DEVELOPMENT AND VALIDATION OF THE COMFORT AND CONFORMITY OF GENDER EXPRESSION SCALE (CAGES)

A Doctoral Dissertation

Presented to

The Faculty of the Educational, School and Counseling Psychology Department

University of Missouri-Columbia

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

By

KATHERINE G. SPENCER

Dr. Roger L. Worthington, Dissertation Supervisor

DECEMBER 2007
The undersigned, appointed by the Dean of the Graduate School, have examined the dissertation entitled

DEVELOPMENT AND VALIDATION OF THE COMFORT AND CONFORMITY OF GENDER EXPRESSION SCALE (CAGES)

presented by Katherine Spencer,

a candidate for the degree of Doctor of Philosophy,

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

_______________________________________________________________
Roger Worthington, PhD
Committee Chair

_______________________________________________________________
Laurie Mintz, PhD

_______________________________________________________________
Lisa Flores, PhD

_______________________________________________________________
Glenn Good, PhD

_______________________________________________________________
Mary Jo Neitz, PhD
Outside Member
ACKNOWLEDGEMENTS

There was something that I said more and more as I neared the end of the dissertation process: “it takes a village to complete a dissertation.” I said this meaning that I had so much love, support, and help from my friends, colleagues, mentors, and family, I knew I could never have completed this project on my own. I am lucky enough to have the most wonderful, loving, kind, generous, and intentional community of friends, family, chosen family, and mentors that one could ever wish for. I know this acknowledgement section will not be enough to express my gratitude to each of them, but I hope it serves as a small thank you to individuals who make my life so fabulous.

First, I want to thank my partner Beth Derenne. I know most people thank their partners last, but Beth needs to be acknowledged first and foremost. Beth, you have made the attainment of this degree, the finishing of this dissertation, and the ability for me to reach my dreams possible. You have done so many things – financial support, chores around the house, backrubs when stressed, editing, listening, theorizing, reassuring, validating, and putting up with me– that you deserve this degree with me. I cherish our nine years together and feel blessed that we have grown together as we have grown into the people we are today. I respect, admire, adore, and trust you completely. Our life together is our greatest accomplishment. Thank you for being my partner and my best friend.

Secondly, I want to thank my family and my chosen family. Ted Spencer, my father, has been my anchor for many years, guiding me through my education, and providing me with a gentle stability that has been my bedrock. Dad, I love you very much, and your pride in my accomplishments touches me deeply. I am proud to continue
the Spencer tradition of higher education. My mother, Paula Ottofy, has provided me with so many of the qualities that make me a good therapist, kindness, joy, and sociability. Mom, I love you, and I thank you for always affirming me in so many ways, and being a role model of a woman who speaks her mind. My grandmother, Annabelle Spencer provided me with support, role modeling, and the financial support to achieve my education. Thank you grandma for everything you have done to help me achieve this goal. My brothers, William and Ray Spencer and Dylan Ottofy, have given me so much, but I want to thank you guys for listening to me, telling me I could do it, and making me laugh.

I want to thank deeply my chosen family and dear friends who have helped me in a multitude of ways. Beth Pickens and Debbi Doerflinger, my best friends, you both supported me tremendously in different ways. Beth, thank you for talking me through so many crises, always being my support, my collaborator, and being there for every step of this process. I am the Shirley to your Esther, you are my true womynfriend. Debs, you are always in my corner, and I cannot thank you enough for being the steady and supportive force in my life. I hope you know how much our friendship means to me. Laura Hacquard, thank you for being my mentor and guiding force. I cannot express in words my gratitude to you. I would not be the person I am today without you in my life. Kendra Yoder, thank you for being my feminist friend, my buddy, my goodness and light, with a sassy side, of course. Struby Struble, KatieBlair, and Becky Smith, thank you for being my crew, my girl gang, and for your data collection and editing help. Carrie Freemen, I cannot thank you enough for your enduring friendship and your Microsoft word skills! I want to thank the family of the MU Women’s Center. It is more than just a space! My
involvement with the Women’s Center made it possible for me to make it through graduate school and to finish this dissertation.

