University of Wisconsin Milwaukee

UWM Digital Commons

August 2023

Masculinity-The Next Generation: Millennials, the MRNI, and Measurement Invariance

Charlotte Maria Shanaver
University of Wisconsin-Milwaukee

Follow this and additional works at: https://dc.uwm.edu/etd

Part of the Cognitive Psychology Commons

Recommended Citation

MASCULINITY-THE NEXT GENERATION: MILLENNIALS, THE MRNI, AND

MEASUREMENT INVARIANCE

by

Charlotte Maria Shanaver

A Dissertation Submitted in

Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

in Educational Psychology

at

The University of Wisconsin-Milwaukee

August 2023

ABSTRACT

MASCULINITY-THE NEXT GENERATION: MILLENNIALS, THE MRNI, AND
MEASUREMENT INVARIANCE
by

Charlotte Maria Shanaver

The University of Wisconsin-Milwaukee
Under the Supervision of Professor Stephen R. Wester

Socialization is, "the process of learning to behave in a way that is acceptable to society" (Merriam-Webster, n.d.). It is inescapable and pervasive. One of the most socialized constructs is gender. Gender roles, norms, expectations, shape how male and female identified folx move through the world and expect others to move through the world. However, this can become problematic when such norms become rigid and insurmountable. Adherence to traditional or hegemonic masculine norms, has been correlated with worse mental health outcomes (Gerdes & Levant, 2018). For decades, psychologists have utilized the Male Role Norms Inventory (Levant et al., 1992; Levant et al., 2007) to study the impact that such norms have on male identified individuals. However, this inventory has not been normed or explored using younger generations of men. Due to recent social and political shifts, it is likely that an inventory normed on older populations, may not resonate with newer generations (Yeazel, 2015). Measurement invariance is statistical evidence that a measure is assessing the same construct between groups (Lee, 2018). For a measure to be truly valid, it must demonstrate strong measurement invariance. The current study explores measurement invariance on the Male Role Norms Inventory-Revised (Levant et al., 2007) between younger and older generations of men to assess whether the construct of masculinity, originally developed in the 90s, still holds true for male identified folx today.

For those who sang the words that carried me to the finish line…

"the sun's gonna rise in a mile…"
"let it be"
"we've all got bruises"
"but singing works just fine for me"
"from what I know, you're always here to stay"
"I can say that lovin' you is easy"
"see there's a reason you're still standing after all you've seen"
"so don't you ever let go"
"I can't remember all the times I tried to tell myself to hold on…to these moments as they pass"
"it's been so long since I've see the ocean…I guess I should"
"but I'm not giving in an inch to fear"
"and when I go there, I go there with you"
"keep on rockin' in the free world"
"yeah"

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I didn't truly understand the phrase, "it takes a village" until now. Looking back on all those who supported me through this journey, I cannot help but feel honored, humbled, and unconditionally grateful. A doctorate brings out your best and your worst, mostly your worst, and I can guarantee that there were moments, days, weeks, years…that I was less than a pleasure to be around. So, to everyone who is still standing, thank you.

Amidst this village, however, are a few folx who deserve more than a mass platitude. Although individual recognition doesn't begin to account for the service provided, nor are there words that can fully express my appreciation, I am going to try.

To my advisor, Dr. Stephen Wester, I could not have crossed the finish line without you. Your unwavering dedication to my success kept me moving even when all I wanted to do was sit still. Few students get to work with someone who is not only invested in their work, but who is invested in them as a person. I am beyond lucky to have been mentored by someone who thinks similarly, and who consistently reminded me, "never forget where you came from".

To my committee, I learned more than I ever imagined I would from your guiding hands. Dr. Dupont, thank you for always being the encouraging, yet logical voice in my ear. You reminded me that there is strength in failure and that failing is progress, not regression. Dr. Zhang, thank you for challenging me to be better, to keep learning, and to never accept second best. Your tenacity parallels mine and allowed me to create something I can take authentic pride in. Dr. Fouad, thank you for your warm voice of reason and fierce defense of my progress. You showed me what it means to move with confidence and humility, in a world that views these traits as mutually exclusive.

To my parents, thank you for holding on, and holding my hand, even when I didn't know what in the world I was saying or doing, you are the real MVPs.

And lastly, to the folx on my internship. You met me at a time of great strife, great pain, and great vulnerability. You met me scared and raw. And I was met with no judgement. No expectations of being perfectly put together. The amount of healing and unlearning this space provided changed my life in ways I cannot put into words. So, to Heather and Jason, thank you for your friendship, allyship, and acceptance. We truly are a cohort for the books. To Dr. Holly Davis, thank you for being human, and letting me do the same. Thank you for supporting and joining my feminist rants and for vehemently defending my right to take up space and be heard. To Dr. Marie Adams, thank you for providing ego strength when I had none left to give. For holding my stories and fears without reprimand. And for sharing my sarcastic, cynical, sense of humor. And to Dr. Pat Galligan, thank you for showing me what genuine safety looks like. For teaching me that gentleness does not mean weakness. That vulnerability is a sign of strength. That I don't have to walk into every room swinging, nor must I always lead with fear.

To each of you individually and as my community, thank you.

Masculinity-the Next Generation: Millennials, the MRNI, and Measurement Invariance

Chapter 1

INTRODUCTION

Socialization is, "the process of learning to behave in a way that is acceptable to society" (Merriam-Webster, n.d.). From a social constructivist approach, learning, or construction, is influenced by things such as historical and political circumstances (Brooks & Elder, 2014). As society changes, so do socialized expectations and constructs, such as gender roles. Depending on the situation, gender roles may be modified to fit specific situations or needs founded on previously experienced or expected consequences (Brooks & Elder, 2014). For example, masculinity, as social construct, allows for contextual variables to be considered and promotes an understanding of both the individual man and the accepted "whole" of what is masculine. Such gender prescriptions are one area where socialization is ever evolving (Ruspini, 2019). Gender roles have been the focus of social and psychological research for years, due to their impact on individuals' actions, interactions, and well-being. They provide a map for how men and women should think, act, look, and feel; how they should be in the world (O'Neil, 1986). Previously accepted gender constructs assigned not only specific behaviors, emotions, and actions to the sexes, they defined what jobs were appropriate, what attire was acceptable, and even what hobbies one could have (Levant et al., 1992; Mahalik et al., 2003, O'Neil, 1086).

According to these prescriptions, men were to be unemotional, stoic, aggressive, and independent. They were to provide, be in control, and unwavering in the face of strife. Traditional Western masculinity has come to be characterized by traits such as self-reliance, assertiveness, and emotional control, (Addis & Hoffman, 2020, Levant et al., 2007) and includes expectations that men should provide, men should be strong, tough, and inexpressive (Levant et

al., 1992). This traditional stereotype has come to be known as hegemonic (Connell, 1995), an enactment of masculinity shaped against femininity, held in high esteem above others (Addis & Hoffman, 2020) and praised by social structures (Brooks & Elder, 2014).

However, younger generations of men seem to be redefining what is hegemonic, in that what was once socially acclaimed and accepted, is now being challenged and questioned (Yeazel, 2015). For example, research shows that younger generations engage in more egalitarian relationships, with more equal divisions of responsibility (Yeazel, 2015). They embrace more open-minded belief systems, including acceptance of non-traditional parenting styles, same sex relationships, and fluid definitions of sexuality, further differentiating them from others (Rimmer, 2015; Yeazel, 2015). In fact, a recent poll showed that millennials were 27 percent more likely to support gay marriage than Gen X's and 77 percent more likely than the Baby Boomer population (Accel + Qualtrics, 2016). Additionally, as millennial gender roles have evolved, expectations have changed between men and women. For example, one poll of millennial men indicated that 75 percent of millennial fathers considered themselves the primary caregiver to their children and 94 percent felt they needed "more opportunities to feel safe discussing emotions" (Advisor Magazine, 2018).

Such assertions directly contradict previously defined traditional masculinity which assumes men will have little paternal engagement (Levant & Richmond, 2007). Millennials are redefining complicity as hegemony. An additional example is male attire. Where in previous decades wearing fitted pants or loud colors would be perceived as solely feminine, today's ads portray men in bold prints, tight jeans, and even the oh so taboo color pink. One of the more rugged brands, AXE cologne recently launched an ad titled, "Find Your Magic", and went so far as to display a man in traditionally woman's heels using a prominent drag queen as their

spokesperson (Wertz, 2020). Another media demonstration of a distinct millennial masculinity is through recent cinema. Take the movie 21 Jump Street, starring Channing Tatum as the hardened heart throb and Jonah Hill as the quirky side kick. The two undercover cops are sent to a high school to bust a notorious drug ring where traditionally masculine Channing Tatum expects to win the school over with his physical prowess and athletic ability, yet he is shocked to find that the new generation of students (millennial and gen z) no longer praise such attributes. Instead, his partner Jonah Hill, soars to popularity for his progressive beliefs and academic achievement. What once was dominant is now compliant. Such examples of campaigning and entertainment, specifically geared towards younger generations, demonstrates how far gender norms have expanded and how distinctive societal values are between generational groups.

The concept of generational differences has been strongly debated, specifically the existence of differences between current younger and older populations (Pew, 2018). Some research questions the legitimacy of a generational paradigm (see Costanza, 2018 or Zacher, 2015) suggesting lacking theoretical evidence, but in fact, when viewing masculinity through a social constructivist approach, the concept of generations fits well. Brooks and Elder (2014) citing Felmange (2000) note, "Social construction emphasizes the role of socialization but also how social context, including historic, economic, political, linguistic, interpersonal, and psychological constructs, affect masculine identity" (p. 12). This theory encompasses more than historical events, as suggested by Zacher (2015), but rather the climate created by the layers of context, distinct to varying age groups, that impact the development of one's own masculine self. The construct of generation is not perfect however, it is the most socially accepted and understood way of categorizing and documenting variation among age groups, which critics even

highlight the importance of (Zacher, 2015), making a generational construct fitting for the purpose studying masculinity.

Recent literature (e.g. Addis & Hoffman, 2020) has highlighted the importance of understanding the contextual underpinnings of masculinity. Context is defined as the circumstances that form the setting for an event, statement, or idea, in terms of which it can be fully understood and assessed (Merriam-Webster, n.d.). Although this term encompasses different environments, situations, and cultures, one could argue that it does not quite encompass a generational narrative. Generation is more than a context; it is a collective lived experience encompassing multiple cultural norms (Merriam-Webster, n.d.). The millennial generation collectively differs from their predecessors (Ruspini, 2019). Though they were raised by baby boomers and Gen X's, their philosophies are remarkably dissimilar, posing the question, if not parental or mentor influence, what has created such a marked shift? What has influenced their new constructional definition of masculinity? This group has grown up in a distinct era, through technological advancement, societal change, and political evolution. Not only are they the first to grow up with computers in hand and constant access to unfiltered opinions, the social and political atmosphere experienced a discernable shift surrounding topics of race, gender, and sex (Accel + Qualtrics, 2016). Perhaps a better term to encompass the millennial experience is zeitgeist, or a "defining spirit or mood of a particular period of history as shown by the ideas and beliefs of the time" (Meriam-Webster, n.d.). Compared to their parents, the setting of millennial masculinity is new and confusing and his been broadly ignored (Yeazel, 2015). Although the study of men is not inherently new, the study of men as gendered beings with specific needs is novel and needs further examination (Brooks & Elder, 2014; Yeazel, 2015). This is especially true for millennials, who have been overlooked by the practicing community (Yeazel, 2015).

Multiple authors (e.g. Addis & Hoffman, 2020; Cuthbert, 2015; Gerdes et al., 2018) note the importance of demographic study, such as age, race, and gender, yet generation is absent from their call to action. It seems obvious that this zeitgeist has influenced what is expected or deemed appropriate of men and demands further exploration, especially as current masculinity measures are outdated.

Historically, theories of traditional masculinity led to the development of masculinity inventories (e.g. Levant et al., 1992; Mahalik et al., 2003; O'Neil, 1986). Although some aspects of masculinity may be visible (e.g. a man's muscular appearance or a public demonstration of cheering for a sports team) masculinity is not specifically observable, making it difficult to research. To address this problem, self-report inventories were created to assist (e.g. GRCS, O'Neil, 1986, CMNI, Mahalik et al, 2003, & MRNI, Levant et al., 1992). These measures attempt to quantify the masculine experience, however, measuring a nuanced construct proves difficult (Levant, Hall, Weingold, & McCurdy, 2015) and, statistically speaking, hypothesized factor structures do not always fit varying populations (Levant et al., 2015). For a measure to be truly generalizable, strong measurement invariance should be present. Looking more specifically at the MRNI, as it is the most psychometrically investigated measure to date (Gerdes et al., 2018) and its operationalization of traditional masculinity ideology (TMI) has been accepted by those researching men and masculinity, scholars have found issues with its generalizability (Levant et al., 2015). Developed in 1992, the original MRNI was a self-report measure consisting seven scales of 53 items rated on a 1-7 Likert type scale (Levant et al., 1992). It defines masculinity as a set of socialized attitudes, behaviors, and beliefs deemed acceptable for men and assesses to what extent the respondent endorses TMI. Although the scale was revised in 2007 and again slightly altered in 2010, the overall factor structure - and norms assessed have remained

predominantly the same (Levant et al., 1992; Levant et al., 2007; Levant et al., 2010), resulting in a definition of masculinity grounded in the late 1990s and posing the question, can a decades old measure truly encompass what masculinity espouses today?

Despite this outdated characterization, the MRNI is still the most widely used masculinity measure among researchers (Gerdes et al., 2018) to demonstrate how varying levels of TMI endorsement impact a range of male experiences. For example, higher levels of TMI have been correlated with increased alcohol and substance abuse (Gerdes & Levant, 2018), lower levels of paternal engagement (Young, 1996 as cited in Levant & Richmond, 2007), more negative attitudes towards physical and psychological help-seeking (Levant et al., 2013; Vogel & Wester, 2003; Vogel et al., 2005) and higher instances of anxiety, depression, and alexithymia (Levant et al., 2010). High usage of the MRNI within the research community has led to further examinations of its psychometric properties (see Levant et al., 2010 or Levant et al., 2015) and the development of shorter versions to decrease participant burn out and increase ease of use (MRNI-SF; MRNI-VB), however, examination of MRNI measurement invariance is severely lacking. As aforementioned, measurement invariance provides statistical support that a measure is assessing the same construct between groups (Lee, 2018). Without such evidence, true generalizability cannot be confirmed and without generalizability prior findings can be called into question. The Male Role Norms Inventory (MRNI; Levant et al., 1992) has also demonstrated variance (e.g. McDermott et al., 2017) with findings stating that "full metric invariance was supported only for heterosexual White men" (p. 735) suggesting that the MRNI is only truly valid for a very narrow population.

The MRNI was not only normed decades ago but was also normed on convenience samples of mostly White college men (Levant et al., 1992; Levant & McCurdy, 2018). Critics of

6

the MRNI have noted the limitation of such a sample and the lacking diversity (e.g. Levant & McCurdy, 2018) including the missing generational representation. Millennial and younger men were not included in this norming group, meaning that their demographic is absent in the accepted understanding of "masculinity". It is then reasonable to assume that the current masculinity measures do not resonate with these generations at all. The MRNI-R (Levant et al., 2007; Levant et al., 2010) measures seven different theorized "norms" of masculinity, however, as discussed in prior examples, millennial men do not fit neatly into these categories. Take avoidance of femininity for example, one item measuring this construct reads, "Men should not wear make-up, cover-up, or bronzer" (Levant et al., 2007, p. 96), yet in the aforementioned AXE commercial the spokesperson was a drag queen, outfit complete with a full face of make-up (Wertz, 2020). Another item, measuring fear and hatred of homosexuals reads, "Men should never hold hands or show affection towards one another" (Levant et al., 2007, p. 97), yet Rimmer (2015) directly states that homo-hysteria is not exhibited by millennials and multiple sources (i.e. Yeazel, 2017, Pew, 2018) note that millennials are openly supportive of same sex relationships. Such examples create a compelling argument that the masculinity construct being measured by the MRNI is not applicable to millennial and younger men.

The simplistic definition of traditional masculinity presented by the MRNI was previously adopted as fact, however, critics have recently called its legitimacy into question (Addis & Hoffman, 2020; Cuthbert, 2015). As a socialized construct, masculinity is influenced by factors including socioeconomic status, race, ethnicity, geographic location, sexual orientation, and more recently, a generational zeitgeist (Addis & Hoffman, 2020). Sticking with a social constructivist approach, masculinity is seen as functional, only if it fits the situation or environment (Brooks & Elder, 2014). It appears that the current environment denies traditional

masculinity's functionality, yet empirical study still relies on outdated tools. Unfortunately, existing masculinity measures are grounded within the time they were written (Gerdes et al., 2018) and although research has started to account for demographic differences by including more diverse samples (e.g. McDermott et al., 2017; Schwing & Wong, 2014), research has not explored longitudinal data nor generational cohort variance. However, if context forms the setting for an idea, an idea which cannot be explained without an understanding of context (Addis & Hoffman, 2020; Merriam- Webster, n.d.) then using an inventory normed on men outside of millennial zeitgeist to measure the millennial masculine experience is invalid and impractical.

Attempting to define masculinity for millennial men, using a construct deemed functional by Gen X or Boomer men, is likely not useful nor accurate and many millennial belief systems directly contradict traditional masculine norms. The MRNI specifically contains a subscale measuring one's fear or negative attitudes towards homosexuality (Levant et al., 1992; Levant et al., 2007; Levant et al., 2010), suggesting that to be masculine, one must think poorly or feel harshly towards gay men. However, such homophobia goes against the millennial identity (Rimmer, 2015). In fact, it has been argued that the homohysteria, or the fear of being perceived as homosexual, that is central to the traditional masculinity posed by the MRNI is not exhibited by millennials (Anderson, 2009). Displays of male to male affection have become commonplace not only in the athletic arena but are also less stigmatized in the general male population as well (Rimmer, 2015; Robinson, Anderson, & White, 2018). Such practices strengthen the argument that millennial men's construction and understanding of masculinity is distinct and different from TMI. As Rimmer so articulately states, "21st century masculinities are radically redefining the

notion of a 'traditional' masculinity that has dominated scholarly research for decades" (Rimmer, 2015, p. 1181) and the literature needs to catch up.

Current Study

The possibility of generational changes in masculinity pose the question, do current masculinity scales, specifically the MRNI, accurately capture masculinity? In their book on masculinity and context, Addis and Hoffman (2020) discuss the importance of "truth" in psychological research, in other words, an agreed upon meaning of a construct. Though a philosophical question, the idea of a true masculinity, whether it differs in context or not, is necessary to operationalize yet is one that researchers remain unsure of (Addis & Hoffman, 2020). It is common in psychological research for constructs to be deemed valid through avenues of favorable anecdotal and statistical evidence. In the context of the current study, and within the field of counseling psychology, a statistical demonstration of generational variance on the MRNI combined with the anecdotal evidence of incongruence between younger and older generations would be hard to argue against. Therefore, the main purpose of this study is to demonstrate the disparities between generational constructs of masculinity by exhibiting measurement variance between younger and older generations on both MRNI-R (Levant et al., 2007) total scores as well as scores on the seven subscales.

Measurement invariance analyses provide a statistical representation of how constructs measure between groups, which could in turn reveal how generational differences impact how men respond on masculinity inventories (Lee, 2018). Put another way, metric invariance allows us to determine whether or not the construct is consistent across the two groups. This is more than mere difference in obtained scores; the differences between generations will appear as variance in responses rather than differences in subscale means. Configural invariance will be

9

assessed first, as it is the most lenient test (Lee, 2018), with no limitations imposed meaning that neither the factor loadings, nor the y-intercepts have been constrained to equal (Lee, 208; Li et al., 2021; Crowson, 2020). If significant invariance holds, metric invariance will be analyzed, restricting the factor loadings of both groups to be equal (Lee, 2018; Crowson, 2020). This is testing the more constrained model, against the less constrained model, to see if when the factor loadings, or slopes, are the same across groups (Pirralha, 2020). If significant invariance is still present, meaning that the model still fits both groups despite the constraints, scalar variance will be assessed, restricting the intercepts of each group to be equal (Crowson, 2020; Lee, 2018; Pirralha, 2020). Should invariance still hold at this point, meaning that when the more constrained models are compared against the less constrained models, significant variance in loadings and intercepts is not present (Crowson, 2020; Lee, 2020; Li et al., 2021), measurement invariance can be confirmed. This means that the construct captured by the MRNI-R holds across the generations. This study will use independent multi-group confirmatory factor analyses (Li et al., 2021; Crowson, 2020), first on the MRNI-R total score, and next on each of the MRNI-R subscales.

Chapter 2

LITERATURE REVIEW

The millennial generation considered those born between 1981 and 1996, and generation Z, those born between 1996 and 2012, defy many norms of their predecessors (Pew, 2018). With 40% of millennials being nonwhite, liberal leaning, and espousing value systems that praise individuality over the status quo, they pose a stark contrast to generation X (1965-1980) and the baby boomers (1946-1960) (Accel + Qualtrics, 2016; Pew, 2018). These generational differences are apparent in many areas, including work, social, and emotional life, which has led researchers

to try and define, specifically, what makes millennials unique (Accel + Qualtrics, 2016; Moore, 2010; Yeazel, 2015). Two opinions seem to dominate the literature. One paints millennials in an astonishingly negative light, labeling them lazy, unmotivated, and easily frustrated (Moore, 2010), the other commends this generation for their open and accepting nature (Accel + Qualtrics, 2016; Yeazel, 2015). Research has shown that the millennial generation thrives on having something to fight for, prioritizing jobs that benefit society over personal financial success, and personal well-being over monetary gain (Yeazel, 2015). While some define this as positive, cause driven behavior, others view it as unmotivated, indecisive, and noncommittal (Moore, 2010). Although there are never ending debates on the adaptiveness of millennial attitudes and behaviors, where there seems to be no dispute is that millennials have redefined what is socially acceptable (Ruspini, 2019).

