterion; *df* = degrees of freedom.
* $p < .01$
** $p < .001$

Table 23

Fit Indices for Stage 2 Sequential MIMIC Models Comparing Heterosexual and Gay Men (Supplemental)

Model	RMSEA	CFI	TLI	AIC	X^2	*df*	ΔRMSEA	ΔCFI	ΔTLI	ΔAIC	ΔX^2	Δ*df*
Baseline Model												
Baseline	.055	.938	.916	3493.3	510.44	126	–	–	–	–	–	–
Stage 2 Sequential Models												
Unstable Relationships	.055	.938	.917	3492.7	509.85	128	.000	.000	.001	0.59	4.55	2
Affective Instability	*Anchor item*											
Efforts to Avoid Abandonment	.055	.938	.917	3487.7	504.84	128	.000	.000	.001	5.6	9.54*	2
Anger	.055	.938	.917	3484.4	501.57	128	.000	.000	.001	8.87	12.82*	2
Identity Disturbance	.055	.938	.917	3489.8	506.95	128	.000	.000	.001	3.49	7.44	2
Emptiness	.055	.938	.917	3497.0	514.17	128	.000	.000	.001	-3.73	0.27	2
Paranoid Ideation	.055	.938	.917	3497.1	514.25	128	.000	.000	.001	-3.81	0.19	2
Suicidality	.055	.938	.917	3475.5	492.60	128	.000	.000	.001	17.84	21.65**	2
Impulsivity	.055	.938	.917	3487.9	505.05	128	.000	.000	.001	5.39	9.38*	2

Note. Model fit indices for baseline MIMIC model and stage 2 sequential models. Anchor item response was constrained to be equal across groups and all other items varied freely. Then, each non-anchor item was iteratively tested for DIF using a 2*df* test in which both the slope and threshold were constrained to be equal across groups. Model fit indices were then compared against the fit of the baseline model. Each item that resulted in significantly worse fit ($p < .01$) according to Likelihood Ratio Tests after being constrained was flagged for DIF.
RMSEA = Root Square Mean Error of Approximation; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; AIC = Akaike Information Criterion; *df* = degrees of freedom.
* $p < .01$
** $p < .001$

Table 24

DIF and ESSD Results for MIMIC Models Comparing Heterosexual and Gay Men (Supplemental)

Criterion	Factor Std. Ld.	Uniform *b*	Non-Uniform ω	ESSD	Uniform *b*	Non-Uniform ω	PMS *b*	PMS*BPD *b*	ESSD
	Final Model				*Perceived Minority Stress Model*				
Unstable Relationships	.552	–	–	–	–	–	–	–	–
Affective Instability	.718	–	–	–	–	–	–	–	–
Efforts to Avoid Abandonment	–	.001	.020	.004	.017	.022	-.021	-.001	.131
Anger	–	-.021*	-.015	-.132	-.005	–	-.013	-.007	1.56
Identity Disturbance	.678	–	–	–	–	–	–	–	–
Emptiness	.659	–	–	–	–	–	–	–	–
Paranoid Ideation	.683	–	–	–	–	–	–	–	–
Suicidality	–	.017*	.035	.067	.009	–	-.001	.035	-.166
Impulsivity	–	-.003	-.020	-.021	.041	.128	-.039	-.152	-1.78

Note. For each item flagged for DIF, nonuniform DIF was tested using a 1*df* test in which the slope was constrained to be equal. If results suggested the presence of nonuniform DIF for a given item, then that item was not tested for uniform DIF. If nonuniform DIF was not detected, then a separate 1*df* test was conducted, in which the threshold was constrained to be equal. If the resulting *b* value was significant ($p <$ .01), then the item was deemed to exhibit uniform DIF. Magnitude of DIF among flagged items is evaluated by the ESSD statistic, with .2, .4, and .6 reflecting small, medium, and large effect sizes, respectively. The effect of DIF is only considered practically meaningful at the item-level if the ESSD value is ≥ .2.

b = standardized regression coefficient from item to latent factor; ω = standardized regression coefficient from item to latent trait-by-group interaction term; ESSD = Expected score standardized difference; PMS = Perceived minority stress.