I am grateful and thankful to my extended friends and family of Columbia. Dan Delgado and Steve Kehnel, thank you for your friendship and for helping me to collect data! Sean Clouse, you have been my friend, confidante, and fellow traveler. Thank you for all your guidance, support, and reassurance. You went first, and I benefited from what you learned! I am blessed to have the support, advisement, and friendship of the fabulous Gretchen Reinders and Rimiko Thomas. Thank you both for the long phone conversations, your friendship, and for being on this path through graduate school together. We made it!! Thank you as well to Meghan Davidson and Oksana Yakushko. You both encouraged me to apply to the doctoral program and always reaffirmed my interests in social justice and feminism in psychology. Thank you to my intern cohort, Jennifer Caldwell, Julie Hau, and Tamba-Kuui Bailey. Each of you helped me in different ways, and always offered me your support. Tamba-Kuui, our conversations have pushed me to grow, and your support pushed me through this final phase.

I believe as a daughter of the women’s movement, as a woman who has benefited from the work of my foremothers, that it is important to honor and acknowledge our feminist mentors. I am lucky to have had many. Thank you Mary Jo Neitz, Debbie Wright, Laurie Mintz, and again Laura Hacquard. You all have challenged me to grow, supported me, and lifted me up through your own hard work. I am deeply indebted to you all.

Thank you to my committee who offered comments, support, and enthusiasm as I proposed and ultimately completed this dissertation: Laurie Mintz, Lisa Flores, Glenn
Good, and Mary Jo Neitz. Finally, I have had the pleasure of working under the tutelage and advisement of Roger Worthington for the past six years. Thank you Roger for all the things you have taught me, and for allowing me to be myself. I know it was not always the smoothest ride, but I have grown so much with you, and am excited for the next phase of our collaboration!
TABLE OF CONTENTS

ACKNOWLEDGEMENTS ........................................................................................................... III
LIST OF TABLES ........................................................................................................................ IX
LIST OF FIGURES ....................................................................................................................... X
ABSTRACT ................................................................................................................................. 1
CHAPTER 1: OVERVIEW ............................................................................................................. 3
  History and Overview of Gender Role Conflict Scale ............................................................. 5
  Implications for Theory, Research, Practice, and Training ...................................................... 13
CHAPTER 2: REVIEW OF LITERATURE ................................................................................. 15
  History of Gender Measurement and Scale Development ...................................................... 16
  Androgyny and Expanding Notions of Gender/Sex ................................................................. 19
  Men and Masculinity Studies ................................................................................................. 24
  New Directions ...................................................................................................................... 28
  Gender as a Social Construct ................................................................................................. 31
  Measurement of Gender ......................................................................................................... 34
CHAPTER 3: METHOD ................................................................................................................. 37
  Study 2: Reliability and Validity Tests of the CAGES ............................................................... 44
  Study 3: Test—Retest Reliability Estimates .......................................................................... 49
CHAPTER 4: RESULTS ............................................................................................................... 52
  Study 1: Scale Construction and Exploratory Factor Analysis ............................................... 52
  Study 3: Test-Retest Reliability .............................................................................................. 70
CHAPTER 5: DISCUSSION ......................................................................................................... 72
  Relationship of the Results to Previous Research and Theory .............................................. 72
  Gender-Role Conflict and Gender Expression ....................................................................... 74
  Continua of Comfort and Conformity .................................................................................... 75
Between Group Differences.................................................................78
Gender Norm Conformity ......................................................................79
Women Completing the CFNI.................................................................82
Women Completing the CMNI.................................................................83
Men Completing the CMNI.................................................................84
Men Completing the CFNI.................................................................85
Test-Retest Reliability.................................................................86
Implications for Future Research and Practice ......................................87
Limitations of the Current Research ..................................................91
Conclusions........................................................................................93
REFERENCES..................................................................................94
APPENDICES..................................................................................99
VITA..........................................................................................138
LIST OF TABLES