Generations as a Construct

The legitimacy of a generational paradigm as a classification system has been contested within the research community (see Costanza, 2018 or Zacher, 2015). It has been suggested that the set cutoff dates are arbitrary, with those on the edges possibly falling into multiple categories, and that such groupings overgeneralize attributes on a large scale. (Zacher, 2015). Generational studies have reported more inner than outer cohort difference in areas such as work behaviors or attitudes and have argued that any existing differences are not legitimate differences, rather they are socially constructed and enacted (Furstenberg, 2017; Rossem, 2016; Zacher, 2015). Zacher (2015) notes that "splitting continuous variables such as age or time into a few discrete units involves arbitrary cutoffs and atheoretical groupings of individuals" (p. 342). In the context of work differences, he suggests using Lifespan Developmental Theory as an alternative, positing three categories that impact individual development: normative age-graded influences, history-

graded influences, and nonnormative influences (Zacher, 2015). Although this may be applicable in career research, using this theory to compare men's varying views of masculinity would be inadequate, as a developmental approach does not fit. In fact, when viewing masculinity through a social constructivist approach, the concept of generations fits well, and even critics of generational theory underline the importance of generational identity. Zacher (2015) notes, "I do believe however that studying generations based on social identity and stereotyping perspective is interesting and important in the work context because, as noted by Costanza and Finkelstein, 'people believe that they exist" (p. 342). Similarly, Rossem's (2016) study on generations as social categories found that people's perceptions on the existence of generational differences resulted in such stereotypes being enacted. Rossem suggested a generational identity-based research approach stating, "generations as a workplace phenomenon must be considered a legitimate phenomenon" (p. 434). Although broad generalizations such as "millennials are all narcissistic" or "all baby boomers are technologically challenged" may be a bit extreme, there is no denying that significant differences exist between older and younger age groups.

Although exact dates have been debated, it is generally accepted that the Baby Boomer generation was born between 1945 and 1960 (Ruspini, 2019). This generation was shaped by historical events including the Civil Rights Movement (1954-1968) and the Vietnam War (1965-1973) (wmfc.org, 2019). As children in a post-war era, theorists have suggested this led to their more radical belief systems and anti-war mentality, participating in anti-war protests and other social movements (Green & McClelland, 2019). Their core values include equal rights and opportunities, personal growth and gratification, and success (wmfc.org). This generation is said to put their careers first, characterized as "living to work", resulting in their expectation that others should do the same (wmfs.org). Much of the Baby Boomer generation's identity is rooted

in their vocation, described as "work-centric" or "workaholics", with many defining their worth

through their career achievement (wmfs.org). Looking more specifically at Baby Boomer men,

research suggests that they demonstrate gender in traditionally masculine ways (Green &

McClelland, 2019). Traditional masculinity is the socialized set of expectations of attitudes,

behaviors, and feelings considered acceptable for men. Such expectations include that men

should be the "breadwinner", with their main focus on work and providing for their family

(Wester & Vogel, 2012). The Baby Boomer generation expects men to be strong, independent,

and dominant, restricting their expression of emotions, remaining a stoic and controlled presence

even in the face of strife. Research has repeatedly demonstrated that this generation holds onto

male norms, despite recent societal shifts, hypothesizing that newer masculinities may feel

uncomfortable or antithetical to their male identity (Green & McClelland, 2019).

The millennial generation born to Baby Boomers, between the years of 1982 and the late

1990's (Ruspini, 2019), greatly differ from their parents'. This generation grew up in an era of

fast paced technological advancement, with easy access world events, opposing opinions, and

cultures outside of the American norm (Ruspini, 2019). In fact, their upbringing in the age of

mass social media, has been described as this generation's defining characteristic (Green &

McClelland, 2019). Millennials root their identity in their "purpose" rather than their career,

hoping to "change the wrong they see in the world today" (wmfc.org). Such goals lead to

behaviors such as "job hopping", always striving for something better. This generation is far

more diverse racially and ethnically than their predecessors, with 40% being nonwhite, and hold

more liberal political leanings past generations (Green & McClelland, 2019; Pew, 2018). This

generation shows vast support for social movements including LGBTQIA+ rights, equal division

of labor between genders, stark anti-racist mentalities, and equal access (Accel + Qualtrics, 2016;

Pew, 2018). This generation also accepts a wide range of gender expression and a transformed

gender climate (Green & McClelland, 2019). Recently it has been acknowledged that rigid

adherence to traditional male roles can be harmful to men (American Psychological Association,

2018). Despite previous development of practice guidelines for females, ethnic, racial, and

sexual minorities, similar guidelines were not published for men until 2018 (APA, 2018; Green

& McClelland, 2019). The millennial generation has not only been a spokesperson for equal

minority rights but has also highlighted that men are also oppressed by their gender roles.

"Millennials have consistently used their voices, now that they are adults, to challenge the

traditional male role" (Green & McClelland, 2019, p. 10).

Heightened awareness has expanded the gender repertoire of millennial men. Popular TV

shows depict men breaking out of traditional male role norms by being emotional, wearing more

traditionally "feminine" attire, and engaging in activities both in line and counterintuitive to

traditional male norms (Green & McClelland, 2019). Research has demonstrated that young boys

engage in behaviors outside of traditional norms, which supports the theory that gender role

expectations are socialized. Men are not born adhering to traditional masculinity, they are taught

to do so through interactions and witnessing such enactments, which are directly influenced by

the social and political climates at the time (Addis & Hoffman, 2020). Millennial men have been

surrounded by vast political advancement, witnessing the first ever Black president-elect, the

first transgender state senator, the legalization of gay marriage, and the open protest of gender

discrimination and sexual harassment in the workplace. It would be naive to say that these men

hold the same values, engage in the same behaviors, or enact the same male role norms that

previous generations of men do. Referring back to the concept of "generational-identity"

(Rossem, 2016) the generational-identity of millennial men, appears to differ significantly from those who came before them.

Generation and the MRNI

These differences, however, are not accounted for by the current masculinity literature nor the construct depicted in the MRNI. Recent social and political movements have also resulted in a heightened , yet negative focus on traditional male norms and the patriarchal structure of society, criticizing them as harmful to both women and men (O'Neil, 2008; Wester & Vogel, 2012). For example, some aspects of traditional masculinity have been correlated with poorer psychological outcomes, higher levels of relationship distress, and higher levels of substance use (O'Neil, 2008; Wester & Vogel, 2012). Such consequences undoubtably impact not only the desire of millennials to enact such masculinity, but also their willingness to admit to such beliefs, falling victim to stereotype threat. Research shows that those experiencing stereotype threat, or the fear of being perceived as a negative stereotype, change their behaviors (Aronson et al., 2013). In fact, students have been shown to alter their self-descriptions to portray a self-image different from such stereotypes (Aronson et al., 2013). It is possible that millennial men not only view masculinity differently than their predecessors but are also making conscious efforts to be perceived as the opposite of previous masculinities, based on the environmental cues they are receiving. Again, focusing back on social constructivism, admitting or enacting traditional masculinity proves dysfunctional for millennial men. Though one could argue these men were not raised "differently" than older generations, the current cultural zeitgeist has provided a strong stance on what it means to be a functional man in today's society. Bringing it back to the MRNI, Eagly et al., (2020) notes that the "likability" or "political correctness" of a construct, can influence how one responds to tools measuring such. If that is the case, then millennial men's

scores on measures such as the MRNI may also be inaccurate due to social desirability. Expressions of masculinity are undoubtably influenced by the social climate and the nature of masculinity has been changing since the development of these instruments (Rimmer, 2015). The more society pushes back against traditional masculinity, the more taboo it becomes, the less functional the construct, the less men will want to be associated.

The study of masculinity is relatively new, compared to other areas in psychology (Addis & Hoffman, 2020, Brooks & Elder, 2016). Though multiple dates of onset could be argued, the late 1980s mark a significant "explosion" of research on the subject (Addis & Hoffman, 2020). Initial research on masculinity suggested that it was a positive trait, however, a marked shift in psychological opinion, beginning in the 1970s, argues a different stance, that masculinity can also be harmful (Addis & Hoffman, 2020; Wester & Vogel, 2008). Such findings were based on self-reports and self-report measures attempting to operationalize the masculine experience. To fully understand this shift, a brief historical examination of the study of men, through the lens of masculinity inventories, is necessary. This literature review will cover the study of men and masculinity from a historical context, the development and current research of the MRNI, and how using statistical procedures such as measurement invariance can strengthen such measures.

History of the Psychological Study of Masculinity

Early years of psychology were paradoxical for men, though they were the majority in the profession, they were rarely viewed as a unique category of subjects being studied, rather they were viewed as the norm (Addis & Hoffman, 2020; Brooks & Elder, 2016, Wester & Vogel, 2012). Beginning in the early 20th century, however, research on sex differences began to emerge. Initially, these differences were based in evolutionary and biological understandings and studies attempted to provide evidence for significant trait variance between the male and female

16

sex (Addis & Hoffman, 2020). However, such efforts proved futile. In fact, more significant variation was found within sex groups than between (Addis & Hoffman, 2020). Despite such findings, researchers continued to search for differences, with the assumption that males were superior, leading to the development social learning theories of sex/gender roles (Addis & Hoffman, 2020). Sandra Bem, a pioneer researcher in the field, developed the Bem Sex Role Inventory (BSRI; Bem, 1974; 1975). This scale assessed both masculinity and femininity with separate scales and operated under the hypothesis that men's or women's strict adherence to either masculinity of femininity was not inherently healthy, rather that androgyny, or a mix of both masculine and feminine was more adaptive (Bem, 1975). However, the specific study of men as gendered beings was still lacking until 1981 with the publication of Joseph Pleck's *The Myth of Masculinity.* In this book, Pleck outlines his social learning theory of gender role strain (GRS). His theory was the first to propose that the rigidity of the male gender role could have negative consequences, suggesting that "men learn a set of societally based prescriptions regarding appropriate masculine behaviors that are maladaptive, problematic, and ultimately unrealistic" (Wester & Vogel, 2012, p. 373). Within GRS are three types of gender role stress: gender role discrepancy, gender role trauma, and gender role dysfunction (Pleck, 1981). Gender role discrepancy suggests that most men will not meet the gender role set out for them, resulting in internal stress when facing this reality. Gender role trauma posits that the process of attempting to meet gender role prescriptions is traumatic, even if they are consistently met. Lastly, gender role dysfunction suggests that the behaviors prescribed to men are harmful, resulting in poor outcomes for everyone involved.

Continuing with a social learning lens, Jim O'Neil built upon Pleck's GRS paradigm and proposed gender role conflict (GRC) instead of strain (O'Neil, 1986/7?). Although similar to

GRS, GRC hypothesizes that socialized masculinity results in internal conflict for men when such gender expectations conflict with other roles (i.e. father, husband, professional, etc.). From this paradigm came the Gender Role Conflict Scale (GRCS, O'Neil, 1987), measuring four patterns of conflict, restricted emotionality (RE), restricted affectionate behavior between men (RABBM), conflict between work and family (CBWFR), and success, power, and competition (SPC). RE is the pattern of stoic and unemotional attitudes and behavior taught to men, despite research showing that such restricted emotionality leads to poor mental health outcomes (Gerdes & Levant, 2018; Hayashi, 1999). RABBM is the socialized notion that affection between men, romantic or not, is considered unacceptable or unmasculine. CBWFR is the conflict men experience between family expectations and the socialized norm of men being fiscally successful and providing for their families. Lastly, SPC is the expectation that men should be successful, independent, and competitive, especially in the work sphere. Each one of these scales demonstrates the societal expectations placed on men and how, depending on context, such expectations can result in internal and external conflict for men. For example, to be successful in one's career, it is typically expected that one works long hours, sometimes even outside of work to one's career. These expectations may directly conflict with family expectations of spending time with one's partner or children.

Moving slightly away from a strict social learning lens, Levant et al. (1992) proposed understanding masculinity thorough a set of socialized norms. They argued that previous understandings relied on stereotypes, or descriptive understandings of men, without an understanding of the prescriptive nature of socialized masculinity (Levant et al., 1992). Rather than simply understanding what people "think men are like", Levant and colleagues wanted to understand what others thought men "should be like" or male role norms (p. 326). From their

research came the Male Role Norms Inventory (MRNI) a 58-item, seven factor, self-report scale theorized to measure seven male role norms: avoidance of femininity, homophobia, self-reliance, aggression, achievement/status, attitudes towards sex, and restrictive emotionality (Levant et al., 1992). In 2007, this subscale was revised, creating the MRNI-R (Levant et al., 2007). The new scale, a 53-item, seven-factor, self-report measure, assess seven male role norms including avoidance of femininity, fear and hatred of homosexuals, extreme self-reliance, aggression, dominance, non-relational attitudes toward sexuality, and restrictive emotionality (Levant et al., 2007). Despite this paradigm's more nuanced understanding of masculinity, researchers still felt that it failed to capture the impact of context on the male role, leading them to propose a social constructivist understanding (Addis & Mahalik, 2003).

Social constructivism places men at the center of their own, individual, masculinity development (Brooks & Elder, 2016). From this perspective masculinity is conceptualized as more fluid, impacted by contextual variables such as age, sexual identity, race, SES, and circumstance (Wester & Vogel, 2012). Such variables may influence what type of conflict arises for men, along with the norms they become accustom to. Social constructivism highlights that not all men have the same understanding of masculinity, nor do they enact or adhere to the same norms (Brooks & Elder, 2016). Addis et al., (2003) attempted to capture true enactments and developed the Conformity to Male Norms Inventory (CMNI). Rather than measuring what men "should" do, the CMNI assesses the individual's adherence to male norms. Measuring individual conformity is important, as although a man may believe men should act in a specific way, this belief does not necessarily imply that he himself engages in these behaviors, beliefs, or attitudes.

Models of the MRNI

As the MRNI is to be used for the purpose of this study, a more in-depth discussion follows. The MRNI model originally proposed by Levant et al., (2007) is bifactorial in nature, with the first being how the 51 items of the MRNI-R load on a single factor while simultaneously loading on seven individual subscales. However, there exists some debate in the psychology of men literature (see, for example, Levant &Richmond, 2016, as well as Wester & Vogel, 2012) as to the viability of this bifactorial understanding. Therefore, this research will test aspects of Levant's model. Briefly, this work will test the viability of Levant (e.g., 2007; see also Lee, 2018) general masculinity construct. It is made up of all 51 items of the MRNI-R. Next, Levant also asserts that each of the seven subscales (Avoidance of Femininity, Fear and Hatred of Homosexuals, Extreme Self-Reliance, Aggression, Dominance, Non-relational Sexuality, and Restrictive Emotionality) were presumed to capture domains of male gender role norms as measured by their individual MRNI-R items. This work will test each of those subscales.

The MRNI is the most psychometrically investigated of all masculinity inventories (Gerdes et al., 2018) and has gone through many revisions and abbreviations over time. Initially proposed by Levant et al., (1992) it suggests seven norms of masculinity ideology: Avoidance of Femininity, Fear and Hatred of Homosexuals, Self-Reliance, Aggression, Achievement/Status, Non-Relational Attitudes towards Sex, and Restrictive Emotionality. The original MRNI also included a non-traditional attitudes subscale, in which higher scores suggested endorsement of non-traditional ideology (Levant et al., 1992), however this scale was removed with the publication of the revised measure (Levant et al., 2007). Each subscale is scored on a seven-point Likert-type scale, with higher scores indicating stronger endorsement of norms. Subscale scores are calculated by finding the mean score for each subscale and the total score is found by calculating the mean of the subscales excluding the non-traditional attitudes subscale (Levant et

al., 1992). Empirical evidence supporting the reliability of the original MRNI has been inconsistent. The MRNI has shown high temporal stability, otherwise referred to as test-retest-reliability, across a three-month period of time with men scoring .65 and women .72 (Heesaker & Levant, 2001) however, subscore reliability has varied. For example, Levant and Majors (1997) examined MRNI scores between European Americans and African Americans, and reported strong reliability statistics for the subscales of avoidance of femininity (.77), fear and hatred of homosexuals (.54), self-reliance (.54), aggression (.52), achievement/status (.67), non-relational attitudes towards sex (.69), restrictive emotionality (.75), non-traditional attitudes (.57) and total score (.84). However, Levant, Cuthbert and colleagues (2003) examined scores between American participants and Russian participants and had unacceptable reliability statistics for subscales of fear and hatred of homosexuals (.42), aggression (.48), and non-traditional attitudes (.47) suggesting that such scores may not be generalizable across populations.

Studies have also examined the MRNI's construct validity. Discriminant validity, or statistical evidence that the MRNI construct is distinct or unique from other measures, was established by comparing the correlation of total MRNI scores to a different measure of gender, the Personal Attributes Scale (PAQ; Spence & Helmreich, 1978). Levant and Richmond (2007) looked at the correlation between the PAQ M scores (M meaning typical male personality traits) and total MRNI scores and found no correlation between the two, demonstrating strong discriminant validity (r = .06). To assess convergent validity, or similarity between one measure of a construct with similar measures of similar constructs, the MRNI was compared with two other gender measures, the Gender Role Conflict Scale-I (GRCS-I; O'Neil, Good, & Holmes, 1995) and the Masculine Gender Role Stress Scale (MGRSS: Eisler, 1995). Convergent validity was also supported, as the MRNI total score was significantly correlated with both the GRCS-I (r

= .52, p < .001) and the MGRSS (r = .52, p < .001). With both discriminant and convergent

validity established, overall construct validity of the MRNI could be concluded.

Revisions to the MRNI

Despite generally sound psychometric properties, the MRNI has undergone multiple

revisions. The Male Role Norms Inventory-49 was developed to focus specifically on traditional

masculine ideology, removing the non-traditional attitudes scale (MRNI-49; Berger, Levant,

McMillan, Keleher, & Sellers, 2005) and also improved the subscale labeled fear and hatred of

homosexuals. In an attempt to create a version applicable to adolescent men, Brown (2002)

developed the Male Role Norms Inventory-Adolescent (MRNI-A) (Levant et al., 2012). This

version utilizes a more appropriate reading and comprehension level for younger men making

two major changes from the MRNI: changing items to be more specific to adolescent

circumstances (e.g., "Boys should not be afraid to go inside a haunted house") and the fear and

hatred of homosexuals and non-relational attitudes towards sexuality subscales were removed

(Brown, 2002; Levant et al., 2012). Unfortunately, the reliability of the MRNI-A was

questionable, resulting in the development of the MRNI-A-revised (MRNI-A-r; Levant et al.,

2012). The MRNI-A-r is a 41-item self-report measure of masculinity ideology for use among

adolescent men. In developing this scale, two items from the MRNI-A were removed due to poor

coefficient alphas (items 38 and 42), language was further simplified, and all negatively worded

items were reworded positively. The five subscales remained the same as the MRNI-A (Levant

et al., 2012). Overall, the MRNI-A-r showed promising initial psychometric properties including

empirical support for the a three factor structure, through an exploratory factor analysis, strong

internal consistency ($\propto = .89$), adequate convergent validity between total scores of the MRNI-

A-r and the Adolescent masculinity Ideology in Relationships Scale (AMIRS; Chu, Porche, &

Tolman, 2005), and discriminant validity by comparing scores on the PAQ-M and the MRNI-A-r (Levant et al., 2012).

In 2007, Levant and colleagues, using a mostly Caucasian (50.6%) sample of 170 undergraduate students (38 male, 132 female) created the MRNI-Revised (MRNI-R) to address both psychometric and construct changes needed to keep the MRNI current. This included updating the language, defining the construct and subscales more accurately, and improving subscale reliability (Levant et al., 2007). Similar to the MRNI-49, the MRNI-R only assess traditional masculinity, rather than both traditional and non-traditional attitudes. Upon initial development, the inventory showed strong reliability with Cronbach's alphas ranging from .75 to .95 for both subscales and total score (Levant et al., 2007). The subscales were more strongly correlated with the total score than they were themselves, suggesting that the scales were distinct aspects of the latent construct of masculinity. Construct validity was determined by analyzing the endorsement patterns between demographics (Levant et al., 2007). It was found that age and education were significantly correlated with MRNI-R total scores, however socioeconomic status was not. Significant differences between sexes were present on all subscales, with men scoring higher than women, excluding the subscale of extreme self-reliance, which mimics previous findings (Levant et al., 2007). Additionally, differences were found between racial and ethnic groups, with both African American and Asian American men endorsing traditional ideology significantly more than European American men, again replicating previous research. However, this was the first study to directly compare Asian and European American men (Levant et al, 2007). Slight revisions have been made to the MRNI-R since it's development, however, with a largely homogeneous sample, the generalizability of the MRNI-R still comes into question,

especially as variance between various groups, such as generational cohorts, remained unexplored.

Levant and colleagues (2010) further assessed the factor structure and construct validity of the revised measure using exploratory factor analysis procedures. Using another majority White (83%) sample of 593 participants (42% women, 58% men, 1 participant indicating other), they found support for the seven-factor structure along with adequate support for item placement within subscales, however three factors were removed due to cross loading (Levant et al., 2010). Construct validity was also supported through examining convergent and discriminant validity. Strong correlations were found between the MRNI-R and three other measures of masculinity, the Conformity to Masculine Norms Inventory (CMNI; Mahalik et al., 2003), the Gender Role Conflict Scale (GRCS; O'Neil et al., 1986) and the Male Roles Attitudes Scale (MRAS; Pleck et al., 1994). Conversely, no correlations were found between the MRNI-R and the Personal Attributes Questionnaire (PAQ; Spence & Helmreich, 1978). Minor adjustments were made to subscale names, to better define the construct of each subscale based on factor loadings. Aggression was renamed "toughness" to account for the retention of only four of the seven original items, non-relational sexuality was renamed "importance of sex" for the same reason, and self-reliance was renamed "self-reliance through mechanical skills" again, due to only three of the seven items loading onto the subscale (Levant et al., 2010). Additionally, five items with poor fit were removed as such was deemed not to impact internal consistency. The resulting scale was a 40-item, seven factor self-report measure. Again, however, the homogenous sample calls the measure's generalizability into question and generational related comparisons were not made. Since the aforementioned revisions to the MRNI, no updates have been made to the measure. Although shortened versions of the scale have been developed, such as the MRNI-

Short Form (MRNI-SF; Levant, Hall, & Rankin, 2013) and the MRNI-Very Brief (MRNI-VB; McDermott et al., 2019), the overall measurement construct has remained the same.