* $p < .01$

** $p < .001$

Table 25

Fit Indices for Stage 1 Sequential MIMIC Models Comparing Heterosexual and Bisexual Men (Supplemental)

Model	BL Std. Ld.	RMSEA	CFI	TLI	AIC	X^2	*df*	ΔRMSEA	ΔCFI	ΔTLI	ΔAIC	ΔX^2	Δ*df*
		Baseline Model											
Baseline		.066	.911	.893	-89911.8	415.36	142	–	–	–	–	–	–
		Stage 1 Sequential Models											
Unstable Relationships	.551	.066	.911	.892	-89908.8	409.65	140	.000	.000	.001	-3.02	0.93	2
Affective Instability	.717	.066	.911	.892	-89908.7	409.80	140	.000	.000	.001	-3.16	0.43	2
Efforts to Avoid Abandonment	.630	.066	.911	.892	-89909.0	409.92	140	.000	.000	.001	-2.79	0.45	2
Anger	.666	.066	.911	.892	-89911.6	409.74	140	.000	.000	.001	-.21	1.82	2
Identity Disturbance	.678	.066	.911	.892	-89913.8	409.72	140	.000	.000	.001	1.99	2.61	2
Emptiness	.660	.066	.911	.892	-89908.3	409.76	140	.000	.000	.001	-3.53	0.29	2
Paranoid Ideation	.685	.066	.911	.892	-89911.4	409.88	140	.000	.000	.001	-.38	1.26	2
Suicidality	.415	.066	.912	.892	-89950.8	409.20	140	.000	.001	.001	39.03	7.82	2
Impulsivity	.611	.066	.911	.892	-89908.8	409.74	140	.000	.000	.001	-3.04	0.60	2

Note. Model fit indices for baseline MIMIC model and stage 1 sequential models. In baseline model, all item parameters were constrained to be equal across groups. In stage 1 sequential models, each item was released individually and model fit indices were compared to the baseline model. Each item resulting in a significant Likelihood Ratio Test upon release was flagged for potential DIF.
BL = Baseline; RMSEA = Root Square Mean Error of Approximation; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; AIC = Akaike Information Criterion; *df* = degrees of freedom.
* $p < .01$
** $p < .001$

Table 26

Fit Indices for Stage 2 Sequential MIMIC Models Comparing Heterosexual and Bisexual Men (Supplemental)

Model	RMSEA	CFI	TLI	AIC	X^2	*df*	ΔRMSEA	ΔCFI	ΔTLI	ΔAIC	ΔX^2	Δ*df*
Baseline Model												
Baseline	.070	.912	.880	-89911.8	415.36	126	–	–	–	–	–	–
Stage 2 Sequential Models												
Unstable Relationships	.069	.912	.882	-89908.8	414.39	128	.001	.000	.002	-3.02	0.97	2
Affective Instability	*Anchor item*											
Efforts to Avoid Abandonment	.069	.912	.882	-89909.0	414.16	128	.001	.000	.002	-2.79	1.20	2
Anger	.069	.912	.882	-89911.6	412.66	128	.001	.000	.002	-0.21	2.70	2
Identity Disturbance	.069	.911	.882	-89913.8	409.42	128	.001	.001	.002	1.99	5.94	2
Emptiness	.069	.912	.882	-89908.3	414.89	128	.001	.000	.002	-3.53	0.47	2
Paranoid Ideation	.069	.912	.882	-89911.4	411.77	128	.001	.000	.002	-0.38	3.59	2
Suicidality	.069	.911	.881	-89950.8	372.64	128	.001	.001	.001	39.03	42.72**	2
Impulsivity	.069	.912	.882	-89908.8	414.40	128	.001	.000	.002	-3.04	0.96	2