1. Item-Total Correlations, Factor Matrix, Communalities, and Item Means and Standard Deviations for the CAGES 125

2. Table 2: Intercorrelations Between Subscales of CAGES and Validity Instruments 127

3. Table 3: Means, Standard Deviations, Skewness, and Kurtosis Studies 1-3 131
LIST OF FIGURES

1. Continua of Comfort and Conformity of Gender Expression  133
2. Frequencies of Responses for CAGES subscale: Discomfort with Nonconformity (DWN)  134
3. Frequencies of Responses for CAGES subscale: Resentful Conformity (RC)  135
4. Frequencies of Responses for CAGES subscale: Active-Physical Comfort Conformity (APCC)  136
5. Frequencies of Responses for CAGES subscale: Comfort with Conformity-Appearance (CCA)  137
ABSTRACT

Gender roles and stereotypes have been shown to be deeply entrenched across race, class, age and location (Bem, 1993). Gender has been a focus of psychological measurement since the early 20th century (Terman & Miles, 1936). The Comfort and Conformity of Gender Expression Scale (CAGES) was developed in order to assess an area not yet explored, comfort and discomfort with conformity and nonconformity of gender expression in a broad sample of adults. After initial item development, three studies were completed to assess the validity of the CAGES as a measure of comfort and conformity of gender expression. Study 1 consisted of the initial exploratory factor analyses (EFA) to determine factor structure and item relevance. A sample of 356 participants completed the initial CAGES. The final EFA resulted in a 24-item, 4-factor scale. The four factors were named; Discomfort with Nonconformity, Resentful Conformity, Active-Physical Comfort Conformity and Comfort with Conformity-Appearance. The CAGES was hypothesized to measure cognitive, affective and behavioral components of gender expression. The hypothesized scale was partially supported in that the final 4-factor CAGES measures behavioral and affective components of comfort-discomfort with conformity-nonconformity in gender expression. Study 2 examined reliability and convergent and divergent validity estimates for the CAGES by correlating the revised CAGES with two established scales, the Conformity to Masculine Norms Inventory (CMNI) and the Conformity to Feminine Norms Inventory (CFNI). A total of 176 participants completed the revised CAGES and the CFNI and CMNI. Moderate to strong correlations were found between all four of the
CAGES subscales and six subscales on the CMNI and five subscales on the CFNI, and no correlation with a measure not related to gender, the Balanced Inventory of Desirable Responding-Impression Management. Study 3 established test-retest reliability through administering the CAGES to 56 participants who completed the revised CAGES twice over a two-week period. Significant correlations were found on three of four CAGES subscales, evidencing test-retest reliability for three CAGES subscales as a measure of gender expression. Thus, the CAGES contributes to the existing literature in gender measurement by introducing a 24-item, 4-factor scale that measures comfort and conformity of gender expression.
Gender is a psychological construct influencing the lives of all individuals. Gender roles and stereotypes have been shown to be deeply entrenched across race, class, age and location (Bem, 1993). Although much of the early research on gender was centered on identifying men that were possibly gay by measuring levels of effeminacy (Terman & Miles, 1936, Bem, 1993), gender has been studied as a psychological construct since the 1930’s. The political nature of gender has been an important and consistent theme throughout the history of gender research. In reviewing the research on gender, writers consistently note the sociocultural context within which gender research occurs, and how this context has consistently shaped gender research from it’s inception through the women’s rights movement in the 1970’s to the men’s psychology movement of the 1980-90’s (Deaux, 1999, Thompson, Pleck, & Ferrera, 1992, Stewart & McDermott, 2004). Many scholars have pointed out that gender has been used as a tool of oppression against women, men, transgender persons, gays and lesbians, and other marginalized groups. An understanding of the sociopolitical nature of gender highlights the difficulties scientists experience in attempting to disentangle gender research, and psychological measurement in particular, from entrenched biopsychosociocultural contexts.