Traditional Masculinity and Men

Traditional masculinity ideology (TMI), initially coined by Pleck (1995) is the "common set of standards and expectations associated with the traditional male role through most of the world" (Levant et al., 2010, p. 25). These standards and expectations are socialized into boys who, across cultures, play the same roles of provider, protector, and parent resulting in similar male role norms international (Levant et al., 2010). Although some cultures have been found to operate more androgynously (e.g. Tahiti and Semai), the gendered norms taught to young boys, who grow into men, are rather universal (Gilmore, 1990; Levant et al., 2010). The MRNI, as previously discussed, is one of the most widely used instruments in the assessment of TMI. It defines the construct of masculinity through these sets of socialized beliefs and behaviors, breaking them down into a set of seven, universal male norms: restrictive emotionality, Self-Reliance through Mechanical Skills, Negativity toward Sexual Minorities, Avoidance of Femininity, Importance of Sex, Toughness, and Dominancy (Levant et al., 2007; Levant et al., 2010). As previously discussed, it was not until the late 70s and early 80s that such norms were challenged by feminist movements and writings, highlighting that TMI may provide both privilege and pain to those who enact such behaviors and attitudes (Wester, 2008). Since then, much research using masculinity measures such as the MRNI have been conducted, attempting to assess how varying levels of TMI may positively or negatively impact men (Levant & Richmond, 2007). A brief review of this literature is to follow.

MRNI Demographic Variance

The endorsement of traditional male norms has been assessed across multiple racial and ethnic groups. Levant and Majors (1997) compared total MRNI scores, (using Levant's original MRNI model from 1992), between African American and European American men and women and found that African American men scored significantly higher than all others, followed by European American men, African American women, and European American women. They suggested this difference lied in the higher levels of both racial and gender role strain experienced by African American men in the United States (Levant & Majors, 1997). In a follow up study, Levant Majors, and Kelly (1997) found that not only was race related to masculinity endorsement, so was geographic location, with those in the south scoring higher than those in northern states. Such results have since been replicated, as previously mentioned, by Levant et al., (2007) in their initial validation of the MRNI-R, finding that African American and Asian American men scored higher than European American men, suggesting higher levels of TMI endorsement for both groups.

A few cross-national studies have been conducted comparing TMI between U.S. college students and Chinese and Russian college students (Levant & Richmond, 2007). Overall, nationality was found to have an even stronger impact than sex on overall TMI endorsement, with both Chinese and Russian men and women endorsing higher levels of TMI than their European American counterparts (Levant, Wu, & Fischer, 1996; Wu, Levant, & Sellers, 2001; Levant, Cuthbert, et al., 2003). Comparisons have also been made between U.S. and Japanese college students, with Japanese men and women endorsing higher TMI (Hayashi, 1999) and between U.S. and Pakistani college students, with Pakistani men and women espousing significantly higher masculine norms than American individuals (Rashid, Yasin, & Massoth, 2000). Overarchingly, the trend seems to follow a similar pattern: European American men

endorse TMI significantly more than women, however, U.S. racial minority and ethnic groups and international participants demonstrate significantly higher MRNI scores than their American counterparts (Levant & Richmond, 2007).

Other demographic variables have been analyzed including sex, social class, age, sexual orientation and marital status. Men tend to endorse higher levels of TMI than women, as to be expected, with younger men scoring higher than older men in most instances, however, one study conducted by Young (1995) demonstrated that sons tended to espouse less traditional beliefs than their fathers, contradicting other reports. Research shows that single men report higher levels of TMI than married men and those in lower class systems also report higher TMI, however, due to the homogeneity of participant samples, reports on social class are not quite clear (Levant & Richmond, 2007). Lastly, gay men show less traditional attitudes than their heterosexual counterparts, excluding the scale of self-reliance (Massoth, Broderick, Festa, & Montello, 1996 as cited in Levant & Richmond, 2007).

MRNI and Well-Being

As scholars proposed that masculinity may negatively impact men, researchers began exploring the link between TMI, as defined by the MRNI, and varying attitudes, behaviors, and overall wellbeing. Although other research exists using various versions of TMI, such as those defined by the CMNI or GRCS, the current focus shall remain on MRNI literature. Smith (1996) looked at risky sex behaviors of men who endorse high TMI and found that those scoring higher on the MRNI had negative beliefs about openly discussing condom use with partners. Higher MRNI scores have also been correlated to lower levels of relationship satisfaction (Mcgraw, 2001) and higher levels of intimacy fear (Maxton, 1994). Wade & Brittan-Powell (2001) explored traditional masculinity related to self-concept and found that those who endorsed

27

traditional masculinity were more likely to hold sexist beliefs along with negative attitudes

towards racial and gender equality. They noted that such attitudes were "conducive to the sexual

harassment of women" (p. 42), suggesting that TMI may be related to such behaviors, however

this study was also limited by its undergraduate college sample of over 80% White men and had

noted low reported consistency reliability (<.70) (Wade & Brittan-Powell, 2001). Important to

note, however, is that this study demonstrated that men seem to base masculinity around what it

is not, rather than what it specifically is, demonstrating how TMI can vary man to man. Bonney,

Kelley, and Levant (1999) attempted to generate a statistical model to understand the interaction

between TMI and paternal involvement. They found that higher scores on the MRNI were

negatively correlated with paternal engagement and positively correlated belief that the father

should play a less integral role than the maternal figure (Bonney, Kelley, & Levant, 1999). In

their discussion of the findings, they note that previous studies theorized paternal engagement as

being determined by societal expectations, however their results indicated that such engagement

may actually be self-determined, based on personal TMI endorsement. However, these results,

specific to the MRNI are over two decades old and have not been extensively explored,

especially with younger generations of fathers. MRNI scores have also been correlated with

sexual aggression, which Levant & Richmond (2007) discuss in their meta-analysis of MRNI

research, however, some of their results rely on dissertation abstracts and should continue

explored. They note that the relationship between MRNI scores and sexual aggression were

shown to be mediated by factors such as motivation, rape supportive attitudes, and peer group

norms (Gale, 1996 as cited in Levant & Richmond, 2007) but surprisingly, men enrolled in

domestic violence treatment programs endorsed lower levels of TMI than previous participant

bases (Teofilo, 1999 as cited in Levant & Richmond, 2007). However, researchers attributed

such results to treatment experience, as it was counter to their original hypotheses. In the context of the current study, peer group norms and societal attitudes towards topics of paternal engagement and sexual aggression seem to have drastically shifted, highlighting the need for further exploration of the MRNI construct.

MRNI, Mental Illness, and Help Seeking

The relationship between traditional masculinity, mental illness, and help-seeking has been of the utmost interest to researchers in recent years. It has been well documented that men more often avoid help seeking because it defies socialized gender roles and expectations (Addis & Mahalik, 2003; Berger et al., 2005; Vogel & Wester, 2014). In fact, it has been found, that traditional masculinity ideology (TMI) is strongly correlated with psychological help seeking avoidance, more so than gender role conflict and alexithymia (Berger et al., 2005). Higher levels of TMI were also correlated with more negative attitudes towards help-seeking (Berger et al., 2005). These findings are problematic as men experience mental health concerns as much, if not more than other genders and, generally, are exposed to more traumatic life events than women (Olff, 2007). More specifically, endorsement of TMI has also been linked to higher levels of substance use, depression, and anxiety (O'Neil, 2008; Wester & Vogel, 2012) and has been correlated with higher levels of alexithymia in men on multiple occasions (Fischer & Good, 1997; Levant, Richmond, et al., 2003; Levant et al., 2006). Research has shown that psychological services are equally beneficial to both men and women, however, male engagement in such is low (Addis & Mahalik, 2003). Although some have attempted to implement interventions geared towards men (e.g. Rabinowitz, 2014) without general male buy in to the benefits of psychotherapy, and the barriers posed by TMI, men still remain undertreated.

General MRNI Research Gaps

A recent content analysis on MRNI research outlined the topics and findings to date on traditional masculinity ideology (Gerdes et al., 2018). They found the most studied topics included mental health, emotions, physical health, relationships, psychometrics, racial/ethnic minority group differences, and help seeking. It was noted that a majority of publications were post 2010 and that the frequency of these six common topics had remained stagnant across time (Gerdes et al., 2018). One major limitation they found amongst all studies was the overarching use of homogenous college student samples who do not mimic the general makeup of the general population. Another important gap, however, not noted, is the lacking breadth of subjects explored. Although the MRNI was revised in 2007 to account for some societal change and advances, it has not been revised since, ignoring the multitude of transitions that have taken place the past 14 years. For example, although comparisons have been made between ages and TMI endorsement, invariance has not been established between generational cohorts' MRNI scores. Additionally, although many studies have analyzed the psychometric properties of the various versions of the MRNI, very few have assessed whether the same construct is being measured between populations, or in other words, few have assessed the measurement invariance of the MRNI and other inventories.

Measurement Invariance

The aforementioned paradigms and accompanying scales each theorize a different, yet similar version of traditional, or hegemonic masculinity, grounded in the time in which they were developed. Hegemonic masculinity, originally proposed by Connell (1995), is the socially accepted and idealized set of traits and behaviors ascribed to men. Originally viewed as a static trait, or a "dominant, societal-wide idealized blueprint", research has shown that masculinity is not fixed over historical eras, "rather, hegemonic masculinities are the most widely accepted

forms of being a man as defined by the historical era, social institution, or community"
(Thompson & Bennett, 2015, p. 116). Such research suggests that there is no singular
masculinity, rather there are masculinities, of which one may be hegemonic at a period of time,
for a group of people (Thompson & Bennett, 2015). As research began to present the idea of
masculinity as both an individual and sometimes collective experience, studies of measurement
invariance were used to assess between group differences on masculinity measures.

Measurement invariance is a statistical property that demonstrates whether or not a construct
is being measured the same way for two or more groups of people (Lee, 2018). As psychological
constructs are frequently theorized from a specific lens, for example traditional masculinity
developed through a young, White, hetero and Eurocentric viewpoint, they frequently differ
between groups (Levant & McCurdy, 2015). Measurement invariance provides insight into such
differences, by comparing scores and factor loadings on psychological measures, to assess if the
same construct is being measured, further demonstrating generalizability. This is evidence that a
given instrument truly measures the same construct, in the same way, across groups (Cieciuch &
Davidov, 2015). There are multiple types or steps of a measurement invariance study (MI)
including configural, metric, and scalar (Lee, 2018). Each of these levels of MI are based on the
constraints or parameters imposed on the factor loadings or intercepts of the sample (Cieciuch &
Davidov, 2015). The weakest or loosest level of invariance is configural invariance (CI) because
it imposes no constraints, or restrictions, on the existing model (Lee, 2018). Rather, it tests to
see, if the model as is, resonates, or is interpreted the same way by two groups (Crowson, 2020;
Li et al., 2021). Testing CI examines the general factor structure of a measure, assuring it fits for
both groups.

Metric invariance, a stronger level of MI, looks at factor loadings, which represent the direction and strength of relation between observed variables and latent variables, between groups by constraining them to be equal (Lee, 2018). Here, constrain means setting the factor loadings to be the same across groups, so that the scale intervals are the same across groups, allowing one to compare the unstandardized regression coefficients or covariances between groups (Pirralha, 2020). In other words, it allows us to see if the two groups are interpreting the measure similarly (Crowson, 2020). Once these constraints are put in place, the model fit is reexamined, assessing whether with these constraints, the model still works for both groups (Lee, 2018). Lastly, the strongest or most constrained model, scalar invariance, looks at item intercepts, again searching for equal group loadings (Lee, 2018). To test for scalar invariance, the factor loadings are left equally constrained and equality constraints are imposed on the item intercepts as well (Pirralha, 2020). Item intercepts are the expected values of the observed variables if the mean of the factor is set to zero (Pirralha, 2020). The model is then assessed again, with all equality constraints imposed, for goodness of fit (Lee, 2018).

Beginning with the CMNI, researchers have questioned whether contextual variables were truly being evaluated for varying cohorts of men. Theoretically the CMNI allows researchers to assess contextual impacts, as its focus is on personal conformity, however this has not always proven accurate. For example, Hsu and Iwamoto (2014) compared CMNI-46 scores between White and Asian American college men and found significant measurement variance between the two groups at the metric level. More specifically, they noted that the CMNI was "more theoretically consistent for White American participants" (p. 11). More simply put, the CMNI was measuring White American masculinity, rather than a "general" masculinity. Hsu and Iwamoto (2014) suggested that more attention be paid to how environmental factors may

influence men's gender role conformity, in this case, racial and ethnic identity. In a similar study, Parent and Smiler (2012) evaluated measurement invariance between men and women on the CMNI-46. Consistent with their hypothesis, scalar variance was present, suggesting that men score higher on conformity than women, however, variance was not found on the configural and metric levels, meaning the scores were comparable across genders (Parent & Smiler, 2012). Although these results provided some support for the CMNI factor structure, the study was limited by its mostly White sample and disregard of age (Parent & Smiler, 2012). This has been a common critique of research samples, as White men in their 20s have been the dominant participant base (Gerdes et al., 2018). And varying levels of masculinity endorsement have been demonstrated between ethnic groups, suggesting that conformity to such norms may vary with culture (Wester & Vogel, 2012). Additionally, as age differences were not examined, these results cannot be generalized across age groups (Parent & Smiler, 2012). In fact, although there are many published MI studies using the CMNI, generational MI analysis is absent.

Similar studies have been conducted using the MRNI, again comparing scores between genders, ethnicities, and sexual orientation, however research on MRNI scores between generational cohorts is lacking. For example, in developing a short form of the MRNI, Levant, Hall, and Rankin (2013) investigated measurement invariance across gender. Although they found full scalar invariance, only partial metric invariance was found, with unequal general masculinity factor loadings across groups. Partial invariance is when parameters of at least two factors, but not all, are equal between groups (Pirralha, 2020). Some argue that this is enough to prove invariance is present (Steenkamp & Baumgartner, 1998) where others argue the opposite (Steinmetz, 2018). However, Levant, Hall, & Rankin (2013) explain that in their study, partial invariance means men and women were not interpreting the construct the same way. In another

study using the MRNI-SF, McDermott et al., (2017) assessed MI between White men, men of Color, and gay men. Similar to previous studies, variance was found between groups. Only partial metric invariance was discovered when comparing men to women, White, Black, and Asian men, and heterosexual to gay men (McDermott et al., 2017). The comparison between heterosexual White men and heterosexual Black men was the only demonstration of full metric invariance, highlighting the bias present in the general factor of masculinity. Instead of measuring a generalizable masculinity factor, the MRNI demonstrates a heteronormative, Eurocentric understanding of the construct (McDermott et al., 2017). In fact, the researchers state, "[the] factors may represent somewhat different constructs in other cultural groups" and that "major differences across cultural groups appear to be on the general TMI" (p. 735). These results proved similar to previous findings exhibiting higher levels of traditional masculinity for racial and ethnic minority men when compared to the majority (see Levant & Richmond, et al., 2003).

Gaps in MI Research

These studies all provide important findings regarding the applicability of current masculinity scales to varying groups of men. However, there seems to be a pattern of MI related topics, gender and racial and ethnic make-up. Although some studies have looked at age related outcomes (e.g. Berger et al., 2005), none have analyzed generational variance. In fact, when analyzing existing MRNI research, Gerdes et al. (2018) note that age-related sample diversity is wanting. Out of 84 published papers, only one looked at older adults and only three examined adolescents. They go on to note that even the topics studied related to the MRNI were repetitive, stating that, "studies using the MRNI examined a wide range of topics, most of which were studied repeatedly" (p. 589). Such topics included, mental health, emotions, physical health,

relationships, psychometrics, racial/ethnic minority group differences, and help seeking. Again, age or generational differences remain absent. Such a gap is problematic. Masculinity for younger generations (e.g., Gen Z and Millennials) is possibly quite different than for older generations (e.g. Gen X or Baby Boomers). As Addis & Hoffman (2020) note, psychology reflects the gender politics of a society, which have changed drastically in the last decade. Each masculinity paradigm and accompanying scale tell a story of the underlying social understanding of masculinity at varying points in time, meaning that the MRNI paradigm represents and understanding from the early 1990's and 2000's (Levant et al., 1992; Levant et al., 2007). At the time of scale validation, Gen Z and very few Millennial men were not of age to take part in such studies, meaning their demographic is missing from the measured general masculinity factor. Additionally, as previously mentioned, samples from previous studies were quite homogenous, White, heterosexual men (Gerdes et al., 2018). However, recent polls show that these younger generations are significantly more racially and ethnically diverse than their predecessors (PEW, 2018).

Expanding Research within the Psychology of Men

Despite recent achievements and awareness within the study of men and masculinity, the field has also come under scrutiny in recent decades. Masculinity measures have been criticized for lacking generalizability, ill-fitting factor structures, and poorly operationalized definitions of male norms or masculinity (Cuthbert, 2015; Thompson & Bennett, 2015). Critics have called for diversifying samples, measures, and research methods, along with reconceptualizing general factors of masculinity (Levant & McCurdy, 2018). In their chapter in Wong and Wester's (Ed) (2016) Handbook of Men and Masculinities, Wong and Horn call for more diverse research practices and samples. As previously discussed, age is a little studied factor in this field, leaving

one to wonder if greater understanding could be achieved by looking at such variables. In fact, Thompson & Bennett (2015) highlight the paucity of research on non-college men and ask, "Are these measures age invariant?" (p. 123). They suggest that the study of masculinity needs to be expanded as measures designed outside of a university setting are scarce, meaning young men separate from college samples, middle-age men, and older men's views are absent within published literature (Thompson & Bennett, 2015).

Looking more specifically at the MRNI measure, research has demonstrated that participants often disagree with many traditional masculinity values (Thompson & Bennett, 2019). Although this measure is theorized to assess masculine ideology, none of them clarify or identify cultural standards of respondents despite the importance of such standards, especially in an age of change surrounding gendered expectations (Cuthbert, 2015). It has been suggested that even the wording of items may influence how different men respond or interpret such statements (Cuthbert, 2015). "Most items within first-generation scales are worded as absolute statements, as if the rule applies equally across generations, contexts, and geographies" (Thompson & Bennett, 2015), yet it is known that such equality is rarely the case. When comparing Baby Boomer and Millennial men for example, it is unlikely that the two groups would construct masculinity the same way, based in their contradictory generational identities and the fact that Millennial male experiences "place their opinions in opposition to Boomers" (Green & McClelland, 2019, p. 15). How could such contrasting value systems conceptualize masculinity equally? Similarly, how can the field even theorize that differences do or do not exist, when the research is completely absent?

Studies of MI are important tools within psychological research, as psychological measures must be generalizable to be valid and useful, yet very few of these studies exist in the

psych of men literature. In conducting a brief analysis of MI studies, only 13 were found between the years of 2001 and 2020. Of these 13 studies, only one assessed age-related differences, where six assessed gender variance, two assessed racial and ethnic variance, and two analyzed variance related to sexual orientation. Despite lacking empirical support for their use with all men, measures such as the MRNI, CMNI, and GRCS are continuously utilized as standard research practice in the psych of men, which is problematic at best.

Summary

Although the study of men as distinct gendered beings is new, there is no excuse for the continued use of measures, like the MRNI, that have not been proven applicable with all men. Studies of measurement invariance provide strong psychometric support for construct invariance across groups, assuring the same concept is being measured between individuals (Lee, 2018). Despite obvious disparities between older and younger generations, little attention has been paid to possible generation related differences on the MRNI, raising the question: Are these inventories applicable to all generational cohorts or are they possibly measuring different things for each? Due to this gap in the research, this study will assess for measurement invariance on the MRNI between generations with the goals of providing either empirical support for use of the measure with all ages, or providing evidence that measures should be updated to fit millennial and younger masculinities, or be reconceptualized all together.

Chapter 3

METHODOLOGY

Levant (e.g., 2007) argues that the MRNI model is bifactorial, in that all 51 items of the MRNI-R load on a single factor while also making up seven subscales. However, psychology of men scholars continue to debate whether or not the MRNI model is in fact truly bifactorial in

nature (see, for example, Levant &Richmond, 2016, as well as Wester & Vogel, 2012) so, this research will instead test individual aspects of Levant's conceptualization. The first research question will address all 51 items of the MRNI-R, while the second will address each of the MRNI-R seven subscales (Avoidance of Femininity, Fear and Hatred of Homosexuals, Extreme Self-Reliance, Aggression, Dominance, Non-Relational Sexuality, and Restrictive Emotionality). Therefore, the general question is as follows: Does the MRNI-R still measure the same construct of masculinity for millennial and younger men as it did for their predecessors? The general prediction was that variance would exist between groups on the MRNI-R's total score, as well as the scores on each of the individual MRNI-R subscales.

Research Question and Hypotheses

1. Does the MRNI-R total score demonstrate measurement invariance between generational cohorts of men?

   a. The MRNI-R total score will demonstrate statistically significant non-invariance between cohorts of men.

2. Do each of the MRNI-R subscales (Avoidance of Femininity, Aggression, Fear and Hatred of Homosexuals, Dominance, Non-Relational Attitudes towards Sex, Extreme Self-Reliance, and Restrictive Emotionality) demonstrate measurement invariance between groups of men?

   a. The subscale Avoidance of Femininity will demonstrate statistically significant non-invariance between generational cohorts.

   b. The subscale Aggression will demonstrate statistically significant non-invariance between generational cohorts.

c. The subscale Fear and Hatred of Homosexuals will demonstrate statistically significant non-invariance between generational cohorts.

d. The subscale Dominance will demonstrate statistically significant non-invariance between generational cohorts.

e. The subscale Non-Relational Attitudes toward Sex will demonstrate statistically significant non-invariance between generational cohorts.

f. The subscale Self Reliance will demonstrate statistically significant non-invariance between generational cohorts.

g. The subscale Restrictive Emotionality will demonstrate statistically significant non-invariance between generational cohorts.

Addis and colleagues have critiqued these inventories for lacking contextual understandings (see Addis, Mansfield, & Syzdek, 2010 or Addis & Hoffman, 2020), noting the importance of understanding how masculinity situationally varies. Although such statistical variance has been explored for variables such as race or gender, it has not been assessed between generational cohorts. Existing constructs of masculinity may not provide an accurate understanding of millennial men, rather they may only provide an understanding of the masculinity grounded in the time the original construct was developed. Referring back to the social constructivist perspective, masculinity is constructed based on the demands of the situation (Brooks & Elder, 2016). The demands of the current climate seem to suggest that traditional masculinity may no longer be seen as functional.

Measurement Invariance

Measurement invariance is statistical evidence that the same construct, in this case masculinity, is being measured the same way amongst groups. In other words, it "assesses the

psychometric equivalence of a construct across groups or across time" (Putnick & Bornstein, 2016, p. 71) and can be demonstrated using an item-response theory framework (IRT) or a structural equation modeling framework (SEM) (Putnick & Bornstein, 2016). For this project, an SEM model was utilized as it is a commonly understood practice (Putnick & Bornstein, 2016; see also Svetlina, Rutkowski, & Rutkowski, 2019). SEM has become regularly used within masculinity research, especially in assessing masculinity measures such as the Conformity to Masculine Role Norms Inventory (CMNI; Mahalik et al., 2003) or the Male Role Norms Inventory-Short Form (MRNI-SF; Levant, Hall, & Rankin, 2013). For example, Levant, Hall, & Rankin (2013) used a CFA to develop and validate the short form measure of the MRNI and compared invariance between self-identified males and females. Similarly, Hsu and Iwamoto (2014) used a multigroup CFA to explore variance between Asian and White men's responses on the CMIN-46.