Note. Model fit indices for baseline MIMIC model and stage 2 sequential models. Anchor item response was constrained to be equal across groups and all other items varied freely. Then, each non-anchor item was iteratively tested for DIF using a 2*df* test in which both the slope and threshold were constrained to be equal across groups. Model fit indices were then compared against the fit of the baseline model. Each item that resulted in significantly worse fit ($p < .01$) according to Likelihood Ratio Tests after being constrained was flagged for DIF.
RMSEA = Root Square Mean Error of Approximation; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; AIC = Akaike Information Criterion; *df* = degrees of freedom.
* $p < .01$
** $p < .001$

Table 27

DIF and ESSD Results for MIMIC Models Comparing Heterosexual and Bisexual Men (Supplemental)

Criterion	Factor Std. Ld.	Uniform *b*	Non-Uniform ω	ESSD	Uniform *b*	Non-Uniform ω	PMS *b*	PMS*BPD *b*	ESSD
	Final Model				*Perceived Minority Stress Model*				
Unstable Relationships	.551	–	–	–	–	–	–	–	–
Affective Instability	.717	–	–	–	–	–	–	–	–
Efforts to Avoid Abandonment	.630	–	–	–	–	–	–	–	–
Anger	.666	–	–	–	–	–	–	–	–
Identity Disturbance	.678	–	–	–	–	–	–	–	–
Emptiness	.660	–	–	–	–	–	–	–	–
Paranoid Ideation	.685	–	–	–	–	–	–	–	–
Suicidality	–	.020	.052	.106	.017	-.011	-.019	.071	.099
Impulsivity	.611	–	–	–	–	–	–	–	–

Note. For each item flagged for DIF, nonuniform DIF was tested using a 1*df* test in which the slope was constrained to be equal. If results suggested the presence of nonuniform DIF for a given item, then that item was not tested for uniform DIF. If nonuniform DIF was not detected, then a separate 1*df* test was conducted, in which the threshold was constrained to be equal. If the resulting *b* value was significant ($p < .01$), then the item was deemed to exhibit uniform DIF. Magnitude of DIF among flagged items is evaluated by the ESSD statistic, with .2, .4, and .6 reflecting small, medium, and large effect sizes, respectively. The effect of DIF is only considered practically meaningful at the item-level if the ESSD value is ≥ .2.

b = standardized regression coefficient from item to latent factor; ω = standardized regression coefficient from item to latent trait-by-group interaction term; ESSD = Expected score standardized difference; PMS = Perceived minority stress.

* $p < .01$

** $p < .001$

Table 28

Fit Indices for Stage 1 Sequential MIMIC Models Comparing Heterosexual and Lesbian Women (Supplemental)

Model	BL Std. Ld.	RMSEA	CFI	TLI	AIC	X^2	*df*	ΔRMSEA	ΔCFI	ΔTLI	ΔAIC	ΔX^2	Δ*df*
		Baseline Model											
Baseline		.049	.952	.943	-85057.2	371.73	142	–	–	–	–	–	–
		Stage 1 Sequential Models											
Unstable Relationships	.574	.050	.952	.942	-85057.3	366.53	140	.001	.000	.001	0.09	3.99	2
Affective Instability	.732	.050	.952	.942	-85058.7	366.53	140	.001	.000	.001	1.49	4.40	2
Efforts to Avoid Abandonment	.667	.050	.952	.942	-85057.2	366.67	140	.001	.000	.001	-0.05	2.23	2
Anger	.685	.050	.952	.942	-85057.2	366.61	140	.001	.000	.001	-0.03	2.74	2
Identity Disturbance	.690	.050	.952	.942	-85056.3	366.74	140	.001	.000	.001	-0.87	1.52	2
Emptiness	.689	.050	.952	.942	-85067.7	366.36	140	.001	.000	.001	10.47	7.30	2
Paranoid Ideation	.689	.050	.952	.942	-85057.6	366.72	140	.001	.000	.001	0.38	2.04	2
Suicidality	.443	.050	.953	.942	-85094.64	365.67	140	.001	.001	.001	37.444	13.15*	2
Impulsivity	.605	.050	.952	.942	-85069.9	366.19	140	.001	.000	.001	12.67	11.52	2