There have been three waves of gender-focused research and scale development of the past 60 years (Bem, 1993, Good, Wallace & Schuster, 1994). These three waves can be typified as (a) the pre-1960’s pathology seeking model, (b) the androgyny and gender freedom models, and (c) the masculinity focused models.
The study of gender previous to the civil rights, women’s, and LGB rights movements of the 1960’s and 1970’s focused on the pathologization of femininity and the utilization of non-conformity to patriarchal gender stereotypes to identify “potential homosexuals” (Terman & Miles, 1936). These measures often proved to be inaccurate assessments of gender. For example, the M-F scale (scale 5) of the Minnesota Multi-Phasic Personality Inventory (MMPI), which was originally developed to assess for potential homosexuality, was found to be a more accurate measure of conformity to gender stereotypes that did not correlate with sexual orientation (Groth-Marnat, 2003).

During the 1970’s, with the rise of women’s liberation movements, psychologists began to question sex difference focused research, and began constructing scales that established the concept of androgyny – a mixture of feminine and masculine traits. The androgyny or gender freedom models of gender research, headed by psychologists such as Janet Spence and Sandra Bem, examined the attributes that were assigned as masculine or feminine (Bem, 1974, Spence & Helmreich, 1976). Utilizing modern conceptions of sex stereotypes, Bem and Spence were able to clarify how sex stereotypes were used to perpetuate sex roles and show the true diversity in masculine and feminine attributes in men and women.

The androgyny and gender freedom scales were foundational to the scale development that would come in the 1980’s and 1990’s. After the breakthroughs of the Bem Sex-Role Inventory (Bem, 1974), the Personality Attributes Questionnaire (Spence, Helmreich & Stapp, 1976) and the Attitudes Towards Women Scale (Spence & Helmreich, 1977), gender measures proliferated into examining sex roles, gender norms, feminist identity development, and sex stereotypes, among others. As psychologists
explored gender as a construct, new questions about how masculinity was constructed and enforced arose.


**History and Overview of Gender Role Conflict Scale**

The Gender Role Conflict Scale (GRCS), developed by O’Neil (1986), has been foundational to the exploration of the interaction of culture and internal experiencing of gender. The GRCS provides a rich framework for the exploration of the restrictiveness of gender norms on men’s gender expressions. The GRCS measures gender role conflict across behavioral, cognitive and affective realms. O’Neill and his colleagues have hypothesized the emotional and psychological effects of experiencing gender role conflict, and have explored the processes through which conflict occurs within multiple contexts. Central to O’Neill’s conceptualization of gender role conflict are what he identified as six patterns of GRC; 1) restrictive emotionality, 2) health care problems, 3) obsession with achievement and success, 4) restrictive sexual and affectionate behavior, 5) socialized control, power, and competition issues, and 6) homophobia. O’Neill, in his initial conceptualization (1986) of GRC stated his hypothesis that these six identified patterns were related to personal and institutional sexism in the larger society.

O’Neill tested his hypotheses with the development of the GRC Scale, and found strong empirical support for four patterns, renamed as; 1) Success, Power,
Competition (SPC), 2) Restrictive Emotionality (RE), 3) Restrictive Sexual and Affectionate Behavior Between Men (RABBM), and 4) Conflict Between Work and Family Relations (CBWFR). The GRCS has strong internal consistency, good divergent validity, and strong support for the four-factor solution. The GRCS has been used in hundreds of published, unpublished and dissertation studies, and continues to prove to be a reliable and valid empirically supported instrument.

There have yet to be any published studies that have attempted to extrapolate the GRCS factor structure to navigating the pressures of institutionalized patriarchal femininity for women (traditional notions of female gender norms). Just as men experience gender-role conflict in expressing or not expressing traditional masculine gender norms, women experience pressure to embody a femininity constructed by traditional and dominant notions of women’s roles. New scales have been developed since the GRCS, utilizing it as a theoretical foundation. Two recent measures, the Conformity to Masculine Norms Inventory (CMNI) (Mahalik et al., 2003), and the Conformity to Feminine Norms Inventory (CFNI) (Mahalik, 2005), built their foundation from the GRCS theory, but stated the GRCS pathologizes masculine gender norm expression, and advocated for a continuum of experience with masculine and feminine gender role. This is an important step, but there are several domains left to explore within the arena of gender role conflict and gender norms in psychology.