A SEM model uses a multigroup confirmatory factor analysis (CFA) to explore model fit (Milfont & Fischer, 2015; Putnick & Bornstein, 2016). In a CFA, the items that create a construct, are loaded onto a latent factor, or, in other words, an unobserved factor that represents the construct (Putnick & Bornstein, 2016). A multigroup CFA divides a dataset into two separate groups analyzing the model for each individually (e.g., Svetlina, Rutkowski, & Rutowski, 2019) in this case, younger and older men. In the current study, the dataset was divided by age, those falling into the millennial generation or younger, and those falling in the generation X category or higher. This allowed for multigroup comparisons to be made and assessed whether groups, in this case older and younger generations of men, interpreted measures of masculinity similarly (Bialosiewicz, Murphy, & Berry, 2013; Crowson, 2020; Li et al., 2021). A description of these types of invariance are as follows:

Configural Invariance

Configural invariance is the most lenient test of measurement invariance (Lee, 2018; Putnick & Bernstein, 2016) and was the first step of this process (e.g., Svetlina, Rutkowski, & Rutkowski, 2019). It assesses if the overarching factor structure or the MRNI-R fits for the two separate groups of men without imposing any constraints, or, in other words, it determines if the construct (masculinity) has the same pattern of both fixed and free loadings across groups (Putnick & Bernstein, 2016; Rönkkö, 2020).

Following a traditional CFA, therefore, the strength of the relationship between each MRNI-R item (i.e. scale questions) and the levels of Levant's proposed bifactor construct of masculinity were analyzed. The strength of this relationship is called a factor loading and must meet the acceptable value of .7 or higher to be deemed strong (Klein, 2005; Lee, 2018). The item's origin term is labeled the "item intercept", or the expected value of Y when X equals zero (Lee, 2018; Rönkkö, 2020). The total score of the MRNI-R, as well as the scores on each subscale, were all assessed for configural invariance by fitting the accepted model to both millennial and generation X groups, leaving factor loadings and item intercepts "free to vary for each group" (Lee, 2018). In this case, vary means the factor loadings and intercepts are allowed to differ from one another. If the model structure fit for both age groups, meaning the factor loadings all surpassed the accepted significance level, it would suggest that the current factor structure held for both age groups. Alternatively, if configural invariance did not hold, suggesting the model does not fit for both age groups of men, no further invariance investigation would be needed, as this would imply the two groups do not interpret or respond to the MRNI-R in a significantly similar way (Lee, 2018).

Metric Invariance

The next step was to test for metric invariance (e.g., Svetlina, Rutkowski, & Rutkowski, 2019), otherwise known as loading invariance, for those models where configural invariance was present (Lee, 2018). Metric is considered a weak form of measurement invariance, but stronger than configural (Putnick & Bernstien, 2016; Rönkkö, 2020). It investigates whether the factor loadings are equal between groups or that each item (i.e. MRNI-R inventory questions) contributes to the latent construct of masculinity in a similar fashion across groups (Putnick & Bernstien, 2016; Lee, 2018). Similar to the process of configural invariance, the same model fit is set for both groups, however this time the factor loadings were constrained to be equivalent, leaving item intercepts free as before (Lee, 2018). Fitting the factor loadings to equal means that the strength and direction of the observed variables to the latent variables were set to be equal between age groups (Pirralha, 2020). More specifically, the first factor for the younger group was set to equal the first factor for the older group on those scales that passed the configural invariance test. Good model fit for both groups would support the presence of metric invariance, where poor fit would indicate that the factor loadings vary, or provide evidence for metric non-invariance (Lee, 2018). This pattern of results would suggest that the current MRNI-R structure does not fit for both groups of men, or that the existing subscales to not capture younger men's definition of masculinity.

Scalar Invariance

Lastly, for those models that demonstrated strong metric invariance, scalar (or intercept) invariance was assessed. Scalar invariance is considered a strong form and examines if the item intercepts, or the expected means of Y when X equals zero, are equivalent between groups (Lee, 2018; Rönkkö, 2020) while factor loadings are also held equal (Pirralha, 2020). Here, the item intercepts are constrained to be uniform, just as the factor loadings were when analyzing for

metric (Lee, 2018; Pirralha, 2020). The first item intercept for one group is set to match the first item intercept of the other group for each construct. Typically, this first intercept otherwise known as the reference intercept, is set to 0, as a base point, meaning all loadings and intercepts are constrained to be equal across groups (Rönkkö, 2020). Should this result in poor model fit, it can be assumed that item intercepts significantly differ, and in this case generational cohorts construe masculinity in different ways. For the scales that demonstrated metric invariance, the constraints imposed were kept, and additionally, item intercepts were set to be equal between younger and older men (Pirralha, 2020). Differences in intercepts were then assessed.

Interpretation of Invariance Results

The current study used IBM SPSS Amos to conduct the SEM analysis. Amos is a SEM software plug-in that expands on the capabilities of SPSS to include the ability to build models utilizing data contained within standard SPSS (IBM, 2022). Multiple protocols exist in the literature around interpretation of measurement invariance analyses (Crowson, 2020; Pirralha, 2020; Svetlina, Rutkowski, & Rutkowski, 2019). For example, some use a single goodness of fit indicator, where others rely on an interpretation of multiple indices (Crowson, 2020). The traditional, and most widely used approach is the chi-square value (CMIN) and a chi-square goodness of fit test (Crowson, 2020; Pirralha, 2020). A chi-square goodness of fit text analyzes if a model significantly varies from the model that perfectly fits the data (Crowson, 2020; Kline, 2016). The p-value associated is considered the significance level and if $p \leq .05$ then the null hypothesis of an exact fitting model is rejected. However, this approach is considered by some to be too permissive (e.g., Svetlina, Rutkowski, & Rutkowski, 2019) and it is in fact sensitive to sample size (Pirralha, 2020). Therefore, other indicators are often employed to buttress (or refute) Chi-square findings.

The CMIN/degrees of freedom (CMIN/DF), for example, is another statistic commonly utilized (Kline, 1998). Generally, it is accepted that CMIN/DF ≤ 3 is considered acceptable fit where CMIN/DF ≤ 5 is considered reasonable fit (Kline, 1998; Uedufy, 2022). Any value that exceeds five is considered poor fit (Kline, 1998; Uedufy, 2022). Another commonly used model fit statistic is the comparative fit index (CFI) (Kline, 2005; Uedufy, 2022). The CFI is another commonly used and reported statistic, and it also is one of the more robust indicators (Bentler, 1990; Pirralha, 2020). CFI scores range from 0 to 1, with 1 indicating a perfect fit (Crowson, 2020; Kline, 2005; Uedufy, 2022). It is generally accepted that a $CFI \geq 0.95$ is excellent fit, $CFI \geq 0.90$ is an acceptable fit, and anything lower is an unacceptable fit (Kline, 2005; Uedufy, 2022). Lastly, a third common fit statistic is the root mean square error of approximation (RMSEA) which measures the "difference between the observed covariance matrix per degree of freedom and the predicted covariance matrix" (Chen, 2007). The accepted rule here is that values greater than .1 are considered poor, values that fall between .08 and .1 are borderline, and values between .05 and .08 are acceptable (MacCallum et al., 1996). Values that fall below or are equal to .05 are considered excellent (MacCallum et al., 1996).

When assessing for measurement invariance using a CFA, it is expected that goodness of fit will decrease as more restrictions are imposed. As configural invariance should hold, if the original model was developed properly, it is assumed that the general model will show acceptable fit more most data (Crowson, 2022). However. as more rigid forms of invariance are tested, it is necessary to see whether or not the decrease in fit is statistically significant. This is accomplished by conducting a chi-square difference test and/or examining the change in the CFI between more and less restrictive models (Crowson, 2022). A chi-square difference test explores whether or not the newly fitted, more restrictive, model represents a significantly worse fit than

the previous model (assuming configural invariance) (Crowson, 2022). In other words, a chi-square difference test looks at whether or not the difference in model fit is due to the equality constraints imposed by metric and scalar invariance assessments (Crowson, 2022). To conduct a chi-square difference test, the chi-square statistic from the less restrictive model is subtracted from the chi-square statistic of the more restrictive model ($X^2diff = X^2more\ restrictive - X^2less\ restrictive$) (Crowson, 2022). Then, the critical p-value for the chi-square test is assessed to determine if the difference between models is statistically significant (Crowson, 2022). For example, Crowson, (2022) states, "…a chi-square difference test, which tests whether the model represents a significantly worse fit to the data than the previous model (assuming configural invariance). Since the model with the equality constraints (i.e. the metric invariance model) is nested within the configural model, we can test whether there is a statistically significant reduction in fit as a result of adding in the equality constraints" (slide 28).

An additional way to assess the change in model fit between less and more restrictive models is by assessing the change in CFI (Crowson, 2022). The typical criteria used to determine if the difference in model fit is significant is $\Delta CFI \leq .01$ (Cheung & Rensvold, 2002; Kline, 2016). To assess this difference, the CFI of the more restrictive model is subtracted from the CFI of the less restrictive model ($\Delta CFI = CFI_{more\ restrictive} - CFI_{less\ restrictive}$) (Crowson, 2022, Cheung & Rensvold, 2002; Klein, 2016). If the difference is significant, then evidence for metric or scalar invariance would be present. If it is insignificant, then evidence for metric or scalar non-invariance would be present. For the current study, the CFI and CMNI statistics were used to determine the presence of configural invariance, and the chi-square difference test and the CFI difference test were utilized to determine the presence of metric and/or scalar invariance.

Recruitment and Participants

Power Analysis

As an SEM framework was used for this study, a power analysis was conducted using Soper's (2016) online structural equation modeling (SEM) calculator. It was determined that a minimum sample size of 195 was necessary to detect an effect and a minimum sample size of 314 was necessary to determine model structure. The following values were used for such calculations: effect size = .3 (as suggested by Levant & McCurdy, 2018 and a generally accepted medium level of effect within psychological sciences), power = .85, and $\propto$ = .05, with eight latent variables or, the variables that cannot be directly observed (i.e. MRNI-R total Score) and 53 observed, or visible variables (i.e. number of total scale items). The goal N was 315 participants and 349 were recruited, however, 116 responses were deemed invalid and removed from data analysis due to incomplete responses or quick response time. Due to financial constraints, further data collection was not possible, leaving the final sample size at 233.

The current study was approved by the University of Wisconsin—Milwaukee's (UWM) Institutional Review Board. Participants were recruited through the crowdsourcing website Prolific Academic. Surveys were created using UWM's Qualtrics survey software, by entering the Male Role Norms Inventory-Revised (MRNI-R) and a brief demographic questionnaire (see Table 1 for demographic questions). The survey was administered to two sets of male-identifying participants, one group millennial age men and younger and one group Gen-X and above. This assured an adequate enough sample sizes for each group to assess measurement invariance. Potential participants were excluded from the study if they were under the age of 18, not proficient in English, did not self-identify as male, or have already completed the survey once.

Participants

Prolific Academic

Crowdsourcing websites such as Amazon Mechanical Turk (MTurk) or Prolific

Academic have become commonplace in academic research (Palan & Schitter, 2017). Such

resources have proven reliable, when compared to laboratory experiments, and have replicated

results from well-known laboratory experiments in both fields of economics and psychology

(Palan & Schitter, 2017). Although MTurk has been used in a plethora of recent works, there are

a few limitations, outlined well by Palan and Schitter (2017) that call it's use into question. First,

although MTurk advertises a large participant base of over 50,000 individual workers, it has been

found that the true size is much smaller (Palan & Schitter, 2017). For example, Stewart et al.,

2015 found that labs commissioned to work for MTurk were made up of around 7,300 persons,

with many of them working for several labs concurrently, significantly decreasing MTurks

proposed sample size. Additionally, the impact of such individuals partaking in hundreds, if not

thousands of online studies has not been adequately explored but is hypothesized to bias

responses and decrease participant novelty (Palan & Schitter, 2017). MTurk workers also have

many online chat rooms where information about surveys is discussed, further decreasing the

validity of participant reports. MTurk also does not control the work environment, meaning

participants may be distracted or engaging in other tasks while completing surveys, nor does it

outline standards for payment of workers, leaving room for unethical treatment of participants

(Palan & Schitter, 2017). Lastly, it is difficult for any crowdsourcing website to verify the

participant identities, meaning workers could lie about various demographics to earn more

money, compromising the true validity and generalizability of research results.

Prolific Academic, a more recently developed crowdsourcing website, was created with

quality research and researchers in mind. Similar to MTurk, well established laboratory studies

have been replicated using Prolific, supporting its ability to produce reliable results (Palan &

Schitter, 2017). It offers a more diverse sample, regarding multiple demographics (i.e. geographic location, ethnicity, naivety) and its participant base has been steadily growing. To assure participants are treated fairly, Prolific has set standards for payment based on time required to complete tasks and provides incentives for participants to provide high quality responses (Palan & Schitter, 2017). Similar to MTurk, participants are subject to rejection by researchers for poor or incomplete performance on tasks. This in turn decreases their acceptance score, which is what gains them access to paid tasks (Prolific Academic, 2014). However, to protect participants, such rejections can be overturned if deemed inappropriate (Palan & Schitter, 2017; Prolific Academic, 2014). Another advantage of Prolific Academic is the ability to "pre-screen" participants based on screening questions they have completed for prior tasks, allowing for more specified participant selection. Participants are given incentive to fill out additional questions, outside of their gender and age, to be considered for a wider range of opportunities, prior to beginning tasks, which helps assure accurate participant identities. Lastly, no survey can be completed twice by the same worker on Prolific, eliminating concerns of repeat responses (Palan & Schitter, 2017).

Based on the aforementioned information, quality participant base, and ease of access, Prolific Academic was chosen for the current study. The diversity of Prolific workers helped address the overuse of cisgender, heterosexual, White college undergraduate students in MRNI research, which is a frequent critique (Levant & McCurdy, 2018; Levant & Richmond, 2007). Additionally, the safeguards put in place allowed for valid and reliable responses, while also allowing researchers to reach a broad demographic in a time efficient manner.

Measures

Demographic Questionnaire

Demographics collected include gender, race, ethnicity, socioeconomic status, sexual

orientation, income, and veteran status to rule out any possible mediating or moderating effects

of identities outside of generational cohort (see Appendix for a full questionnaire). Although the

current study did not focus on veterans, it is important to collect such data due to the traditionally

masculine culture of the military (Connell, 1985) and how this may impact both generation of

men's responses to the MRNI-R.

Male Role Norms Inventory-Revised

The Male Role Norms Inventory-Revised (MRNI-R; Levant, et al., 2007; Levant et al.,

2010) is a self-report measure intended to assess endorsement of traditional masculine norms. It

consists of a total score as well as scores on seven subscales: Avoidance of Femininity, Fear and

Hatred of Homosexuals, Extreme Self-Reliance, Aggression, Dominance, Non-relational

Sexuality, and Restrictive Emotionality (Levant et al., 2007). Participants are asked to indicate

their level of agreement or disagreement on items using a 7-point Likert-type scale, where 1 =

strongly disagree and 7 = strongly agree. Levant et al., (2007) noted Cronbach's alpha's ranging

from .79 to .91 and .96 for the total score, suggesting high reliability for both subscale and total

score. Factor structure and construct validity were reassessed by Levant and colleagues (2010)

using exploratory factor analysis (EFA). Correlations between the MRNI-R and other

masculinity measures were also examined (Levant et al., 2010). As hypothesized, the EFA

confirmed the MRNI-R's model structure, reporting strong factor loadings for the subscales.

Additionally, the MRNI-R was significantly correlated with other masculinity measures such as

the Gender Role Conflict Scale (GRCS; O'Neil et al., 1986) suggesting strong construct validity.

The MRNI-R also showed strong discriminant validity, showing no correlations with other

measures including the Personal Attributes Questionnaire (Levant et al., 2010). Although there

are many measures of masculinity, the MRNI-R was chosen for its psychometric strength. As previously mentioned, the MRNI is the most psychometrically investigated measure out of the existing masculinity inventories (Levant et al., 2010) and also one of the most frequently used (Levant et al., 2010). For this reason, it is imperative that the MRNI exhibit strong metric invariance to assure truly generalizable findings.

Reliability and Validity

Reliability

Reliability is sample dependent and assesses whether an instrument measures what it intends to measure. When conducting research utilizing psychological measures, it is important examine internal consistency, or the consistency of participant responses within a group, to assure that sampling error is not behind the findings. Multiple methods assess internal consistency including split half tests and examinations of covariance such as Cronbach's Alpha (UCLA, 2020). Cronbach's Alpha assesses how closely related participant responses are within a group and is widely used in the field of counseling psychology. A high Cronbach's alpha indicates strong internal consistency, or strong likelihood that participant responses could be generalized to similar measures, however this does not mean the measure itself is unidimensional (UCLA, 2020). This is important to note as the MRNI-R used in this study is not a unidimensional measure. Reliability scores lie on a scale ranging from 0 to 1, with .9 indicating excellent reliability and .5 or less indicating unacceptable reliability. However, it is suggested that scores should lie above .7 to be used for analysis (UCLA, 2020). Prior to any factor analyses, reliability of MRNI-R total and sub scores was assessed for each group using Cronbach's alpha. Scores falling below a .7 were excluded from further analyses.

Validity

Measures of validity assess the accuracy of a measure. In this case, is the MRNI-R measuring general masculinity? Strong validity is hard to come by in measures such as the MRNI, due to differing definitions of constructs between individuals. Three types of validity are usually assessed, construct, content and criterion, with construct being considered the more important in examinations of metric invariance.

Construct Validity. Construct validity is evidence that a construct measures what it is supposed to measure. Studies have well examined the statistical evidence that the MRNI construct is distinct or unique from other measures, typically by comparing the correlation of total MRNI scores to different measures of gender or the gender role. Levant and Richmond (2007) looked at the correlation between the PAQ M scores (M meaning typical male personality traits) and total MRNI scores and found no correlation between the two, demonstrating strong discriminant validity ($r = .06$). To assess convergent validity, or similarity between one measure of a construct with similar measures of similar constructs, the MRNI was compared with two other gender measures, the Gender Role Conflict Scale-I (GRCS-I; O'Neil, Good, & Holmes, 1995) and the Masculine Gender Role Stress Scale (MGRSS: Eisler, 1995). Convergent validity was also supported, as the MRNI total score was significantly correlated with both the GRCS-I ($r = .52, p < .001$) and the MGRSS ($r = .52, p < .001$). With both discriminant and convergent validity established, overall construct validity of the MRNI could be concluded.

Content Validity. Content validity assures that a scale measures what it intends to measure, including that the scale is relevant to the target population (Boateng et al., 2018). For the current scale, this means that it measures the masculine experience. For content validity to be met, five conditions are suggested. First, the content's meaning has a general, societal, understanding. In this case the definition of masculinity, though possibly outdated, has been

51

proposed and accepted across multiple disciplines (Gerdes et al., 2018; Levant & Richmond, 2007). This construct of masculinity has also been evidenced across cultures and nationalities (Gerdes et al., 2018; Levant & Richmond, 2007). Second, a construct must be operationally defined, assuring it is distinct enough from previously accepted definitions. The construct of traditional masculinity ideology (TMI) measured by the MRNI has been operationally defined both upon initial development and revised when the measure was updated in 2007 (Levant et al., 1992; Levant et al., 2007). Although similar, it differs from prior gender role paradigms (e.g. Pleck, 1981, 1995; O'Neil et al.; 1986) and masculinity paradigms (e.g. Brannon & Juni, 1984). Where gender role strain focuses on the pressure or conflict resulting from mismatched gender norms and environments, the MRNI assesses endorsement of TMI, defined as "an individual's internalization of cultural belief systems and attitudes towards masculinity and men's roles" (Levant & Richmond, 2007, p. 131). This operational definition is also distinct from more recent inventories such as the Conformity to Male Role Norms Inventory (CMNI; Mahalik et al., 2003) that instead assesses not one's endorsement, but true engagement in TMI. The third standard needed for strong content validity is that the subscales (or domains) used in the scale must be relevant to the study of the construct. As previously discussed, the bifactor structure of the MRNI, with a general TMI factor and seven subfactors (scales), has been consistently supported (e.g. Levant et al., 2010), meaning that the subscales are relevant aspects of TMI. Fourth, "qualified judges agree that the domain has been adequately sampled based on consensus" (Boateng et al., 2018), as has been demonstrated by key researchers in the field continuing to utilize and enhance the MRNI. Lastly, "the response content must be reliably observed and evaluated" (Boateng et al., 2018) through the use of expert and target population judges taking the measure and reporting back, as has been consistently demonstrated since the MRNI's

development (see Gerdes et al., 2018 or Levant & Richmond, 2007). Although the MRNI-R has previously demonstrated acceptable content validity (Levant et al., 1992; Levant et al., 2007), and continues to do so, evidence of measurement invariance is somewhat absent from the literature. This study hopes to begin to fill that gap.

Criterion Validity. Criterion validity is the extent to which a measure can predict behaviors or performance in another area. For example, as high levels of traditional masculinity have been associated with higher levels of substance use (Vogel & Wester, 2012), one could hypothesize that those who score higher on the MRNI-R drink more than those who do not. A more general example would the GRE, which has been shown to predict graduate school performance (Glen, 2015). The social sciences struggle to meet acceptable criterion validity, as operationalizing social constructs proves difficult (Glen, 2015). However, this makes validity even more important to demonstrate, as sociological and psychological measures are inherently flawed. As previous research has shown that higher scores on the MRNI-R are correlated with things such as substance use, poor mental health outcomes, (Wester & Vogel, 2012) and avoidance of help-seeking (Wester et al., 2010) it is important to assure its applicability to all ages, as not to overgeneralize such associations. The current study hopes to provide more insight into the validity of the MRNI, as demonstrated invariance would further enhance the measure's criterion validity, where demonstrated non-invariance would suggest improvements to the measure need to be made so that it can remain relevant into the 21st century.