Note. Model fit indices for baseline MIMIC model and stage 1 sequential models. In baseline model, all item parameters were constrained to be equal across groups. In stage 1 sequential models, each item was released individually and model fit indices were compared to the baseline model. Each item resulting in a significant Likelihood Ratio Test upon release was flagged for potential DIF.
BL = Baseline; RMSEA = Root Square Mean Error of Approximation; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; AIC = Akaike Information Criterion; *df* = degrees of freedom.
* $p < .01$
** $p < .001$

Table 29

Fit Indices for Stage 2 Sequential MIMIC Models Comparing Heterosexual and Lesbian Women (Supplemental)

Model	RMSEA	CFI	TLI	AIC	X^2	*df*	ΔRMSEA	ΔCFI	ΔTLI	ΔAIC	ΔX^2	Δ*df*
Baseline Model												
Baseline	.052	.953	.936	-85057.2	371.73	126	–	–	–	–	–	–
Stage 2 Sequential Models												
Unstable Relationships	.052	.953	.937	-85057.3	367.64	128	.000	.000	.001	0.09	4.09	2
Affective Instability	*Anchor item*											
Efforts to Avoid Abandonment	.052	.953	.937	-85057.2	367.84	128	.000	.000	.001	-0.05	3.89	2
Anger	.052	.953	.937	-85057.2	367.80	128	.000	.000	.001	-0.03	3.93	2
Identity Disturbance	.052	.953	.937	-85056.3	368.62	128	.000	.000	.001	-0.87	3.11	2
Emptiness	.052	.953	.937	-85067.7	357.48	128	.000	.000	.001	10.47	14.25*	2
Paranoid Ideation	.052	.953	.937	-85057.6	367.38	128	.000	.000	.001	0.38	4.35	2
Suicidality	.052	.953	.937	-85094.64	330.59	128	.000	.000	.001	37.44	41.14**	2
Impulsivity	.052	.953	.937	-85069.9	355.76	128	.000	.000	.001	12.67	15.97**	2

Note. Model fit indices for baseline MIMIC model and stage 2 sequential models. Anchor item response was constrained to be equal across groups and all other items varied freely. Then, each non-anchor item was iteratively tested for DIF using a 2*df* test in which both the slope and threshold were constrained to be equal across groups. Model fit indices were then compared against the fit of the baseline model. Each item that resulted in significantly worse fit ($p < .01$) according to Likelihood Ratio Tests after being constrained was flagged for DIF.

RMSEA = Root Square Mean Error of Approximation; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; AIC = Akaike Information Criterion; *df* = degrees of freedom.

* $p < .01$

** $p < .001$

Table 30

DIF and ESSD Results for MIMIC Models Comparing Heterosexual and Lesbian Women (Supplemental)

Criterion	Factor Std. Ld.	Uniform *b*	Non-Uniform ω	ESSD	Uniform *b*	Non-Uniform ω	PMS *b*	PMS*BPD *b*	ESSD
	Final Model				*Perceived Minority Stress Model*				
Unstable Relationships	.574	–	–	–	–	–	–	–	–
Affective Instability	.732	–	–	–	–	–	–	–	–
Efforts to Avoid Abandonment	.667	–	–	–	–	–	–	–	–
Anger	.685	–	–	–	–	–	–	–	–
Identity Disturbance	.690	–	–	–	–	–	–	–	–
Emptiness	–	-.017	-.007	-.139	-.019	.001	.002	-.002	-.153
Paranoid Ideation	.689	–	–	–	–	–	–	–	–
Suicidality	–	.016*	.044	.094	.009	–	-.004	.046	.045
Impulsivity	–	.020**	.020	.135	.018	–	-.002	.014	.121

Note. For each item flagged for DIF, nonuniform DIF was tested using a 1*df* test in which the slope was constrained to be equal. If results suggested the presence of nonuniform DIF for a given item, then that item was not tested for uniform DIF. If nonuniform DIF was not detected, then a separate 1*df* test was conducted, in which the threshold was constrained to be equal. If the resulting *b* value was significant ($p <$ .01), then the item was deemed to exhibit uniform DIF. Magnitude of DIF among flagged items is evaluated by the ESSD statistic, with .2, .4, and .6 reflecting small, medium, and large effect sizes, respectively. The effect of DIF is only considered practically meaningful at the item-level if the ESSD value is ≥ .2.

b = standardized regression coefficient from item to latent factor; ω = standardized regression coefficient from item to latent trait-by-group interaction term; ESSD = Expected score standardized difference; PMS = Perceived minority stress.