These three waves of gender research have created a vast research base in gender. Gender scale construction has been a rich area for scientific inquiry, but there are theoretical and construct issues that remain not fully explored. Specifically, there remains a continuing divide in recent years between gender research focused on men’s
experiences and a holistic approach to gender. Additionally, the scales that have been developed have been based in gender stereotypes. The construct of gender expression has gone largely unquantified, and the literature still fails to reflect the true diversity of gender. In order to address these concerns within the field of gender measurement, interdisciplinary approaches, specifically the fields of feminist and queer theory, are helpful in expanding current understandings of gender.

From a feminist perspective, there have been arguments for the development of research that is inclusive of gender nonconformant people and acknowledges the social context. Several theorists in feminist literature have begun to hypothesize about a range of masculinities and femininities and what these “other” genders might embody (hooks, 2000, Faludi, 2002, Taylor, 2005). Within psychology, much of the men and masculinity literature appears to be exploring the constructions of masculinity and examining how men experience the rigid enforcement of patriarchal masculine norms (O'Neil, et al., 1995). Theorists continue to argue over what constitutes masculinity, how it is enacted, and the meanings of these constructions. However, many feminist theorists agree that individuals experience gender within a white supremacist capitalist heterosexist patriarchy where men are given power and privilege with their maleness, where women are denigrated as less than and femininity is seen as weakness (hooks, 2000). Research done within the arena of gender continues to be interpreted within dominant western social context and thus continues to have political ramifications, thus any measure developed must acknowledge these contextual and societal factors.

Anne Fausto-Sterling (2000) in her book, Sexing the Body, discusses the existence of five sexes. She argued the current two sex system should be revamped to include the
categories herms (named after "true" hermaphrodites), merms (named after male "pseudohermaphrodites"), and ferms (named after female "pseudohermaphrodites"). A "true" hermaphrodite has an ovary and a testis, or a combined gonad called an ovo-testis. A "pseudohermaphrodite" has either an ovary or a testis, along with genitals from the "opposite" sex. The term hermaphrodite has mostly been replaced with the term “intersexual” in the literature and by intersexual activists (ISNA, 2007). This proposal has served as a starting point for many within the transgender and intersexual movement for an expansion of recognized gender and sex identities (Wilchins, 2002, Bernstein-Sycamore, 2007)

Current theoretical conceptualizations utilized in gender scale development are limiting to the true diversity of gender expression. The men and masculinity theoretical developments mainly apply to men’s experiences of gender, and although many of the main tenets may be transferable, gender is a phenomena experienced by all sexes. Additionally, there has not been a scale developed with a framework that is inclusive of transgender, gender variant, and genderqueer identities as normative. Many of the scales that have been developed work from a gender binary theoretical assumption and do not address non-conformant gender expressions from a queer perspective.

Gender role conflict, gender norms, sex roles, and the scales that assess these phenomena are lacking the assessment of the domain of gender expression. Although most of these scales address gender expression tangentially, and the conflict or norm endorsement that is being assessed would logically be an outcome of gender expression, it has not been fully articulated or addressed. Gender expression is one component of the
broader gender constructs that have already been developed, and it is of primary
importance in deepening the measurement of gender to fully explore gender expression.

The proposed project is the development and validation of the Comfort and
Conformity of Gender Expression Scale (CAGES). This scale is intended to measure the
levels of comfort or discomfort people feel in expressing their own sense of gender
within a cultural context, or how conformant or nonconformant they feel in their gender
expression within the social context. The underlying principles of this scale are based in
the theoretical foundations of the gender and sex-role scales developed between the
1970s-2000s stemming from the women’s and men’s psychology movements. A key
difference of the CAGES from previously developed scales is that the CAGES is
hypothesized to measure gender expression, whereas previous scales focused on utilizing
gender stereotypes to measure attitudes or adherence to sex-role mandates.