Chapter 4

RESULTS OF THE ANALYSIS

The purpose of the current study was to analyze for generational variance between older and younger men's scores on the Male Role Norms Inventory-Revised (MRNI-R; Levant et al.,

2007). The MRNI-R is an accepted measure within the field of men and masculinity and is

utilized frequently within men and masculinities research (Levant et al., 2013). It was generally

hypothesized that the model fit would not be adequate for both groups, or in other words, that

invariance would not be present between groups. Such results would suggest that the MRNI-R's

proposed construct of masculinity significantly differs between generational cohorts of men. The

research questions, again, are as follows:

1. Does the MRNI-R total score demonstrate measurement invariance between
   generational cohorts of men?

2. Do each of the MRNI-R subscales (Avoidance of Femininity, Aggression, Fear
   and Hatred of Homosexuals, Dominance, Non-Relational Attitudes Towards Sex,
   Extreme Self-Reliance, and Restrictive Emotionality) demonstrate measurement
   invariance between groups of men?

To answer these research questions a measurement invariance investigation was

conducted. Prior to any analysis, however, the collected data was cleaned to assure valid results.

Incomplete survey responses, repeat responses, invariant responses, and surveys that were

completed too quickly (under 5 min) were removed. The most common reason for data removal

was survey completion time. Due to Prolific Academic's researcher requirements, responses

under 5 minutes could not be returned or replaced, resulting in a large number of completed, but

unusable responses. Overall, 116 responses were deemed invalid data. The remaining data (N =

233) consisted of 119 men falling into the "Younger" category (18-40 years old) and 114 men in

the "Older" category (40+ years old). This total sample size, although not technically large

enough to fully detect model structure, did exceed the minimum sample size of 195 necessary to

detect an effect. However, due to the smaller sample size, it is possible that these results lack statistical power.

Participants & Demographics

Age Delineation

Using Prolific Academic's participant filtering capabilities, two groups of men were surveyed. One survey was published for men between the ages of 18 and 40, to encompass Millennials and younger, and another was published for men over the age of 40, to capture those in generation X and higher. Unfortunately, the specific age of each participant was not captured, and will be discussed further in the limitation section of this document.

Younger

The group of younger men was slightly larger than the group of older men (N = 119). Ages ranged from 18-40 (see Table 6 for more detailed breakdown). Of this group, 98 identified as heterosexual, 6 identified as homosexual, 12 identified as bisexual, 1 identified as pansexual, and 2 identified as asexual (see Table 1). Regarding racial identity, the sample was predominantly White (N = 71) with 11 identifying as Black, 14 identifying as Hispanic, 2 identifying as American Indian, 13 identifying as Native Hawaiian, 7 identifying as Asian, and 1 identifying as two or more races (see Table 2). Over 50 percent of this sample had obtained a bachelor's degree or higher (51.2%) and 63.8 percent earned an income above $40,000 per year (see Tables 3 and 4). Only one individual endorsed military background or .8 percent of the sample (see Table 5).

Older

The group of older men was slightly smaller than the group of younger men (N = 116). Ages ranged from 41-84 (see Table 12 for more detailed breakdown). Of this group, 98

identified as heterosexual, 10 identified as homosexual, and 6 identified as bisexual (see Table 7). None in this group identified as pansexual or asexual. Regarding racial identity, 98 identified as White, 4 identified as Black, 4 identified as Hispanic, 6 identified as Asian, and 2 identified as two or more races (see Table 8). Over 50 percent of this sample had obtained a bachelor's degree or higher (52.7%) and 64 percent earned an income above $40,000 per year (see Table 9). Of note in this sample is that 20.2 percent of the respondents endorsed a military background. This is quite different from the .8 percent endorsement of the younger group (see Table 5 and 11).

Analyses

For both Research question 1, as well as each of the seven subscales mentioned in Research Question 2, configural invariance was assessed first, thereby testing the overall model fits for each group. If the general model fit was sufficient, equality constraints were then imposed on the factor loadings or item intercepts, assessing metric or "weak" invariance, and scalar invariance analyses were conducted as needed. As mentioned prior, imposing restrictions on the models was expected to decrease the overall fit of the data to the model (Crowson, 2022). To show evidence of metric or scalar invariance, chi-square difference and CFI difference tests were conducted. If the resulting statistics were statistically significant, evidence of metric or scalar non-invariance was present (Crowson, 2022).

Reliability & Validity

As mentioned above, reliability is sample dependent and must be assessed prior to any statistical analyses (Lee, 2018). Cronbach's Alpha was evaluated for the total score and each individual subscale score. All the subscales, excluding Extreme Self-Reliance, Aggression, and Non-Relational Attitudes Towards Sex demonstrated excellent reliability ($\alpha \geq .9$). Those subscales that did not show excellent reliability still demonstrated good reliability ($\alpha \geq .8$). The

total score also demonstrated excellent reliability ($\alpha = .98$) (see Appendix D for all scores). A

Pearson Correlation test was then conducted to assure items were significantly correlated to the

total score. Each item was significantly correlated at the .01 alpha level to the total score (see

Appendix D, Figure 1). As the sample size was not quite large enough, based on the power

analysis (Soper, 2016), an exploratory factor analysis was also conducted. The Bartlett's test was

significant ($p = .000$), which indicates that the matrix is appropriate for factor analysis. The

Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO MSA) was also in the great range

(.958) (Crowson, 2019; Kaiser & Rice, 1974) which suggests the matrix was acceptable for

factoring (see Table 21). Additionally, all but two items of the MRNI-R (Levant et al., 2010)

loaded onto the general masculinity factor at an acceptable range ($\geq .5$). The two loadings that

did not meet criteria were likely impacted by content, meaning that the item itself was likely not

written well. These two items were numbers 24, "It is okay for a man to use any and all means to

'convince' a woman to have sex" (Levant et al., 2007, p. 97) and 39, "Men should get up to

investigate a strange noise in the house at night' (Levant et al., 2007, p. 98). It is possible that

these items' specific content, was not relatable and or possibly disavowed by men in the current

sample. As mentioned prior, the social and political climate has shifted significantly, meaning it

is possible the current sample no longer adheres to these two items. As all other factor loadings

fell within an acceptable range, despite the slightly small sample size, enough statistical power

was present. With both reliability and validity of the data demonstrated, statistical analyses were

then assumed.

As discussed, prior in this document, the MRNI demonstrates strong construct validity.

Discriminant validity, or statistical evidence that the MRNI construct is distinct or unique from

other measures, was previously established by comparing the correlation of total MRNI scores to

a different measure of gender, the Personal Attributes Scale (PAQ; Spence & Helmreich, 1978).

Levant and Richmond (2007) looked at the correlation between the PAQ M scores (M meaning

typical male personality traits) and total MRNI scores and found no correlation between the two,

demonstrating strong discriminant validity ($r = .06$). Convergent validity, or similarity between

one measure of a construct with similar measures of similar constructs, has also been previously

assessed. The MRNI was compared with two other gender measures, the Gender Role Conflict

Scale-I (GRCS-I; O'Neil, Good, & Holmes, 1995) and the Masculine Gender Role Stress Scale

(MGRSS: Eisler, 1995). Convergent validity was supported, as the MRNI total score was

significantly correlated with both the GRCS-I ($r = .52$, $p < .001$) and the MGRSS ($r = .52$, $p <$

.001). With both discriminant and convergent validity being previously established, overall

construct validity of the MRNI was concluded for the current study.

Results of Analysis

As discussed above, there are multiple statistics used to assess model fit. For the purpose

of the current study, the statistics utilized to assess the presence of configural invariance are the

Comparative Fit Index (CFI), the Chi-Square/degrees of freedom (CMIN/DF) and the Root

Mean Square Error of Approximation (RMSEA). The cutoff point used for the CFI was .95,

however, if acceptable model fit (i.e. .9) was present, further analyses were conducted to assure

model fit (West et al., 2012; Uedufy, 2022), the cutoff for the CMIN/DF was 5 (Kline, 1998) and

the cutoff used for RMSEA was $\geq$ .06 (MacCallum et al., 1996). If configural invariance held,

equality restraints were imposed on the factor loadings, and the model was re-fit to the data to

assess for metric invariance. Chi-square difference tests were used to determine if the decrease in

fit was statistically significant. If the decrease was significant, this provided evidence for metric

non-invariance (Crowson, 2022). The same steps were conducted for those scales that required scalar invariance testing.

Research Question 1: MRNI-R Total Score

The total score consists of the sum of all 51 MRNI-R items. The configural invariance test indicated poor model fit, CFI = .616, CMIN/DF = 2.809, RMSEA = .089 (see Table 22, Table 23, & Table 24). Although the CMIN/DF indicated acceptable fit, the CFI and RMSEA indicated poor model fit, which suggests that the two groups did not interpret the measure the same way and configural invariance is not present. Although this result may suggest that one of the two groups could fit into a single-factor model or that neither of them do, configural invariance is not fully present. The traditional or hegemonic definition of masculinity, that men are to be assertive, unwavering, emotionless; that they are to avoid the feminine and revel in the rough, has indeed evolved with time, meaning this masculine construct, as proposed by the MRNI, has not held throughout history, as evidenced by the ever-present variance in interpretation between older and younger men. Or, in other words, older and younger men interpret and enact masculinity significantly differently (Crowson, 2022), at least in terms of how the MRNI-R samples this construct.

Research Question 2: Subscale Model

This question involved an individual analyses of each MRNI-R subscale (Fear and Hatred of Homosexuals, Aggression, Avoidance of Femininity, Dominance, Self-Reliance, Non-Relational Attitudes Towards Sex, and Restricted Emotionality).

Fear and Hatred of Homosexuals: The first scale assessed was Fear and Hatred of Homosexuals (FHH) (see all models in Appendix E). The configural invariance test for the subscale indicated poor model fit, CFI = .869, CMIN/DF = 5.279, and RMSEA = .136,

suggesting the two age groups did not interpret this subscale the same way (see Table 25, Table 26, & Table 27). This means no further investigation was needed, as there was no evidence for measurement invariance and general model fit was poor between the two groups. In other words, the FHH subscale, as theorized by the MRNI, does not apply across generations, or is not interpreted the same between the groups of older and younger men.

Aggression: The configural invariance test for the subscale Aggression (AGG) indicated strong model fit, CFI =.964, CMIN/DF = 1.867, and RMSEA = .061, which suggests strong configural invariance, or that both groups interpreted the AGG construct similarly (see Table 28, Table 29, & Table 30). Although the RMSEA is slightly above the cutoff point, a value of .061 is not considered unacceptable (MacCallum et al., 1996) and when paired with two strong fit indices, shows model fit. This is to be expected, as the general model has proven good model fit in prior studies (Levant et al., 2010). As the configural invariance held, Metric, or weak invariance was tested. Constraints were imposed on the factor loadings and the model fit worsened slightly, CFI = .932, CMIN/DF = 2.366, and RMSEA = .077 (see Table 31 & Table 32). To determine if this was a significant drop, a chi-square difference test was conducted, X^2_{config} - X^2_{metric} = $X^2_{diff,}$ (80.454 - 52.276 = 28.178) and df_{metric}-$df_{configural}$ = df_{diff} (34 − 28 = 6). The resulting p-value was significant with a p = .000087 < .05 indicating significant variance between groups meaning that the two groups of men interpreted the construct differently, or that the model does not hold between groups (Crowson, 2020; Lee, 2018). This suggests that the construct AGG as defined by the MRNI does not fit between older and younger men.

Avoidance of Femininity: The configural invariance test for the subscale Avoidance of Femininity (AVF) indicated acceptable model fit, CFI = .912, CMIN/DF = 4.755, RMSEA = .127, again suggesting acceptable levels of configural invariance (see Table 33, Table 34, &

Table 35). Meaning, the general model demonstrated acceptable fit for both groups. Although the RMSEA statistic suggested poor model fit, the other two indices suggested acceptable model fit, therefore in the interest of exploration it was determined that metric invariance should be assessed (e.g., Crowson, 2020). Metric invariance was tested, imposing equality constraints on the factor loadings. The model fit worsened slightly, CFI = .908, CMIN/DF = 4.330, RMSEA = .111 and a chi-square difference test was conducted (X^2_{config} - X^2_{metric} = $X^2_{diff)}$, (203.520 – 190.206 = 13.314) and df_{metric}-$df_{configural =} df_{diff}$ (47 – 40 = 7) (see Table 36 & Table 37). The resulting p-value was not significant with a p = .064818 > .05, providing evidence for metric invariance. As both configural and metric invariance held, meaning the model still fit both groups with equality constrained factor loadings and intercepts, scalar invariance was also assessed. Equality constraints were imposed on the item intercepts and the model fit was evaluated. The resulting model improved slightly, CFI = .910, CMIN/DF = 3.839 (see Table 38 & Table 39). A Chi Square Difference test was conducted to assess whether the increase was significant, X^2_{config} - X^2_{metric} = X^2_{diff}, (207.283 - 203.520 = 3.763) and df_{metric}-$df_{configural =} df_{diff}$ (54 - 47= 7). The resulting p-value was not significant p = .806635 > .05. Despite this, the relative improvement of fit and demonstrated configural and metric invariance still suggested that in fact AVF subscale showed significant measurement invariance between groups (Crowson, 2022). In other words, AVF fit both groups in that the two groups interpreted this scale similarly. As this subscale is one of the only constructs that held across groups, it should be noted that the centrality of avoidance of femininity is still prevalent and valid for younger and older men today. This may suggest the strong societal importance placed on men avoiding the feminine, no matter the current definition.

Dominance: The configural invariance test for the Dominance (DOM) subscale indicated acceptable model fit, CFI = .921, CMIN/DF 4.807, RMSEA .128, suggesting acceptable levels of configural invariance, or similar initial construct interpretation (see Table 40, Table 41, & Table 42). As discussed before, even though the RMSEA statistic indicated poor fit, the other two indices indicated strong fit, and therefore, it was determined, configural invariance held (Crowson, 2020; Klein, 2015). This was, again, expected, as the general model fit held in previous studies (Crowson, 2020). Metric invariance was examined, imposing equality constraints on the factor loadings. The model fit worsened, CFI = .916, CMIN/DF = 4.330, RMSEA = .120 (see Table 43 & Table 44). To determine if this drop was significant, a chi-square difference test was conducted, X^2_{config} - X^2_{metric} = $X^2_{diff,}$ (147.227 - 134.602 = 12.625) and df_{metric}-$df_{configural}$ = df_{diff} (34 − 28 = 6). The resulting p-value = .049393 was significant. This suggests a statistically significant drop when equality constraints were imposed, providing evidence for variance between the two groups interpretation of the subscale construct. This suggests the presence of metric non-invariance or poor model fit between groups. In other words, the two age groups of men did not interpret this subscale the same way or the subscale DOM as outlined by the MRNI is not relevant across generations.

Self-Reliance: The CFA for the Self-Reliance (ESR) subscale indicated poor model fit, CFI = .894, CMIN/DF = 3.951, RMSEA = .113, suggesting the two age groups did not interpret this subscale the same way (see Table 45, Table 46, & Table 47). In other words, there was no evidence for invariance, or the general model fit was poor for both groups. This suggests that the ESR construct delineated by the MRNI is not applicable to these groups of older and younger men.

Non-Relational Attitudes Toward Sex: The configural invariance test for the Non-Relational Attitudes Towards Sex (NRAS) subscale indicated strong model fit, CFI = .966, CMIN/DF = 2.271, RMSEA = .074 (see Table 55, Table 56, & Table 57). Equality restraints were imposed on the factor loadings, and the model fit was again assessed. The model fit worsened slightly, CFI = .964, CMIN/DF = 2.042, RMSEA = .67 (see Table 58 & Table 59). The significance of this drop was calculated, X2metric = X2diff, (46.965 - 40.884 = 6.081) and dfmetric-dfconfigural = dfdiff (23 – 18 = 5). The resulting p-value = .298417 > .05 was not statistically significant, suggesting invariance between groups. As metric invariance held, scalar invariance was evaluated. Equality restraints were imposed on the item intercepts and the model fit was assessed. The model fit worsened slightly again, CFI = .952, CMIN/DF = 2.142, RMSEA = .07 (see Table 60 & Table 61). The significance of this drop was calculated X2metric = X2diff, (59.977 - 46.965 = 13.012) and dfmetric-dfconfigural = dfdiff (28 - 23 = 5). The resulting p-value = .023267 < .05, suggesting a significant decrease in fit. This suggests that the theory of NRAS as proposed by the MRNI is not applicable for both groups of older and younger men, or that non-invariance was present.

Restrictive Emotionality: The CFA for the Restrictive Emotionality (RE) subscale indicated acceptable model fit, CFI = .942, CMIN/DF = 2.571, RMSEA = .082 (see Table 48, Table 49, & Table 50). Therefore, metric invariance was tested by restricting the factor loadings to equal between groups. The model fit improved slightly, CFI = .947, CMIN/DF = 2.224, RMSEA = .064 (see Table 51 & Table 52). The significance of this increase was assessed X^2_{metric} = X^2_{diff}, (104.532 - 102.846 = 1.686) and $df_{metric}\text{-}df_{configural} = df_{diff}$ (47 – 40 = 7), resulting in a non-significant p-value of .975164. As this was an unusual result, the scalar invariance was also assessed by imposing equality constraints on the intercepts. The model fit again improved

slightly, CFI = .952. CMIN/DF = 1.959 (see Table 53 & Table 54). This result was also

statistically insignificant, p = .995946. As these p-values were statistically insignificant, the

results of the invariance study were inconclusive for the RE scale. It is possible that this scale

was misinterpreted by both groups of men or that the subscale was not an appropriate fit for

either generation of men. However, no concrete inferences can be made.

Demographics—Younger Group

Table 1
Sexual Orientation—Younger Group

	Frequency	Percentage
Heterosexual	98	82.4
Homosexual	6	5
Bisexual	12	10.1
Pansexual	1	.8
Asexual	2	1.7

Table 2
Racial Identity—Younger Group

	Frequency	Percentage
Black	11	9.2
White	71	59.7
Hispanic	14	11.8
American Indian	2	1.7
Native Hawaiian	13	10.9
Asian	7	5.9
More Than Two Races	1	.8

Table 3
Income—Younger Group

	Frequency	Percentage
Less than $15k	15	12.6
$15-40k	28	23.5
$40-60k	23	19.3
$60-80k	21	17.6
$80-100k	14	11.8
$100k+	18	15.1

Table 4
Education Completed—Younger Group

	Frequency	Percentage
No HS Diploma	11	9.2
High School Diploma/GED	71	59.7
Some College	14	11.8
Associate's	2	1.7
Bachelor's	13	10.9
Master's	7	5.9
Doctorate	1	.8

Table 5
Veteran Status—Younger Group

	Frequency	Percentage
Yes	1	.8
No	118	99.2

Table 6
Age—Younger Group

	Frequency	Percentage
18-24	30	25.2
25-34	60	50.4
35-40	29	24.4

Demographics—Older Group

Table 7
Sexual Orientation—Older Group

	Frequency	Percentage
Heterosexual	98	86
Homosexual	10	8.8
Bisexual	6	5.3

Table 8
Racial Identity—Older Group

	Frequency	Percentage
Black	4	3.5
White	98	86
Hispanic	4	3.5
Asian	6	5.3
More Than Two Races	2	1.8

Table 9
Income—Older Group

	Frequency	Percentage
Less than $15k	6	5.3
$15-40k	35	30.7
$40-60k	23	20.2
$60-80k	16	14
$80-100k	8	7
$100k+	26	22.8

Table 10
Education Completed—Older Group

	Frequency	Percentage
No HS Diploma	1	.9
High School Diploma/GED	16	14
Some College	22	19.3
Associate's	15	13.2
Bachelor's	41	36
Master's	18	15.8
Doctorate	1	.9

Table 11
Veteran Status—Older Group

	Frequency	Percentage
Yes	23	20.2
No	91	79.8

Table 12
Age—Older Group

	Frequency	Percentage
41-44	25	21.6
45-54	32	27.6
55-64	33	28.4
65-74	18	15.5
75-84	6	5.2

Table 13

Group Statistics T-Test

Subscale	Age	N	Mean	Std. Deviation	Std. Error Mean
AVT	Younger	119	2.4181	1.39711	.12807
	Older	114	2.9857	1.68123	.15746
FHH	Younger	119	1.7916	1.12774	.10338
	Older	114	2.3553	1.62732	.15241
ESR	Younger	119	3.4478	1.26660	.11611
	Older	114	3.6604	1.39394	.13055
AGG	Younger	119	3.3037	1.34264	.12308
	Older	114	3.4599	1.27049	.11899
DOM	Younger	119	2.2797	1.22840	.11261
	Older	114	2.4712	1.43058	.13399
NRAS	Younger	119	2.0630	1.05810	.09700
	Older	114	2.2544	1.18189	.11069
RE	Younger	119	2.1565	1.09605	.10047
	Older	114	2.5833	1.36107	.12748
TMS	Younger	119	2.4549	1.04388	.09569
	Older	114	2.8070	1.27924	.11981

Chapter 5

DISCUSSION

The goal of the current study was to explore the possible generational differences between older and younger men's interpretation of masculinity using the Male Role Norms Inventory-Revised (MRNI-R; Levant et al. 2007). According to historically socialized gender prescriptions, men are expected to be unemotional, stoic, aggressive, and independent. They are to act as providers, protectors, and never waver in the face of danger. Traditional Western masculinity has become characterized by the traits included in the Male Role Norms Inventory (Levant et al., 1992) including aggression, restricted emotionality, non-relational attitudes towards sex, avoidance of femininity, fear and hatred of homosexuals, dominance, and extreme self-reliance (Levant et al., 1992).

Younger generations of men have begun to push back against this traditional, or hegemonic definition of masculinity (Yeazel, 2015). For example, younger men are more likely to be stay at home fathers (Advisor Magazine, 2018), engage in egalitarian relationships (Yeazel, 2015), and accepting of fluid definitions of sexuality (Rimmer, 2015). These new norms directly contradict those measured by the MRNI, yet it is still the most commonly utilized masculinity measure in both research and practice (Levant et al., 2007). This poses a problem for the validity of results from studies and therapeutic practices that glean information from the MRNI, as if varying generations of men do not interpret the construct proposed by the MRNI the same way, then any results drawn would be invalid.