* $p < .01$

** $p < .001$

Table 31

Fit Indices for Stage 1 Sequential MIMIC Models Comparing Heterosexual and Bisexual Women (Supplemental)

Model	BL Std. Ld.	RMSEA	CFI	TLI	AIC	X^2	*df*	ΔRMSEA	ΔCFI	ΔTLI	ΔAIC	ΔX^2	Δ*df*
		Baseline Model											
Baseline		.044	.969	.962	80157.8	503.97	142	–	–	–	–	–	–
		Stage 1 Sequential Models											
Unstable Relationships	.578	.044	.969	.962	80148.2	496.43	140	.000	.000	.000	9.61	13.93**	2
Affective Instability	.738	.044	.969	.962	80156.5	497.15	140	.000	.000	.000	1.24	3.93	2
Efforts to Avoid Abandonment	.672	.044	.969	.962	80142.4	496.62	140	.000	.000	.000	15.39	8.87	2
Anger	.690	.044	.969	.962	80154.4	497.22	140	.000	.000	.000	3.38	4.11	2
Identity Disturbance	.693	.044	.969	.962	80153.6	497.51	140	.000	.000	.000	4.15	3.25	2
Emptiness	.691	.044	.969	.962	80152.6	497.03	140	.000	.000	.000	5.22	5.63	2
Paranoid Ideation	.690	.044	.969	.962	80158.4	497.67	140	.000	.000	.000	-0.66	1.53	2
Suicidality	.451	.044	.970	.963	80003.5	488.12	140	.000	.001	.001	154.3	44.32**	2
Impulsivity	.613	.044	.969	.962	80119.6	494.92	140	.000	.000	.000	38.21	23.82**	2

Note. Model fit indices for baseline MIMIC model and stage 1 sequential models. In baseline model, all item parameters were constrained to be equal across groups. In stage 1 sequential models, each item was released individually and model fit indices were compared to the baseline model. Each item resulting in a significant Likelihood Ratio Test upon release was flagged for potential DIF.
BL = Baseline; RMSEA = Root Square Mean Error of Approximation; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; AIC = Akaike Information Criterion; *df* = degrees of freedom.
* $p < .01$
** $p < .001$

Table 32

Fit Indices for Stage 2 Sequential MIMIC Models Comparing Heterosexual and Bisexual Women (Supplemental)

Model	RMSEA	CFI	TLI	AIC	X^2	*df*	ΔRMSEA	ΔCFI	ΔTLI	ΔAIC	ΔX^2	Δ*df*
Baseline Model												
Baseline	.045	.971	.960	80157.8	503.97	126	–	–	–	–	–	–
Stage 2 Sequential Models												
Unstable Relationships	.045	.971	.961	80148.2	499.88	128	.000	.000	.001	9.61	13.23*	2
Affective Instability	*Anchor item*											
Efforts to Avoid Abandonment	.045	.970	.960	80142.4	500.08	128	.000	.001	.000	15.39	18.93**	2
Anger	.045	.970	.960	80154.4	500.04	128	.000	.001	.000	3.38	7.24	2
Identity Disturbance	.045	.970	.960	80153.6	500.86	128	.000	.001	.000	4.15	8.01	2
Emptiness	.045	.970	.961	80152.6	489.72	128	.000	.001	.001	5.22	9.15	2
Paranoid Ideation	.045	.970	.961	80158.4	499.62	128	.000	.001	.001	-0.66	3.36	2
Suicidality	.046	.969	.959	80003.5	462.83	128	.001	.002	.001	154.3	153.62**	2
Impulsivity	.045	.970	.960	80119.6	488.00	128	.000	.001	.000	38.21	40.92**	2