Gender expression is defined as the external manifestation (physical
presentations, social interactions, behaviors) of the internal experience of social and
cultural roles associated with biological sex. Expression may be congruent with social
mandates or incongruent, or a mix. Expressions may vary within contexts, be fluid, or
remain static over time. Gender expression can be fluid and is impacted by socialization,
internalization of sexist gender norms and the sociocultural context.

In constructing the CAGES, three conceptual content areas were identified as
constituting the construct of gender expression. Utilizing the framework provided by
O’Neill in the development of the construct of gender conflict, the domains of behavior,
cognition, and affect were identified as important in the conformity-comfort of gender
expression. It is expected that gender will be expressed through these three domains, and
that respondents will be positioned on a continuum of comfort or discomfort with conformity or nonconformity in behaviors, cognitions, and affect that will impact their gendered expression. This gendered behavior is not only expressed within a patriarchal sociocultural context, but is impacted and thus expressed on differing levels through this context.

While the GRCS measures the level of conflict men feel about challenging gender norms, the CAGES will measure the expression of gender in the cultural context through cognitions, behaviors and affect. The CAGES also differs from the CFNI and CMNI in that Mahalik’s scales measure the endorsement of social gender norms, while the CAGES seeks to measure how the person expresses their gender through their thoughts, behaviors, and feelings. Mahalik and colleagues worked to identify sets of social gender norms. O’Neill identified feelings of conflict based on identified social gender norms. The CAGES items seek to measure how the person expresses their gender within the context of highly regulated gender roles.

There are similarities between the CAGES and the GRCS. The key differences between the CAGES and the GRCS is that while the GRCS items name an emotion that is felt in relation to specific masculine gender norms, the CAGES is meant to be utilized with men and women, and names behaviors, thoughts or feelings felt in relation to expressing gender within the cultural context.

It is imperative that a paradigm shift in the understandings of gender and gender roles begins. We must contextualize gender within a capitalistic, patriarchal system that devalues femininity and privileges masculinity (hooks, 2000). The privileging of masculinity and devaluing of femininity does not occur within a vacuum, and the
development of gender identity does not occur within a vacuum either. To fully understand and assist clients struggling with issues of gender and identity (all clients to some degree, as we exist within a gendered world) we must begin to explore the complex interactional forces of gender within the cultural context. The CAGES seeks to examine the co-existing and interactional forces of gender expression within a cultural context and the gender expression congruency or incongruence within the self.

The CAGES is being developed as one scale that can be administered to men, women, transgender, transsexual and intersex (previously known as “hermaphrodite”) individuals. In the past, many gender scales have been developed with separate forms for men and women. As mentioned above, this methodology, although methodologically simpler, tends to reify the patriarchal dualistic gender binary and ignores the existence of other sex identities such as transgender, genderqueer and intersex (see Table 1 for language and definitions used in this dissertation).

The proposed project intends to engage in a paradigmatic shift that steps away from measuring gender stereotypes and towards measuring the subjective experiences of gender that are mediated by biopsychosociocultural processes through measuring comfort and conformity in gender expression. Gender stereotypes are produced and maintained through hegemonic patriarchal society (Bem, 1997). The previous gender scales that have been developed have often utilized stereotypes of masculinity and femininity to guide what is considered gender norms. This reproduction of hegemonic gender actually reinforces the measurement of gender stereotypes as gender instead of the individuals’ internal and external experience of gender. The current proposed scale intends to measure
gender expression within the social-cultural context and use inclusive language to include gender nonconformant individuals.