Although generational differences, specifically comparisons between older and younger generations, have been highly debated in the literature (Pew, 2018), a generational paradigm is fitting for the study of socialized masculinity, when using a social constructivist lens. Brooks and

Elder (2014) citing Felmange (2000) note, "[s]ocial construction emphasizes the role of socialization but also how social context, including historic, economic, political, linguistic, interpersonal, and psychological constructs, affect masculine identity" (p. 12). This theory encompasses more than historical events, but rather the climate created by layering contexts distinct to varying age groups, that impact the development of the masculine self. This understanding of context has been recently emphasized in the literature (e.g. Addis & Hoffman, 2020) as imperative to the true understanding and treatment within the psychology of men, however one could argue that even more than context is necessary to understand the generational underpinnings of masculinity. The millennial and generation z experiences have been markedly different than their predecessors (Green & McClelland, 2019; Pew, 2018). They live and grow in an era fueled by technology and vision, social change and political evolution. This Zeitgeist has created a new and confusing environment for men attempting to define and enact what they see as masculine. Although the hegemonic remains dominant for many, it is possible that millennial and Gen Z men are different than their fathers, their grandfathers (Green & McClelland, 2019) and if this is the case, the construct measured by the MRNI will not hold for them, threatening future masculinity research.

The MRNI-R aims to measure the construct of masculinity, first through the use of a total score calculated by summing the responses to all 51 items. This construct was based in that of hegemonic, or traditional masculine norms, including that men should be aggressive, assertive, brawny, and unemotional (Levant et al., 2007; Levant et al., 2010). As initially hypothesized, and demonstrated by the previously discussed data, the total score showed poor fit between groups at the configural level. Meaning, with no equality constraints imposed, Levant's (2007) conceptualization did not fit for the varying generations of men (Lee, 2018; Pirralha, 2020). In

70

other words, the formerly accepted definition of masculinity as being strong, stoic, unemotional, being the provider, protector; the prideful commanding presence that simultaneously commands respect, is no longer relevant for younger generations of men. Recent social and political practices, media influence, and evolving education and research have remolded what is generally acceptable and expected for men. For example, the idea of male emotional suppression, has more recently been denounced, based on the negative mental health implications found (Addis & Hoffman, 2020; Berger et al., 2005; Vogel & Wester, 2014). In fact, there is a push, both within media, and psychological spaces, to allow men to be more openly vulnerable, aiding not only their personal well-being, but the well-being of their partners and relationships (Addis & Hoffman, 2020).

Studies have also found that younger generations engage in more egalitarian relationships (Pew, 2018) meaning that the role of the man as the "sole provider" has become less prevalent, and less expected. In 2019 it was found that over half of American households were dual income (DeMarco, 2019) and with current political and economic atmospheres shifting, it is likely that this number will continue to rise. Not only are expectations of men and maleness shifting, behaviors that were once praised as being tough and rugged, are now being described as harmful and destructive (Addis & Hoffman, 2020). The current findings speak to this shifting paradigm. suggesting that the previously accepted MRNI-R definitions of masculinity may be irrelevant for the coming generation's definition and understanding of maleness.

Subscales

The subscales of Dominance (DOM), Fear and Hatred of Homosexuals (FHH), Aggression (AGG), Non-Relational Attitudes Towards Sex (NRAS) and Extreme Self-Reliance (ESR), did not prove invariant. In fact, the two groups of men interpreted each subscale

significantly differently. However, at the same time, the subscale of Avoidance of Femininity (AVF) held between groups while the subscale Restrictive Emotionality (RE) was inconclusive.

Subscale One: Dominance (DOM)

As hypothesized, the subscale DOM significantly differed between groups, suggesting the two generational groups had different interpretations, represented by the lacking metric invariance (Lee, 2018; Pirralha, 2020). When factor loadings were held equal, the model significantly worsened. Like the Aggression subscale, the trait labeled "dominance" has come under recent scrutiny. Men who enact such traits have been deemed "toxic", rather than in control as they were previously seen (Addis & Hoffman, 2020). Additionally, as previously discussed, younger generations present strong support for social movements of global acceptance, rather than a singular dominant power (Pew, 2018). Such pushbacks and ideological shifts have likely led to younger generations straying away from typical dominant behaviors.

Subscale Two: Fear and Hatred of Homosexuals (FHH)

The current data suggests that the model proposed by the MRNI for the subscale FHH does not fit for both groups of men. More specifically, there was no evidence for configural invariance, or the general model fit was poor for both groups of men. This differs from the subscale's original validation and proceeding revisions to the subscale (e.g. Levant et al., 2007). The proposition by the MRNI that to be masculine, one must avoid, distain, or flat out hate anything wavering on homosexual, seems to no longer apply to younger generations of men. As previously discussed, it is possible that recent shifts in societal acceptance of a non-binary understanding of sexual orientation and gender expression have changed the way millennial and younger men view the acceptability of non-heterosexuality (Accel + Qualtrics, 2016; Pew, 2018). This shift has also led to a less heteronormative narrative, which may influence how

younger generations express and interpret the meaning of their own sexuality. It is also possible, that due to the recent political push against derogatory language surrounding gender minorities, that younger men felt the obligation to answer the subscale questions in what has become considered a "politically correct" or socially acceptable nature. However, exploration of this is outside of the scope of the current study.

Subscale Three: Aggression (AGG)

The data suggests that the AGG model proposed by the MRNI was not defined similarly by both groups of men, or they interpreted the construct differently. This was evidenced by lacking metric invariance between groups. In other words, when equality constraints were imposed on the factor loadings, the model fit significantly worsened. Current views of masculinity as "toxic" or harmful have recently led to societal pushback against the stereotypical "aggressive" male (Addis & Hoffman, 2020). Media attention and new understandings of male expression and coping styles have also resulted in more acceptance for men demonstrating emotions and tender actions (Green & McClelland, 2019). For example, popular television programming, social media outlets, and media coverage broadcast men engaging in expression deemed outside the generally accepted male role including wearing more "feminine" attire or outwardly demonstrating softer emotions (Green & McClelland, 2019). Such new expectations and enactments have likely impacted the way younger generations view and accept acts of aggression as part of their masculine self. It is possible that younger men do not view aggression as masculine at all, whereas older generations were expected to be aggressive possibly making more nurturing emotions feel contradictory to their internalized masculine-self.

Subscale Four: Non-Relational Attitudes Towards Sex (NRAS)

In line with our stated hypothesis, the NRAS subscale showed significant variance, at the scalar level, between groups (Crowson, 2020; Pirralha, 2020). This suggests that the construct outlined by the MRNI, labeled NRAS did not hold across generations, or that the two generations did not interpret the subscale in the same manner. Due to recent societal emphasis on consent and active communication around intimacy, it is likely that younger generations have less rigid and gendered expectations surrounding sex. Education around safe and consensual sex practices have been further integrated into educational curriculums. Studies have shown that greater education around sex leads to less instances of nonconsensual sex and impulsive sexual acts (SIECUS, 2020). It is possible that such educational programs and societal practices and movements such as the #MeToo Movement have led to a more nuanced relationship with sex for men than previously espoused.

Subscale Five: Extreme Self-Reliance (ESR)

As hypothesized, the ESR subscale demonstrated poor model fit, or lacking configural invariance. This suggests that the construct outlined by the MRNI did not hold between generations of men. Like the subscales AGG and DOM, recent societal shifts have criticized extreme independence and lack of help seeking among men, labeling this practice as unhealthy and harmful (Addis & Hoffman, 2020; Green & McClelland, 2019). It is likely that younger generations of men have addressed or disavowed such behaviors due to this pushback. It is also possible that increased levels of societal acceptance of help-seeking and community programming geared towards aiding men (i.e. male focused self-help or therapy groups) have normalized this behavior for younger generations of men. Additionally, younger men have likely seen the detriment this set of behaviors had on elder friends and family (see Berger et al., 2005 and Wester & Vogel, 2012), leading to introspection and conscious change.

Subscale Six: Avoidance of Femininity (AVF)

Contrary to the stated hypotheses, the AVF scale proved invariant between groups. This suggests that both younger and older generations interpret this subscale similarly, or that the originally normed construct still holds with younger generations. It is possible that despite the version of masculinity evolving, the desire to be distinctly masculine, whatever that may be, is still strongly instilled in all men. For example, although it has become more socially acceptable for men to express emotions (Green & McClelland, 2019), certain emotional expression may be seen as strongly female. Although a man may express sadness, it may still be seen as weak or feminine to cry openly. Another example is occupation. Although men are being represented more in stereotypically feminine occupations such as nursing, they tend to lean towards more "masculine" subfields, such as trauma, flight, or intensive care practices.

Subscale Seven: Restrictive Emotionality (RE)

Contrary to the stated hypothesis, the analysis of this subscale was inconclusive, as the model fit increased with equality constraints, however this increase was not statistically significant. Similar to other subscales, it was theorized that due to recent societal shifts and more societal acceptance and endorsement of men freely expressing emotion, this subscale would not hold between generations of men. However, the results did not indicate this to be so. It is possible that the sample size was not large enough to produce a significant result and it is also possible that the groups misinterpreted the meaning of the items on the subscale, leading to an inconclusive result.

Limitations

There were a few limitations to the current study. First, each of the two research questions were analyzed independently rather than in true bifactorial fashion. Research Question

1 consisted of all 51 items loading on total score analyzed with a multi-group CFA, whereas Research Question 2 explored each of the seven subscales (also with their own independent multigroup CFA). Ideally, a complete bifactor model would have been constructed and tested (see Geng, Ni, Wang, Fan, & Qian, 2022 as an unrelated example) using SEM to conduct a single multigroup CFA. This approach ended up, however, outside the scope of this research. Future research should consider a more integrated approach, which would allow for a full and complete examination of Levant's (2007) proposed bifactorial model.

A second limitation was the smaller than ideal sample size. However, small sample size does automatically impact the statistical power of this research in a negative fashion. As noted above, while a sample of 314 was deemed necessary to determine model structure an N of only 195 was needed to detect an effect (Soper, 2016). Although the sample surpassed the threshold of 195, not reaching 314 may have undermined this study's ability to detect a true effect. However, as the exploratory factor analysis showed strong factor loadings and the Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO MSA) was in the remarkable range (.958) (Crowson, 2019; Kaiser & Rice, 1974) which suggests the MRNI-R matrix was acceptable for factoring. Or, in other words, the loadings were strong enough to demonstrate statistical power and were appropriate for conducting a confirmatory factor analysis (i.e. measurement invariance study). The two items (item number 39 and number 24) that demonstrated less than acceptable factor loadings, were likely limited by their specific content. As discussed previously, due to the shifting social and political climate, it is likely few men would openly endorse sexual coercion (item 24) or feel the need to put themselves in danger to inspect a loud noise at night (item 39) (Levant et al., 2007). Though, on a similar note, social desirability bias, or the answering of questions in what is considered an "acceptable" manner versus being fully transparent, was not

assessed. It is possible that some participants were answering questions as they believed they "should" rather than representing their true beliefs.

Another limitation of this study was the online nature. Full validity of responses cannot be guaranteed in a virtual setting. For example, it is possible that individuals obtained multiple Prolific accounts to double take surveys or lied on the demographic questionnaire. However, online research is the most cost effective and efficient way to reach large and diverse samples of participants. Finally, despite Prolific Academic's diverse sampling pool, the sample of this study was predominantly White and Heterosexual. This does limit the applicability of these results to other populations or those who hold marginalized identities. Lastly, not gathering participant specific age was a limitation. Although age-range was gathered, the true age of each participant was not gathered. This could have offered more insight into age-related differences, however, an exploration of that nature was outside of the scope of the current study.

Implications

The current study provided evidence against future use of the MRNI, specifically with younger generations of men. Although older male generations may still espouse traditional masculine beliefs, it could be argued that the measure should be avoided with all groups. This variance suggests that survey itself is not a valid measure of masculinity, or that the construct of masculinity is too nuanced to be measured by a pencil and paper self-report. Rather, masculinity appears to be rooted in specific points in time, and as such exists in the eye of the beholder, or the man himself. Such rigid, stereotypical, boxes do not allow for the personal masculine persona to be expressed, rather it locks men into such a tight bind, it does more harm than good. Especially as the measure has demonstrated invalid results, for the mental well-being and best

research and therapeutic practice, it is suggested that use of the MRNI be avoided or that results using the measure be taken lightly.

Future Directions

Previous studies have explored various types of measurement invariance between groups using the MRNI. This was the first to explore a generational difference between interpretations of masculinity. As significant differences were found between groups, it is suggested that further examination of the statistical properties of the MRNI be explored or re-analyzed. Specifically, the subscales of AVF and RE should be re-examined using a larger sample size or possibly incorporating a qualitative aspect with the qualitative nature of this full study to explore the nuances of men's answers to various items and various subscales. It is important for researchers and practitioners to stay privy to the ever-changing nature of social and political expectations surrounding gender and gender expression, taking the results of self-report, Likert-Type surveys with a grain of salt, as masculinity is not a one size fits all, nor is it meant to be.

Although outside the scope of the current study, the impact of varying demographics or identities should be explored in relation to these results. As our sample was predominantly White and Heterosexual, it is possible that these majority identities impacted the results. A more diverse sample should be collected, and invariance testing should be reconducted to assess whether the same invariance holds for other folks. The younger group of men had a more diverse set of identities around sexual orientation compared to the older group of men. It is possible that this group adhered to a less rigid version of masculinity due to these identities, rather than their age or generational status. Further investigations of this nuance should be explored, especially in relation to the use of masculinity inventories. Of note is also the difference in Veteran status between age groups. The older group of men had a significantly larger sampling of Veterans. It is

possible that this identity impacted the way these men responded to the MRNI-R in a statistically significant manner. These nuances should be further analyzed, as they may impact the results.

CONCLUSION

The aim of the current study was to explore possible variance between generational understandings of masculinity. Based on the aforementioned findings, it is suggested that the MRNI, as it currently stands, does not fit for newer generations of men. Furthermore, while not every subscale was found to be significantly variant five (AGG, NRAS, DOM, ESR, and FHH) of seven subscales were. When considered together, this provides evidence for the re-evaluation of this measure and its applicability in current research. If younger men are not enacting or defining masculinity the same as their predecessors, then the MRNI, in all its variations, is irrelevant. Indeed, recent social, political, and theoretical shifts have led to more nuanced understandings of gender and gender roles (e.g. Addis & Hoffman, 2020), suggesting that the previously normed MRNI may not be relevant to younger males.

It was hypothesized that significant variance would exist between older and younger groups of men on the individual subscales along with the total score. In other words, it was hypothesized that the model would not fit for both groups of men or, that varying generations define the construct of masculinity significantly differently. This overarchingly proved to be the case. The subscales of Dominance, Fear and Hatred of Homosexuals, Aggression, Non-Relational Attitudes Towards Sex, and Extreme Self-Reliance all demonstrated significant variance. These results suggest that the validity of the measure is poor between younger and older groups of men, or that these generations define the masculine disparately. Contrary to the stated hypotheses, the subscales of Avoidance of Femininity and Restrictive Emotionality did not demonstrate significant variance. It is possible that no matter the societal definition of

79

masculinity, it is ingrained in men to avoid the feminine, whatever that may be. However, due to the sample size and virtual nature of the study, further exploration should be considered.

Gender has always and may forever be a debated construct. However, what seems to remain constant is that what is considered masculine or feminine continues to evolve. At this point in time, the traditional, unemotional, unwavering, infallible male norm does not seem to fit nor serve current generations of men. Roles that were once deemed essential are now seen as archaic, un-useful, even harmful and painful. Although there may always be a debate on what masculinity is, was, or should be, or whether or not the Millennial way is adaptive, progressive or lazy and entitled, Green and McClelland (2019) seem to say it best, "Millennials have consistently used their voices, now that they are adults, to challenge the traditional male role" (p. 10) and their challenging has proved significant enough to call measures like the MRNI into question.

References

Accel and Qualtrics, (2016). *The millennial study.* https://www.qualtrics.com/millennials/

Addis, M.E. & Hoffman, (2020). *The psychology of men in context.* Routledge.

Addis, M.E. & Mahalik, J. (2003). Men, masculinity, and the contexts of help seeking. *American Psychologist, 58,* 5-14. DOI: 10.1037/0003-066X.58.1.5

Addis, M. E., Mansfield, A. K., & Syzdek, M. R. (2010). Is "masculinity" a problem?: Framing the effects of gendered social learning in men. *Psychology of Men & Masculinity, 11*(2), 77–90. https://doi.org/10.1037/a0018602.

Advisor Magazine. (2018). *Millennial men feel the pressure to "Be it all".* https://www.lifehealth.com/millennial-men-feel-pressure/

Anderson, E. (2009). Inclusive masculinity: The changing nature of masculinities. *Sociology of Sport Journal, 27,* 319-322.

American Psychological Association, Boys and Men Guidelines Group. (2018). *APA guidelines for psychological practice with boys and men.* Retrieved from https://www.apa.org/about/policy/boys-men-practice-guidelines.pdf

Aronson, J., Burgess, D., Phelan, S., Juarez, L. (2013). Unhealthy interactions: The role of stereotype threat in health disparities. *American Journal of Public Health, 103*(1), 50-56.

Ayres, M. (2017). Easy [Song]. On *Drive slow.* Dixon Court Records.

Bem, S.L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology, 42,* 155-162. DOI: 10.1037/h0036215

Bialosiewicz, S., Murphy, K., & Berry, T. (2013). *An introduction to measurement invariance testing: Resource packet for participants.* Retrieved from

http://comm.eval.org/HigherLogic/System/DownloadDocumentFile.ashx?DocumentFile
Key=63758fed-a490-43f2-8862-2de0217a08b8

Beatles, (1970). Let it be [Song]. On *Let it be.* Apple Records.

Berger, Levant, McMillan, Keleher, & Sellers, (2005). Impact of gender role conflict, traditional
masculinity ideology, alexithymia, and age on men's attitudes toward psychological help
seeking. *Psychology of Men and Masculinity, 6,* 73-78. DOI: 10.1037/1524-9220.6.1.73

Boateng, G.O., Neilands, T.B., Frongillo, E.A., Melgad-Quiñonez, H.R., Young, S.L. (2018).
Best practices for developing and validating scales for health, social, and behavioral
research: A Primer. *Frontiers in Public Health.* DOI: 10.3389/fpubh.2018.00149

Bonney, J. F., Kelley, M. L., & Levant, R. F. (1999). A model of fathers' behavioral involvement
in childcare. *Journal of Family Psychology, 13*, 401-415.

Brannon, R., & Juni, S. (1984). A scale for measuring attitudes about masculinity. *Psychological
Documents, 14*(1). (University Microfilms No. 2612).

Brown, J. M. (2002). *Statistical analysis of a measure of adolescent gender ideology, the Male
Role Norms Inventory-Adolescent: A Pilot Study*. Unpublished doctoral directed study,
Nova Southeastern University

Brooks, G.R. & Elder, W.B. (2016). History and future of the psychology of men and
masculinities. In J.R Wong & S.R. Wester's *APA Handbook of Men and Masculinities.*

Chu, J.Y., Proche, M.V., & Tolman, D.L. (2005). The adolescent masculinity ideology in
relationships scale: Development and validation of a new measure for boys. *Men and
Masculinities, 8,* 93-115. http://dx.doi.org/10.1177/1097184X03257453

Cieciuch, J. & Davidov, E. (2015). Establishing measurement invariance across online and offline samples: A tutorial with the software packages AMOS and MPLUS. *Methodological Problems in Psychological Research.* 83-99.

Connell, R. W. (1995). *Masculinities*. Berkeley, CA: University of California Press.

Cope, C. (2004). Son's gonna rise [Song]. On *The clarence greenwood recordings.* RCA Records.

Costanza, D. (2018). Can we please stop talking about generations as if they are a thing? Millennials are not all narcissists and boomers are not inherently selfish. The research on generations is flawed. *Slate,* retrieved from https://slate.com/technology/2018/04/the-evidence-behind-generations-is-lacking.html

Crosby, Stills, Nash, & Young. (1970). Almost cut my hair [Song]. On *Déjà vu.* Atlantic.

Crowson, H. M. (2020). Testing for measurement invariance using AMOS. Downloaded from

https://drive.google.com/file/d/1GVi5dqRiScVdxJdJ27LGpP_x_Aixm2o0/

Cuthbert, A. (2015). Current and possible future directions in masculinity ideology research. *Psychology of Men and Masculinity, 16*(2), 134-136. http://dx.doi.org/10.1037/a0038998

Eagly, A.H., Nater, C., Miller, D.I., Kaufmann, M. & Sczesny, S. (2020). Gender stereotypes have changed: A cross-temporal meta-analysis of U.S. public opinion polls from 1946-2018. *American Psychologist, 75*(3), 301-315. http://dx.doi.org/10.1037/amp0000494.

Eisler, R. M. (1995). The relationship between masculine gender role stress and men's health risk: The validation of a construct. In R. F. Levant & W. S. Pollack (Eds.), *A new psychology of men* (pp. 207–225). New York, NY: Basic Books.

Furstenberg, Frank (2017): The Use and Abuse of Millennials as an Analytic Category. *Council*

on Contemporary Families. Retrieved from https://contemporaryfamilies.org/8-
furstenberg-millennials-analytic-category/

Faiyaz, B. (2022). Angel [Song]. On *Wasteland.* Dreamville Records.

Fischer, A. R., & Good, G. E. (1997). Men and psychotherapy: An investigation of alexithymia,
intimacy and masculine gender roles. *Psychotherapy, 34*, 160-170.

Gale, S. R. (1996). *Male role norm endorsement and acquaintance sexual aggression among
college students*. Paper presented at the annual meeting of the American Psychological
As- sociation, Toronto, Ontario, Canada.

Gerdes, Z.T., Alto, K.M., Jadaszewski, S., D'Auria, F., & Levant, R.F. (2018). A content
analysis of research on masculinity ideologies using all forms of the male role norms
inventory (MRNI). *Psychology of Men and Masculinity, 19*(4), 584-599.
http://dx.doi.org/10.1037/men0000134.

Gerdes, Z.T. & Levant, R.F. (2018). Complex relationships among masculine norms and
health/well-being outcomes: Correlation patterns of the conformity to masculine norms
inventory subscales. *Mental Health & Wellbeing, 12*(2), 229-240. DOI:
10.1177/1557988317745910

Gilmore, D. (1990). *Manhood in the making: Cultural concepts of masculinity*. New Haven, CT:
Yale University Press.