Note. Model fit indices for baseline MIMIC model and stage 2 sequential models. Anchor item response was constrained to be equal across groups and all other items varied freely. Then, each non-anchor item was iteratively tested for DIF using a 2*df* test in which both the slope and threshold were constrained to be equal across groups. Model fit indices were then compared against the fit of the baseline model. Each item that resulted in significantly worse fit ($p < .01$) according to Likelihood Ratio Tests after being constrained was flagged for DIF.
RMSEA = Root Square Mean Error of Approximation; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; AIC = Akaike Information Criterion; *df* = degrees of freedom.
* $p < .01$
** $p < .001$

Table 33

DIF and ESSD Results for MIMIC Models Comparing Heterosexual and Bisexual Women (Supplemental)

Criterion	Factor Std. Ld.	Uniform *b*	Non-Uniform ω	ESSD	Uniform *b*	Non-Uniform ω	PMS *b*	PMS* BPD *b*	ESSD
	Final Model				*Perceived Minority Stress Model*				
Unstable Relationships	–	-.010	-.012	-.065	-.012	-.286	.020	.268	-1.338
Affective Instability	.738	–	–	–	–	–	–	–	–
Efforts to Avoid Abandonment	–	–	.033*	.008	–	.391	-.018	-.344	.963
Anger	.690	–	–	–	–	–	–	–	–
Identity Disturbance	.693	–	–	–	–	–	–	–	–
Emptiness	.691	–	–	–	–	–	–	–	–
Paranoid Ideation	.690	–	–	–	–	–	–	–	–
Suicidality	–	–	.091**	.198	–	-.017	.005	.102	-.079
Impulsivity	–	–	.047**	.164	–	.060	.008	-.023	.237

Note. For each item flagged for DIF, nonuniform DIF was tested using a 1*df* test in which the slope was constrained to be equal. If results suggested the presence of nonuniform DIF for a given item, then that item was not tested for uniform DIF. If nonuniform DIF was not detected, then a separate 1*df* test was conducted, in which the threshold was constrained to be equal. If the resulting *b* value was significant ($p < .01$), then the item was deemed to exhibit uniform DIF. Magnitude of DIF among flagged items is evaluated by the ESSD statistic, with .2, .4, and .6 reflecting small, medium, and large effect sizes, respectively. The effect of DIF is only considered practically meaningful at the item-level if the ESSD value is ≥ .2.

b = standardized regression coefficient from item to latent factor; ω = standardized regression coefficient from item to latent trait-by-group interaction term; ESSD = Expected score standardized difference; PMS = Perceived minority stress.

* $p < .01$

** $p < .001$

Appendix B: IRB Exemption Letter

NOT HUMAN SUBJECTS RESEARCH DETERMINATION

June 1, 2020

Erin Carsten

Dear Ms.Carsten:

On 5/30/2020, the IRB reviewed the following protocol:

IRB ID:	STUDY001002
Title:	Sexual Orientation: Examining Latent Structure and Measurement Invariance of Personality Disorder Criteria Across Subgroups

The IRB determined that the proposed activity does not constitute research involving human subjects as defined by DHHS and FDA regulations.

IRB review and approval is not required. This determination applies only to the activities described in the IRB submission. If changes are made and there are questions about whether these activities constitute human subjects research, please submit a new application to the IRB for a determination.

While not requiring IRB approval and oversight, your project activities should be conducted in a manner that is consistent with the ethical principles of your profession. If this project is program evaluation or quality improvement, do not refer to the project as research and do not include the assigned IRB ID or IRB contact information in the consent document or any resulting publications or presentations.

Sincerely,

Various Menzel
IRB Research Compliance Administrator

Institutional Review Boards / Research Integrity & Compliance
FWA No. 00001669
University of South Florida / 3702 Spectrum Blvd., Suite 165 / Tampa, FL 33612 / 813-974-5638