An important grounding tenet in this paradigmatic shift is the focus on the gender blended self versus the idea that women can only identify as feminine or men as masculine. Based on evidence from responses from the BSRI, PAQ, and others, men and women embody a range of attributes considered masculine and feminine (Gilbert & Scher, 1999). Furthermore, Spence, the author of the PAQ, has argued that the widely used measures for masculinity-femininity are really measures of instrumentality (rational or concreteness) or expressiveness (emotionality) (Spence, 1993, 1999; Spence & Buckner, 1995, 2000; Spence & Helmreich, 1981). This points to the stereotypes inherent in the proposed measurement of gender. Hoffman (2000) in the development of the Hoffman Gender Scale (HGS) utilizes separate forms for women and men and in these separate forms includes items such as “I feel confident in my femininity (or masculinity).” She also allows the participant to define their own sense of femininity or masculinity, but retains the idea of them as separate, discrete concepts for men and women. No matter how femininity or masculinity (assumed to be mandated and shaped through the sexist, heterosexist, patriarchal context of Western society) is defined, correlating gender identity health with sex identity reifies the ideology that gender is naturally connected with biological sex. Diamond (2000) has argued that instead of using the term gender identity as a correlate to biological sex, psychologists should free gender from conflation with sex, and instead frame what is traditionally thought of as gender identity as sex identity. One’s sense of maleness or femaleness is their sex identity not their gender identity (Diamond, 2000). This depathologizes gender nonconformant
behavior and still acknowledges transgender person’s sense of sex identity. It is also consistent with the writings on gender from psychology, women’s studies and queer studies that argue that gender is a social construct that is socially enforced and individually experienced (Butler, 1990, Wilchins, 1997).

**Implications for Theory, Research, Practice, and Training**

The range of gendered identities is of central importance to the theoretical foundation of the proposed scale, the Comfort and Conformity of Gender Expression Scale (CAGES). The CAGES is being constructed to measure an individual’s sense of conformity or comfort in expressing their internal sense of gender identity within a gender rigid context. By engaging in this theoretical, conceptual and applied paradigmatic shift, the proposed scale will further the existing canon of gender research in beginning to normalize a continuum of gendered experiences.

By articulating clearly the differences between sex identity and gender identity, the CAGES shifts the existing language and framework utilized in gender measurement. This is important to the field of counseling psychology in several ways. By expanding traditional notions of gender, researchers have increasingly more precise tools for exploring and understanding how gender impacts individuals within diverse contexts. As discussed, gender is an amorphous construct and because it is intrinsically linked to sexuality, it is important for researchers to be able to articulate gendered expressions specifically. Practitioners are in need of finely tuned understandings of their clients’ experiences and expressions of gender because the archaic conceptions of gender are limited in working with gender diverse populations.
For practitioners working with gender nonconformant populations, it is important to understand how contextual variables impact the variance of expression within different environments. Specifically, for practitioners working with clients who wish to transition, many transgender health agencies encourage clients to live “full time” for a year prior to surgery. This can be difficult within gender rigid contexts. Understanding how external pressures to conform versus internal comfort factors are affecting the client’s gender expression could move towards a more flexible understanding of meeting these requirements.

Finally, as counseling psychology is positioned at the forefront of social justice in psychology, the grounding of the proposed scale in feminism and consciously resistant to oppressive models of gender theories pushes the field to create space within society for a multiplicity of gender expressions.
Chapter 2: Review of Literature

Gender frames a set of questions that guide scientific inquiry. Psychologists use gender as a descriptor, as a variable, and as a tool of difference. In order to engage in an exploration of the past and current research on gender, we must first define it within a psychological context.

Gender and sex are used interchangeably in everyday use as well as in the literature. In the early 1970s, the distinction between sex and gender was clarified. Sex was defined as the biological/physical attributes that are anatomically and physiologically determined (Fausto-Sterling, 2000). Gender is defined as the psychological manifestation of sex which is socially constructed and maintained through gender norms (Money & Erhardt, 1972). Other researchers have begun to expand and clarify this definition in order to understand how gender can be socially constructed yet also internally experienced and deeply felt phenomenon (Spence & Buckner, 2000). This complexity continues to make gender difficult to operationalize. To further understandings of gender, psychologists have tried to delineate differences between the terms gender, gender identity, gender role, sex role, gender socialization, gender schema and gender expression, among others, to help articulate the complexity of gender as a construct.