Glen, S. (2015, July 15). Criterion validity: Definition, types of validity. *Statistics How To:
Statistics for the Rest of Us.* https://www.statisticshowto.com/criterion-validity/

Green, A. & McClelland, C. (2019). Male gender expression conflict between baby boomers and
millennials. *Pepperdine Journal of Communication Research, 7*(6), 6-19.
https://digitalcommons.pepperdine.edu/pjcr/vol7/iss1/6

Hayashi, S. (1999). Masculinity and mental health: The relationship between men's attitudes

toward male gender roles and mental health indices: Anxiety, depression, affect

regulation, and self-esteem. *JASS Proceedings, 11*(1), 2-13.

Heesacker, M., & Levant, R. F. (2001). Cross-lagged panel data. Unpublished raw data,

University of Florida.

Hsu, K., & Iwamoto, D. K. (2014). Testing for measurement invariance in the Conformity to

Masculine Norms-46 across White and Asian American college men: Development and

validity of the CMNI-29. *Psychology of Men & Masculinity, 15,* 397–406.

http://dx.doi.org/10.1037/a0034548.

Lee, S.T.H. (2018). Testing for measurement invariance: Does your measure mean the same

thing for different participants? *Association for Psychological Science.* Retrieved from

https://www.psychologicalscience.org/observer/testing-for-measurement-invariance

Levant, R., Hirsch, L., Celentano, E., Cozza, T., Hill, S., MacEachern, M., Marty, N., &

Schnedeker, J. (1992). The male role: An investigation of contemporary norms. *Journal

of Mental Health Counseling, 14,* 325-337.

Levant, R.F., Richmond, K, Majors, R. G., Inclan, J. E., Rosello, J. M., Rowan, G., & Sellers, A.

(2003). A multicultural investigation of masculinity ideology and

alexithymia. *Psychology of Men and Masculinity, 4,* 91-99. DOI: 10.1037/1524-

9220.4.2.91.

Levant, R. F., Good, G. E., Cook, S. W., O'Neil, J. M., Smalley, K. B., Owen, K., & Richmond,

K. (2006). The Normative Male Alexithymia Scale: Measurement of a gender-linked

syndrome. *Psychology of Men and Masculinity, 7,* 212–224.

http://dx.doi.org/10.1037/1524-9220.7.4 .212

Levant, R. F., Smalley, K. B., Aupont, M., House, A., Richmond, K., &. Noronha, D. (2007). Initial validation of the Male Role Norms Inventory- Revised. *Journal of Men's Studies, 15,* 85-100. DOI: 10.3149/jms.1501.83

Levant, R. F., Rogers, B. K., Cruickshank, B., Kurtz, B. A., Rankin, T. J., Williams, C. M., Colbow, A, (2012). Exploratory factor analysis and construct validity of the Male Role Norms Inventory Adolescent-revised (MRNI-A-r). *Psychology of Men and Masculinity, 13*, 354-366. DOI: 10.1037/a0029102

Levant, R. F., Hall, R. J., Weigold, I., & McCurdy, E. R. (2015). Construct distinctiveness and variance composition of multi-dimensional instruments: Three short-form masculinity Measures. *Journal of Counseling Psychology, 62*, 488-502. http://dx.doi.org/10.1037/cou0000092

Levant, R., & Majors, R. (1997). An investigation into variations in the construction of the male gender role among young African-American and European-American women and men. *Journal of Gender, Culture and Health, 2,* 33-43.

Levant, R.F., Majors, R.G., & Kelly, M.L. (1997). Masculinity ideology among young African American and European American women and men in different regions of the United States. *Cultural Diversity and Mental Health, 4*(3), 227-236.

Levant, R.F., Hall, & Rankin (2013). Male role norms inventory short form. *Journal of Counseling Psychology, 60*(2), 228-238. DOI: 10.1037/a0031545.

Levant, R.F. & McCurdy, E.R., (2017). Toward diversifying research participants: Measurement invariance of the male role norms inventory-short form (MRNI-SF) across recruitment method. *Psychology of Men and Masculinity, 19*(4), 531-539. http://dx.doi.org/10.1037/men0000138

Levant, R.F., Rankin, T.J., Williams, C.M., Hasan, N.T., & Smally, K.B. (2010). Evaluation of
the factor structure and construct validity of scores on the male role norms inventory-
revised (MRNI-R). *Psychology of Men and Masculinity, 11,* 25-37.
http://dx.doi.org/10.1037/a0017637

Levant, R.F. & Richmond, K.R., (2007). A review of research on masculinity ideologies using
the male role norms inventory. *The Journal of Men's Studies, 15*(2), 130-146. DOI:
10.3149/jms.1502.130

Levant, R.F., Stefanov, D.G., Rankin, T.J., Halter, M.J., Mellinger, C., & Williams, C.M. (2013).
Moderated path analysis of the relationships between masculinity and men's attitudes
toward seeking psychological help. *Journal of Counseling Psychology, 60*, 392-406.
http://dx.doi.org/10.1037/a0033014

Levant, R., Wu., R., & Fischer, J. (1996). Masculinity ideology: A comparison between U. S.
and Chinese young men and women. *Journal of Gender, Culture and Health, 1*, 207-220.

Mathison, D.L. (2001). Sex differences in the perception of assertiveness among female
managers. *The Journal of Social Psychology, 126*(5), 599-606. DOI: 10.1080/00224545

Massoth, N., Broderick, C., Festa, J., & Montello, N. (1996). *A comparison of gay and head
injured males with normative data on the Male Role Norms Inventory*. Unpublished
manuscript, Fairleigh Dickinson University.

Maxton, R. A. (1994). How do men in mid-life conceptualize masculinity, and how do these
conceptualizations relate to intimacy? (Doctoral Dissertation, Boston University).
Dissertation Abstracts International, 54, B4432

McDermott, R. C., Levant, R. F., Hammer, J., Hall, R., McKelvey, D., & Jones, Z. (2017).
Further examination of the factor structure of the Male Role Norms Inventory-Short

Form (MRNI-SF): Measurement considerations for women, men of color, and gay

men. *Journal of Counseling Psychology, 64(6)*, 724-738.

http://dx.doi.org/10.1037/cou0000225

McDermott, R. C., Levant, R. F., Hammer, J., Borgogna, N.C., & KcKelvey, D.K. (2019).

Development and validation of a 5-Item Male Role Norms Inventory using bifactor

modeling. *Psychology of Men and Masculinities, 20,* (4), 467-477.

http://dx.doi.org/10.1037/men0000178

Mcgraw, S. L. (2001). Masculinity ideologies, men's relationship behavior, and relationship

satisfaction in heterosexual couple relationships. *Dissertation Abstracts International,*

62(3), 1588B.

Merriam-Webster. (n.d.). Socialization. In *Meriam-Webster.com dictionary*.

https://www.merriam-webster.com/dictionary/socialization.

Merriam-Webster. (n.d.). Context. In *Meriam-Webster.com dictionary.*

 https://www.merriam-webster.com/dictionary/context.

Milfont, T. L., & Fischer, R. (2010). Testing measurement invariance across groups:

Applications in cross-cultural research. *International Journal of Psychological Research*,

3, 111–121. DOI: 10.21500/20112084.857

Moore, B. (2010, May 10). The worst generation? *New York Post.*

https://nypost.com/2010/05/10/the-worst-generation/

Morton, P.J. (2019). Don't let go [Song]. On *Paul.* Morton Records.

Norton, G., (1996). Long December [Song]. On *Recovering the satellites.* DGC Geffen.

Olff, M., Langeland, W., Draijer, N., & Gersons, B.P.R. (2007). Gender differences in

posttraumatic stress disorder. *Psychological Bulletin, 133*(2), 183-204. DOI:

10.1037/0033-2909.133.2.183

O'Neil, J.M., Helms, B.J., Gable, R.K., David, L., Wrightsman, L.S. (1986). Gender-role

conflict: College men's fear of femininity. *Sex Roles, 14*(5/6), 335-350.

O'Neil, J.M. (2008). Summarizing 25 years of research on men's gender role conflict using the

gender role conflict scale: New research paradigms and clinical implications. *The*

Counseling Psychologist, 36, 358-445. DOI: 10.1177/0011000008317057

Palan, S. & Schitter, C. (2017). Prolific.ac—A subject pool for online experiments. *Journal of*

Behavioral and Experimental Finance, 17, 22-27.

https://doi.org/10.1016/j.jbef2017.12.004.

Parent, M. C., & Smiler, A. P. (2012, March 26). Metric Invariance of the Conformity to

Masculine Norms Inventory-46 Among Women and Men. *Psychology of Men &*

Masculinity. Advance online publication. DOI: 10.1037/a0027642

Pew Research Center, (2018, March 1). *The generation gap in American politics.* Retrieved

from https://www.pewresearch.org/politics/2018/03/01/4-race-immigration-same-sex-

marriage-abortion-global-warming-gun-policy-marijuana-legalization/

Pirralha, A. (2020). Testing for Measurement Invariance with Many Groups. Retrieved from

https://bookdown.org/content/5737/

Pleck, J.H. (1981). *The myth of masculinity.* Cambridge, MA: MIT Press.

Pleck, J.H. (1995). The gender role strain paradigm: An update. In R.F. Levant & W.S. Pollack

(Eds.), *The new psychology of men* (pp. 11-32). New York: Basic Books.

Rabinowitz, F. E. (2014) Counseling men in groups In C. A. Brumley & B. E. Gatson (Eds.)

A counselor's guide to working with men, (55-70). Alexandria, VA: American

Counseling Association.

Rashid, T. Yasin, M. & Massoth, N. (2000). Pakistani males: An example of traditional

    masculine socialization. *Society for the Psychological Study of Men and Masculinity*

    Bulletin, 5(3), 25-26.

Reit, R., (2017). *The Relationship between the Military's Masculine Culture and Service*

    Members' Help-Seeking Behaviors. (2017). Master's Theses.

    http://epublications.marquette.edu/theses_open/410

Rimmer, J. (2015), Eric Anderson: 21st century jocks: Sporting men and contemporary

    heterosexuality, *Journal of Youth and Adolescence, 44,* 1180-1182, DOI:

    10.1007/s10964-015-0264-6

Robinson, S., Anderson, E., & White, A. (2018). The bromance: Undergraduate male friendships

    and the expansion of contemporary homosocial boundaries. *Sex Roles, 78,* 94-106. DOI:

    10.1007/s11199-017-0768-5

Rossem, A.H.D. (2016). Generations as social categories: An exploratory cognitive study of

    generational identity and generational stereotypes in a multigenerational workforce.

    Journal of Organizational Behavior, 40, 434-455. DOI: 10.1002/job.2341

Rönkkö, M. (2020, June, 15). *Measurement invariance.* [Video]. YouTube.

    https://www.youtube.com/watch?v=G5PEgxoRVC4

Ruspini, E. (2019). Millennial men, gender equality and care: The dawn of a revolution? *Teorija*

    in Praska, 56(4), 985-1000.

Schwing AE, Wong YJ, Fann M. (2014). The development and validation of the African

    American Men's Gendered Racism Stress Inventory. Psychology of Men & Masculinity.

    2013; 14:16–24.10.1037/ a0028272

Soper, D. (2016). A-priori sample size calculator for structural equation models [Computer

software]. Retrieved from https://www.danielsoper.com/statcalc/calculator.aspx?id=89

Spence, J. & Helmreich, R. (1979). On assessing androgyny. *Sex Roles, 5,* 721-738. DOI:

10.1007/BF00287935

Taylor, J. (1970). Sweet baby james [Song]. On *Sweet baby james.* Warner Bros.

Teofilo, C. N. (1999). An investigation of male role norms in battering men. *Dissertation*

Abstracts International, 60(5), 2372B.

Thompson, E.H. & Bennett, K.M. (2015). Measurement of masculinity ideologies: A critical

review. *Psychology of Men and Masculinity, 16*(2), 115-133.

http://dx.doi.org/10.1037/a0038609

Train. (2012). Bruises [Song]. On *California 37.* Sony Columbia.

UCLA. (2020). *What does Cronbach's Alpha mean?*

https://stats.idre.ucla.edu/spss/faq/what-does-cronbachs-alpha-mean/

U2. (1987). Where the streets have no name [Song]. On *The Joshua tree.* Bertelsmann Music

Group.

Vogel, D.L., & Wester, S.R. (2003). To seek help or not to seek help: The risks of self-

disclosure. *Journal of Counseling Psychology, 50*, 351-361.

Vogel, D.L., Wester, S.R., Wei, M., & Boysen, G.A. (2005). The role of outcome expectations

and attitudes on decisions to seek professional help. *Journal of Counseling Psychology,*

52, 459-470.

Vogel, D. L. & Wester, (2014) Referring men to seek help: The influence of gender role conflict

and stigma, Psychology of Men and Masculinity 15(1), 60-67, DOI: 10.1037/a0031761

Wade, J.C. & Brittan-Powell, C. (2001). Men's attitudes toward race and gender equality: The

importance of masculinity ideology, gender-related traits, and reference group identity dependence. *Psychology of Men and Masculinity. 2,* 42-50. DOI: 10.1037/1524-9220.2.1.42

Wertz, B. (2020). The new rules of style, according to millennials. *Men's Journal.*

https://www.mensjournal.com/style/new-rules-style-according-millennials/

Wester, S.R., Vogel, D.L., O'Neil, J.M., & Danforth, L. (2012). Development and evaluation of the gender role conflict scale-short form (GRCS-SF). *Psychology of Men and Masculinity, 13*(2), 199-210. DOI: 10.1037/a0025550

Wester, S.R., Arndt, D., Sedivy, A., & Arndt, L. (2010). Male police officers and stigma associated with counseling: The role of anticipated risks, benefits, and gender role conflict. *Psychology of Men and Masculinity, 11*(4), 286-302. DOI: 10.1037/a0019108

West Midland Family Center. (2019). *Generational Differences Chart.*

http://www.wmfc.org/uploads/GenerationalDifferencesChartUpdated2019.pdf

Wu, R., Levant, R. F., Sellers, A. (2001) The influence of sex and social development on masculinity ideology of Chinese undergraduate students. *Psychological Science (published by the Chinese Psychological Society), 24*, 365-366.

Yeazel, M.R. (2015). Your millennial toolbox: Developing culturally competent strategies for psychotherapy with men. *Journal of Men's Studies, 23*(2), 226-237. DOI: 10.1177/1060826515582522

Young, N. (1989). Rockin' in the Free World [Song]. *Freedom*. Reprise.

Young, P.A. (1996). A transgenerational study of sex role norms and development in adult men. *Dissertation Abstracts International, 56*(11), 6413B.

Zacher, H. (2015). Using lifespan developmental theory and methods as a viable alternative to

the study of generational differences at work. *Organizational Psychology: Perspectives on Science and Practice 8*(3), 342-346. https://doi.org/10.1017/iop.2015.47.

Appendix A: Informed Consent

<div align="center">

University of Wisconsin-Milwaukee
Informed Consent to Participate in Research
</div>

Study Title: The Next Generation—Millennials, the MRNI, and Measurement Invariance

Researcher[s]:

Stephen Wester, Ph.D.
Principle Investigator
Professor of Counseling Psychology
Department of Educational Psychology
University of Wisconsin—Milwaukee

Charlotte M. Shanaver M.S.Ed., M. Phil. Ed.
Student Co-Principle Investigator
Doctoral Student in Counseling Psychology
Department of Educational Psychology
University of Wisconsin—Milwaukee

We're inviting you to participate in a research study. Participation is completely voluntary. If you agree to participate, you can always change your mind and withdraw. There are no negative consequences, whatever you decide.

What is the purpose of this study?
We want to understand if different generations answer questions about masculinity the same way.

What will I do?
This survey will ask questions about your opinion of different personal attributes, attitudes, beliefs, and behaviors associated with traditional and non-traditional masculine gender roles. This survey will take approximately 15 minutes.

Risks
- Some questions may be uncomfortable. You may skip any questions you don't want to answer or stop the survey entirely by exiting the screen and not completing any further questions. However, you will not be compensated for incomplete surveys.
- Online data being hacked or intercepted: This is a risk you experience any time you provide information online. We're using a secure system to collect this data, but we can't completely eliminate this risk.
- Prolific Academic could link your worker ID (and associated personal information) with your survey responses. Make sure you have read Prolific Academic's participant and privacy agreements to understand how your personal information may be used or disclosed.
- Breach of confidentiality: There is a chance your data could be seen by someone who shouldn't have access to it. We're minimizing this risk in the following ways:
    - Data is confidential.
    - All identifying information is removed and replaced with a study ID.

- We'll store all electronic data on a password-protected, encrypted computer.

Possible benefits: No personal benefits identified. Possible benefits to society include a broader understanding of the construct of masculinity and possible generational variances.
Estimated number of participants: 360-500
How long will it take? Approximately 15 minutes. To rule out possible invalid survey completions, you will not be compensated if it takes you less than 5 minutes to complete the survey.
Costs: None.
Compensation: Each participant will be compensated $2.50 for their survey completion. As mentioned above, to rule out possible invalid survey completions, you will not be compensated if it takes you less than 5 minutes to complete the survey.
Future research: De-identified data (all identifying information removed) may be shared with other researchers. You won't be told specific details about these future research studies.
Funding source: None.
Confidentiality and Data Security
No personal identifying information will be collected for the purpose of this study. However, Prolific Academic could link your worker ID (and associated personal information) with your survey responses as Prolific Academic IDs will be collected to issue payment and destroyed after payment has been issued. Make sure you have read Prolific Academic's participant and privacy agreements to understand how your personal information may be used or disclosed.

Where will data be stored? Data is anonymous, stored on non-networked account with password protection on a secure computer. Participants are given PI contact information. Prolific Academic identifications will not be stored with data set.
How long will it be kept? The American Psychological Association requires that all data be kept for 7 years post publication. Data is anonymous, stored on non-networked account with password protection on a secure computer. Participants are given PI contact information. Prolific Academic identifications will not be stored with data set.
Who can see my data?
- We (the researchers) will have access to de-identified (no names, birthdate, address, etc.). This is so we can analyze the data and conduct the study.
- The Institutional Review Board (IRB) at UWM, the Office for Human Research Protections (OHRP), or other federal agencies may review all the study data. This is to ensure we're following laws and ethical guidelines.
- We may share our findings in publications or presentations. If we do, the results will be presented in aggregate (grouped) and de-identified (no names, birthdate, address, etc.) data. If we quote you, we'll use pseudonyms (fake names).
- Prolific: Because they own the Prolific Academic internal software, and to issue payment, Amazon will have access to your Prolific Academic worker ID. There is a possibility Prolific Academic could link your worker ID (and associated personal information) with your survey responses.

Contact information:

For questions about the research, complaints, or problems: Contact Charlotte Shanaver at shanaver@uwm.edu or Dr. Stephen Wester at srwester@uwm.edu

For questions about your rights as a research participant, complaints, or problems: Contact the UWM IRB (Institutional Review Board; provides ethics oversight) at 414-229-3173 / irbinfo@uwm.edu.

Please print or save this screen if you want to be able to access the information later.

IRB #:

IRB Approval Date:

Agreement to Participate

If you meet the eligibility criteria below and would like to participate in this study, click the button below to begin the survey. Remember, your participation is completely voluntary, and you're free to withdraw at any time.

- I am at least 18 years old
- I identify as male
- I am proficient in English

Appendix B: Male Role Norms Inventory-Revised

Please complete the questionnaire by circling the number which indicates your level of agreement or disagreement with each statement. Give only one answer for each statement.

Strongly Disagree	**Disagree**	**Slightly Disagree**	**No Opinion**	**Slightly Agree**	**Agree**	**Strongly Agree**
1	**2**	**3**	**4**	**5**	**6**	**7**

1. Homosexuals should never marry.
   1   2   3   4   5   6   7

2. The President of the US should always be a man.
   1   2   3   4   5   6   7

3. Men should be the leader in any group.
   1   2   3   4   5   6   7

4. A man should be able to perform his job even if he is physically ill or hurt.
   1   2   3   4   5   6   7

5. Men should not talk with a lisp because this is a sign of being gay.
   1   2   3   4   5   6   7

6. Men should not wear make-up, cover-up or bronzer.
   1   2   3   4   5   6   7

7. Men should watch football games instead of soap operas.
   1   2   3   4   5   6   7

8. All homosexual bars should be closed down.
   1   2   3   4   5   6   7

9. Men should not be interested in talk shows such as Oprah.
   1   2   3   4   5   6   7

10. Men should excel at contact sports.
    1   2   3   4   5   6   7

11. Boys should play with action figures not dolls.
    1   2   3   4   5   6   7

12. Men should not borrow money from friends or family members.
    1   2   3   4   5   6   7

13. Men should have home improvement skills.

1	2	3	4	5	6	7
Strongly Disagree	**Disagree**	**Slightly Disagree**	**No Opinion**	**Slightly Agree**	**Agree**	**Strongly Agree**
1	**2**	**3**	**4**	**5**	**6**	**7**

14. Men should be able to fix most things around the house.
    1   2   3   4   5   6   7

15. A man should prefer watching action movies to reading romantic novels.
    1   2   3   4   5   6   7

16. Men should always like to have sex.
    1   2   3   4   5   6   7

17. Homosexuals should not be allowed to serve in the military.
    1   2   3   4   5   6   7

18. Men should never compliment or flirt with another male.
    1   2   3   4   5   6   7

19. Boys should prefer to play with trucks rather than dolls.
    1   2   3   4   5   6   7

20. A man should not turn down sex.
    1   2   3   4   5   6   7

21. A man should always be the boss.
    1   2   3   4   5   6   7

22. A man should provide the discipline in the family.
    1   2   3   4   5   6   7

23. Men should never hold hands or show affection toward another.
    1   2   3   4   5   6   7

24. It is ok for a man to use any and all means to "convince" a woman to have sex.
    1   2   3   4   5   6   7

25. Homosexuals should never kiss in public.
    1   2   3   4   5   6   7

26. A man should avoid holding his wife's purse at all times.
    1   2   3   4   5   6   7

27. A man must be able to make his own way in the world.

1	**2**	**3**	**4**	**5**	**6**	**7**

Strongly Disagree	**Disagree**	**Slightly Disagree**	**No Opinion**	**Slightly Agree**	**Agree**	**Strongly Agree**
1	**2**	**3**	**4**	**5**	**6**	**7**

28. Men should always take the initiative when it comes to sex.
    1   2   3   4   5   6   7

29. A man should never count on someone else to get the job done.
    1   2   3   4   5   6   7

30. Boys should not throw baseballs like girls.
    1   2   3   4   5   6   7

31. A man should not react when other people cry.
    1   2   3   4   5   6   7

32. A man should not continue a friendship with another man if he finds out that the other man is homosexual.
    1   2   3   4   5   6   7

33. Being a little down in the dumps is not a good reason for a man to act depressed.
    1   2   3   4   5   6   7

34. If another man flirts with the women accompanying a man, this is a serious provocation and the man should respond with aggression.
    1   2   3   4   5   6   7

35. Boys should be encouraged to find a means of demonstrating physical prowess.
    1   2   3   4   5   6   7

36. A man should know how to repair his car if it should break down.
    1   2   3   4   5   6   7

37. Homosexuals should be barred from the teaching profession.
    1   2   3   4   5   6   7

38. A man should never admit when others hurt his feelings.
    1   2   3   4   5   6   7

39. Men should get up to investigate if there is a strange noise in the house at night.
    1   2   3   4   5   6   7

40. A man shouldn't bother with sex unless he can achieve an orgasm.
     1   2   3   4   5   6   7

Strongly Disagree	**Disagree**	**Slightly Disagree**	**No Opinion**	**Slightly Agree**	**Agree**	**Strongly Agree**
1	**2**	**3**	**4**	**5**	**6**	**7**

41. Men should be detached in emotionally charged situations.
     1   2   3   4   5   6   7

42. It is important for a man to take risks, even if he might get hurt.
     1   2   3   4   5   6   7

43. A man should always be ready for sex.
     1   2   3   4   5   6   7

44. A man should always be the major provider in his family.
     1   2   3   4   5   6   7

45. When the going gets tough, men should get tough.
     1   2   3   4   5   6   7

46. I might find it a little silly or embarrassing if a male friend of mine cried over a sad love story.
     1   2   3   4   5   6   7

47. Fathers should teach their sons to mask fear.
     1   2   3   4   5   6   7

48. I think a young man should try to be physically tough, even if he's not big.
     1   2   3   4   5   6   7

49. In a group, it is up to the men to get things organized and moving ahead.
     1   2   3   4   5   6   7

50. One should not be able to tell how a man is feeling by looking at his face.
     1   2   3   4   5   6   7

51. Men should make the final decision involving money.
     1   2   3   4   5   6   7

52. It is disappointing to learn that a famous athlete is gay.
     1   2   3   4   5   6   7

53. Men should not be too quick to tell others that they care about them.

1   **2**   **3**   **4**   **5**   **6**   **7**

Appendix C: Demographic Questionnaire

1. What gender to identify with?
   a. Male
   b. Female
   c. Other
2. What is your age?
   a. 18-24
   b. 25-34
   c. 35-44
   d. 45-54
   e. 55-64
   f. 65-74
   g. 75-84
   h. 85+
3. What race do you identify with?
   a. Black or African American, non-Hispanic
   b. White, non-Hispanic
   c. Hispanic or Latinx of any race
   d. American Indian or Alaska Native
   e. Native Hawaiian or Pacific Islander
   f. Asian American
   g. Two or more races
   h. Other
4. What is your current household income?
   a. Less than $15K
   b. $15K-$40K
   c. $40K-$60K
   d. $60K-$80K
   e. $80K-$100K
   f. Over $100K
5. What is the highest level of education you have completed?
   a. Less than a high school diploma
   b. High school diploma or equivalent (GED)
   c. Some college, no degree
   d. Associates degree (e.g. AA, AS)
   e. Bachelor's degree (e.g. BA, BS)
   f. Master's degree (e.g. MA, MS, Med)
   g. Doctorate (e.g. PhD, EdD)
6. Have you ever served in the US Military?
   a. Yes
   b. No

Appendix D: Reliability and Validity Statistics

Table 14
Total Score Model

Cronbach's Alpha	N of Items
.980	53

Table 15
Avoidance of Femininity

Cronbach's Alpha	N of Items
.945	8

Table 16
Fear and Hatred of Homosexuals

Cronbach's Alpha	N of Items
.960	10

Table 17
Extreme Self-Reliance

Cronbach's Alpha	N of Items
.857	7

Table 18
Aggression

Cronbach's Alpha	N of Items
.866	7

Table 19
Dominance

Cronbach's Alpha	N of Items
.926	7

Table 20
Nonrelational Attitudes Towards Sex

Cronbach's Alpha	N of Items
.857	6

Table 21
Restrictive Emotionality

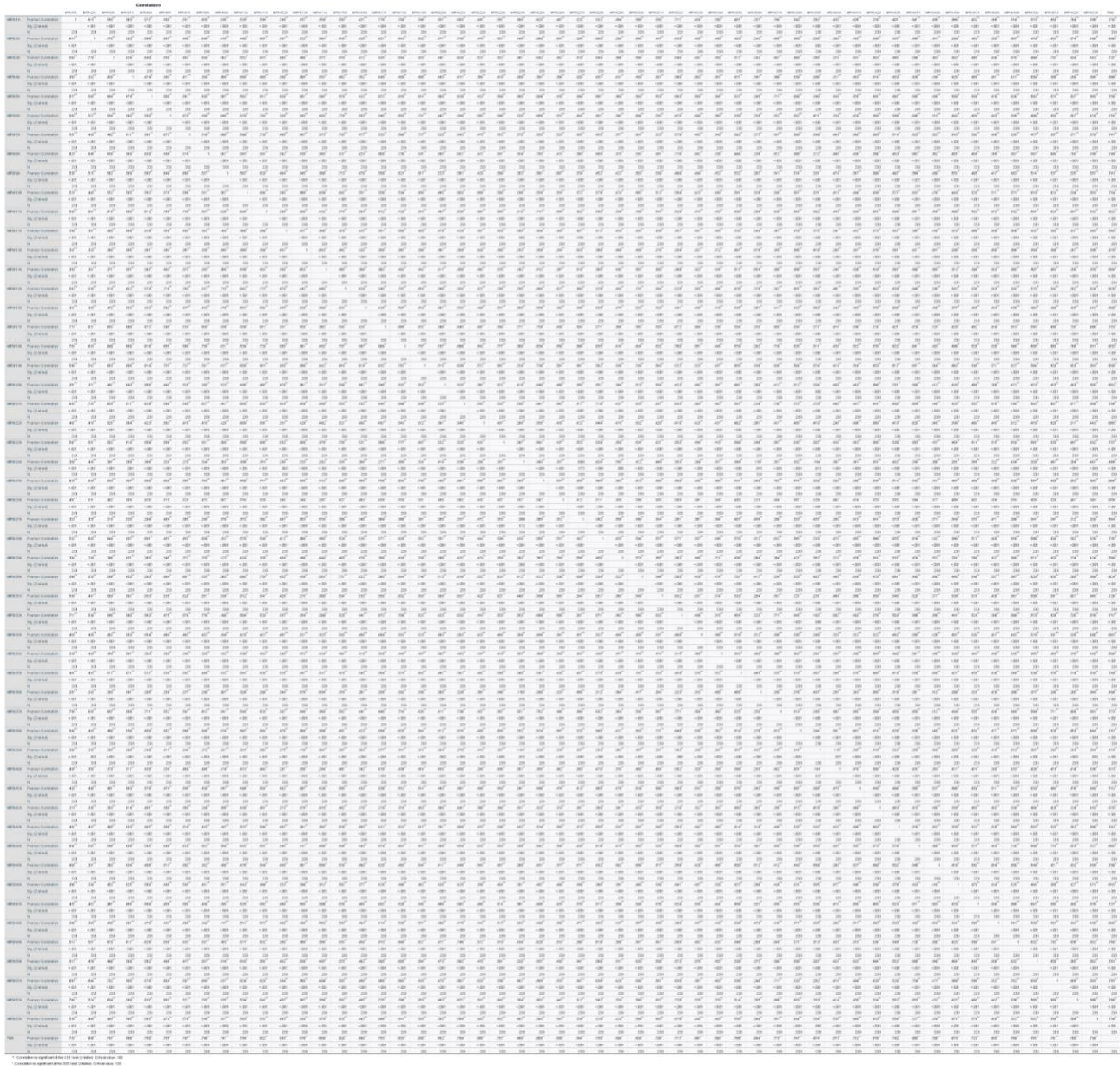| Cronbach's Alpha	N of Items
.911	8

Figure 1
Pearson's Correlation Matrix

Table 22

Exploratory Factor Analysis of MRNI-R (Levant et al., 2010)
KMO & Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		.958
Bartlett's Test of Sphericity	Approximate Chi Square	11816.570
	df	1378
	Sig	.000

Appendix E: Models

Figure 2
Total Score Model
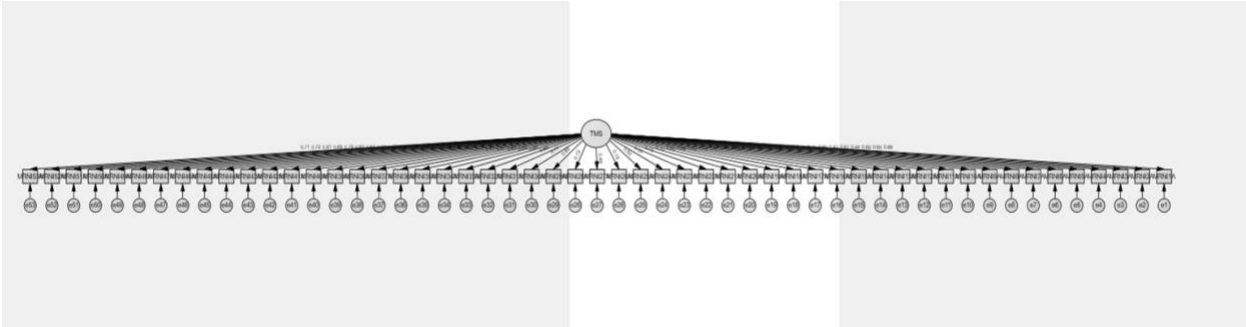
Note: Only items, onto general factor of masculinity

Figure 3
Full Model
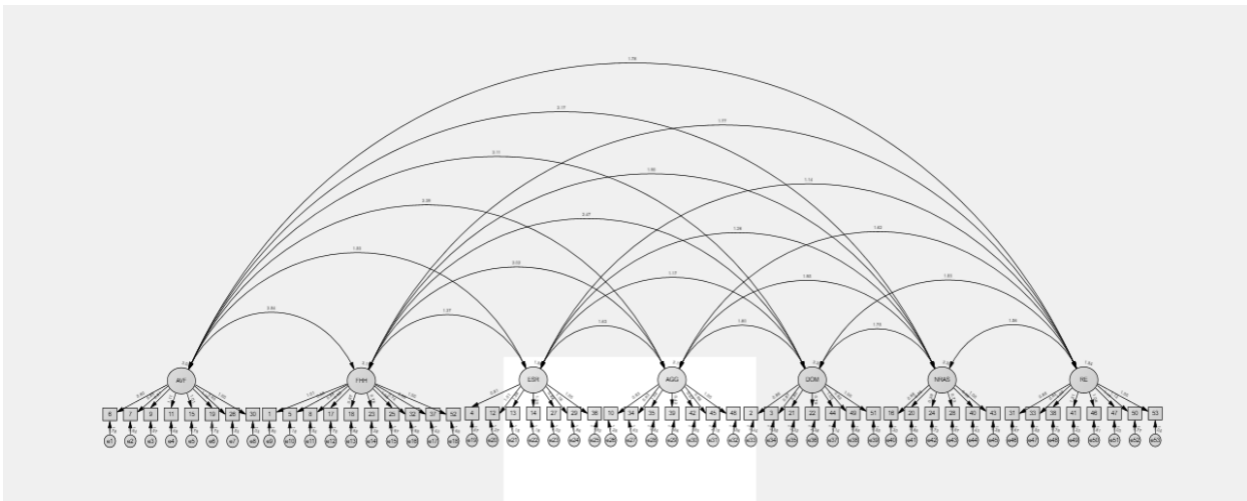
Note: Items, onto subscales, create general factor

Figure 4
Subscale: Aggression
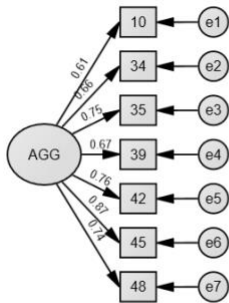
Note: Items 10, 34, 35, 39, 42, 45, 48, create factor or subscale, Aggression

Figure 5
Subscale: Avoidance of Femininity
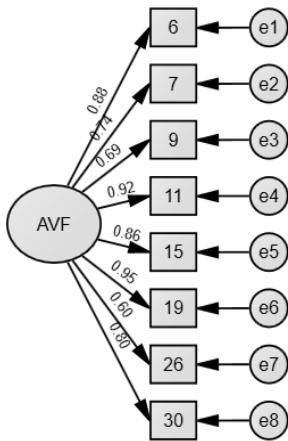
Note: Items 6, 7, 9, 11, 15, 19, 26, 30, create factor or subscale, Avoidance of Femininity

Figure 6
Subscale: Dominance
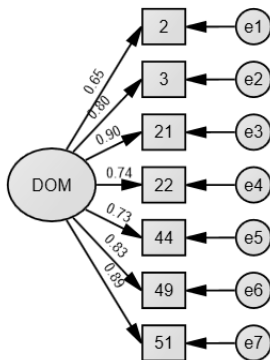
Note: Items 2, 3, 21, 22, 44, 49, 51, create subscale Dominance

Figure 7
Subscale: Fear and Hatred of Homosexuals
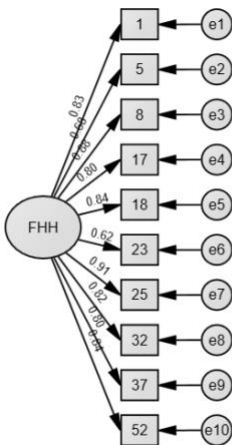
Note: Items 1, 5, 8, 17, 18, 23, 25, 32, 37, 52, create subscelae Fear and Hatred of Homosexuals

Figure 8
Subscale: Extreme Self-Reliance
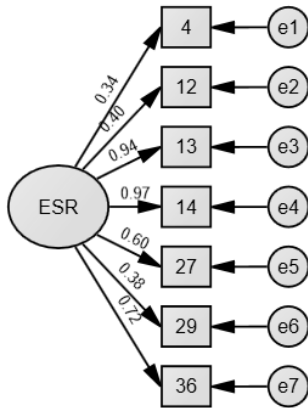
Note: Items 4, 12, 13, 14, 27, 29, 36, create subscale Extreme Self-Reliance

Figure 9
Subscale: Non-Relational Attitudes Towards Sex
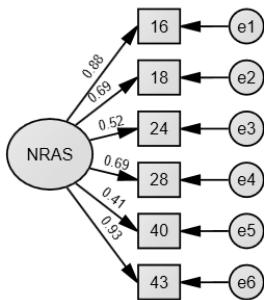
Note: Items 16, 18, 24, 28, 40, 43, create subscale Non-Relational Attitudes Towards Sex

Figure 10
Subscale: Restrictive Emotionality
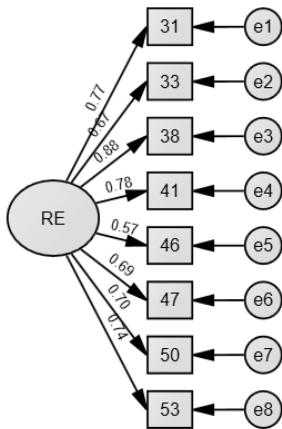
Note: Items 31, 33, 38, 41, 46, 47, 50, 53, create subscale Restrictive Emotionality

Appendix F: Invariance Output

Table 23
Total Masculinity Score—Configural Invariance: Chi Square

Model	CMIN	DF		CMNI/DF
Default	7444.659	2650	.000	2.809
Saturated	.000	0		
Independence	15252.298	2756	.000	5.534

Table 24
Total Masculinity Score—Configural Invariance: Comparative Fit Index

Model	CFI
Default	.616
Saturated	1.000
Independence	.000

Table 25
Total Masculinity Score—Configural Invariance: Root Mean Square

Model	RMSEA
Default	.089
Independence	.140

Table 26
Fear and Hatred of Homosexuals—Configural Invariance: Chi-Square

Model	CMIN	DF		CMNI/DF
Default	369.511	70	.000	5.279
Saturated	.000	0		
Independence	2530.474	90	.000	28.116

Table 27
Fear and Hatred of Homosexuals—Configural Invariance: Comparative Fit Index

Model	CFI
Default	.877
Saturated	1.000
Independence	.000

Table 28

Fear and Hatred of Homosexuals—Configural Invariance: Root Mean Square

Model	RMSEA
Default	.136
Independence	.343

Table 29

Aggression—Configural Invariance: Chi-Square

Model	CMIN	DF		CMNI/DF
Default	52.276	28	.000	1.867
Saturated	.000	0		
Independence	723.350	42	.000	17.223

Table 30

Aggression—Configural Invariance: Comparative Fit Index

Model	CFI
Default	.964
Saturated	1.000
Independence	.000

Table 31

Aggression—Configural Invariance: Root Mean Square

Model	RMSEA
Default	.061
Independence	.265

Table 32

Aggression—Metric Invariance: Chi-Square

Model	CMIN	DF		CMNI/DF
Default	80.454	34	.000	2.366
Saturated	.000	0		
Independence	723.350	42	.000	17.223

Table 33

Aggression—Metric Invariance: Comparative Fit Index

Model	CFI
Default	.932
Saturated	1.000
Independence	.000

Table 34

Avoidance of Femininity—Configural Invariance: Chi-Square

Model	CMIN	DF		CMNI/DF
Default	190.206	40	.000	4.755
Saturated	.000	0		
Independence	1762.994	56	.000	17.223

Table 35

Avoidance of Femininity—Configural Invariance: Comparative Fit Index

Model	CFI
Default	.912
Saturated	1.000
Independence	.000

Table 36

Avoidance of Femininity—Configural Invariance: Root Mean Square

Model	RMSEA
Default	.127
Independence	.363

Table 37

Avoidance of Femininity—Metric Invariance: Chi-Square

Model	CMIN	DF		CMNI/DF
Default	207.283	54	.000	3.839
Saturated	.000	0		
Independence	1762.994	56	.000	31.482

Table 38

Avoidance of Femininity—Metric Invariance: Comparative Fit Index

Model	CFI
Default	.91
Saturated	1.000
Independence	.000

Table 39

Avoidance of Femininity—Scalar Invariance: Chi-Square

Model	CMIN	DF		CMNI/DF
Default	207.283	54	.000	3.839
Saturated	.000	0		
Independence	1762.994	56	.000	31.482

Table 40

Avoidance of Femininity—Scalar Invariance: Comparative Fit Index

Model	CFI
Default	.910
Saturated	1.000
Independence	.000

Table 41

Dominance—Configural Invariance: Chi-Square

Model	CMIN	DF		CMNI/DF
Default	134.602	28	.000	4.807
Saturated	.000	0		
Independence	1391.881	42	.000	17.223

Table 42

Dominance—Configural Invariance: Comparative Fit Index

Model	CFI
Default	.921
Saturated	1.000
Independence	.000

Table 43

Dominance—Configural Invariance: Root Mean Square

Model	RMSEA
Default	.128
Independence	.373

Table 44

Dominance—Metric Invariance: Chi-Square

Model	CMIN	DF		CMNI/DF
Default	147.227	34	.000	4.330
Saturated	.000	0		
Independence	1391.881	42	.000	33.140

Table 45

Dominance—Metric Invariance: Comparative Fit Index

Model	CFI
Default	.916
Saturated	1.000
Independence	.000

Table 46

Extreme Self-Reliance—Configural Invariance: Chi-Square

Model	CMIN	DF		CMNI/DF
Default	110.623	28	.000	3.951
Saturated	.000	0		
Independence	823.788	42	.000	19.614

Table 47

Extreme Self-Reliance—Configural Invariance: Comparative Fit Index

Model	CMIN
Default	.894
Saturated	1.000
Independence	.000

Table 48

Extreme Self-Reliance—Configural Invariance: Root Mean Square

Model	RMSEA
Default	.113
Independence	.284

Table 49

Restrictive Emotionality—Configural Invariance: Chi-Square

Model	CMIN	DF		CMNI/DF
Default	102.846	40	.000	2.571
Saturated	.000	0		
Independence	1138.993	56	.000	20.339

Table 50

Restrictive Emotionality—Configural Invariance: Comparative Fit Index

Model	CFI
Default	.942
Saturated	1.000
Independence	.000

Table 51

Restrictive Emotionality—Configural Invariance: Root Mean Square

Model	RMSEA
Default	.082
Independence	.289

Table 52

Restrictive Emotionality—Metric Invariance: Chi-Square

Model	CMIN	DF		CMNI/DF
Default	105.798	54	.000	1.959
Saturated	.000	0		
Independence	1138.993	56	.000	20.339

Table 53

Restrictive Emotionality—Metric Invariance: Comparative Fit Index

Model	CFI
Default	.952
Saturated	1.000
Independence	.000

Table 54

Restrictive Emotionality—Scalar Invariance: Chi-Square

Model	CMIN	DF		CMNI/DF
Default	105.798	54	.000	1.959
Saturated	.000	0		
Independence	1138.993	56	.000	20.339

Table 55

Restrictive Emotionality—Scalar Invariance: Comparative Fit Index

Model	CFI
Default	.952
Saturated	1.000
Independence	.000

Table 56

Nonrelational Attitudes Towards Sex—Configural Invariance: Chi-Square

Model	CMIN	DF		CMNI/DF
Default	40.884	18	.002	2.271
Saturated	.000	0		
Independence	693.312	30	.000	23.110

Table 57

Nonrelational Attitudes Towards Sex—Configural Invariance: Comparative Fit Index

Model	CFI
Default	.966
Saturated	1.000
Independence	.000

Table 58

Nonrelational Attitudes Towards Sex—Configural Invariance: Root Mean Square

Model	RMSEA
Default	.074
Independence	.309

Table 59

Nonrelational Attitudes Towards Sex—Metric Invariance: Chi-Square

Model	CMIN	DF		CMNI/DF
Default	46.965	23	.002	2.042
Saturated	.000	0		
Independence	693.312	30	.000	23.110

Table 60

Nonrelational Attitudes Towards Sex—Metric Invariance: Comparative Fit Index

Model	CFI
Default	.964
Saturated	1.000
Independence	.000

Table 61

Nonrelational Attitudes Towards Sex—Scalar Invariance: Chi-Square

Model	CMIN	DF		CMNI/DF
Default	59.977	28	.000	2.142
Saturated	.000	0		
Independence	693.312	30	.000	23.110

Table 62

Nonrelational Attitudes Towards Sex—Scalar Invariance: Comparative Fit Index

Model	CFI
Default	.952
Saturated	1.000
Independence	.000