Beginning with feminist psychologists’ attempts to differentiate between sex and gender in the 1970’s and continuing with the research on men and masculinity, there has been a rich psychological canon developed in exploring gender constructs.

Western psychologist researchers have struggled with defining and making tangible what constitutes gender as a construct since the early 20th century. Research has
focused on defining gender itself, exploring biological sex and its relation to gendered behaviors, as well as numerous studies attempting to comprehend how gender stereotypes work, and how they influence many different areas of cognition, decision making, choice and schematic processing (Deaux, 1999). An important area within gender research has been gender measurement and assessment. Within this area, there is a diverse history of gender scale development.

Gender measurement has always been connected to the sociopolitical climate. Social contextual factors impact gender measurement. Beginning with efforts to operationalize gender with prevalent gender stereotypes, identifying of deviant populations, through the reevaluation of gender incited by the women’s liberation movement and the emergence of masculinity studies gender measurement has remained intertwined with social factors. In reviewing the relevant literature, there are three identifiable waves of gender research; (a) the pre- 1960’s pathology seeking model, (b) the androgyny and gender freedom models, and (c) the masculinity focused models. Along with these three historical approaches, there are new directions within gender measurement that are utilizing the knowledge gained from this earlier research and the acknowledgement of the impact of sociopolitical contextual variables on gender measurement.

*History of Gender Measurement and Scale Development*

Psychologists began developing assessments to measure gender as early as the 19th century. The focus of much of this theorizing and development was in proving women’s inherent inferiority. Some of the first research establishing gender as a psychological construct was Lewis Terman’s work on “mental masculinity and
femininity” (Bem, 1993). Terman and his coauthor Catherine Cox Miles pre-tested and retested scale items that seemed to differ between men and women. Individuals could either positively or negatively endorse items such as preferences for hunting or people with loud voices (named as masculine), or nursing and babies, (named feminine). Masculine items also included dislike for taking baths, denial of fear of the dark, and denial of care for dress. Feminine items included denial of extreme disobedience as a child and endorsement of preference for someone else to take the lead (as described in Bem, 1993).

These early measurements were developed as single dimension scales with socially constructed stereotypical hyper-masculine or hyper-feminine traits as anchors. In the Terman and Miles example, this forced individuals to choose either masculine or feminine traits, enforcing a dualistic gender construct. Terman and Miles supported the existing social values around “healthy” male and female norms conflating psychological health with socially mandated gender conformant behavior.

Terman and Miles (1936) also developed the Attitude-Interest Analysis Test which served as a model for the development of Scale 5 (masculinity-femininity) of the Minnesota Multi-Phasic Personality Inventory (MMPI). The development of Scale 5 for the MMPI was initially a project to identify male “sexual inverts” (homosexual men). Male inpatients seeking treatment for “homosexual tendencies” were compared to non-patient men who were inducted into the army pre-World War II (Peterson & Dahlstrom, 1992). A comparable sample of female “sexual inverts” was not achieved.

The item development of these early scales was achieved through comparing this sample of homosexual men versus presumably heterosexual men, and by comparing
assumedly heterosexual male and female non-patient responses. High levels of femininity were also correlated with several psychologically undesirable traits such as misuse of prescription drugs, hyper-criticalness and emotional lability. Masculinity was correlated with high self-confidence, honesty and openness to new experiences (Peterson & Dahlstrom, 1992). Masculine personality attributes and feminine personality attributes were conceptualized as binary opposites. After tallying masculine and feminine personality variables, a person could be ranked as to how feminine or masculine they were with masculinity for males and femininity for females being seen as psychologically healthy.

An early version of the Strong Vocational Interest Blank (Strong, 1936) included a masculinity-femininity measure. Campbell (1966) removed this measure. The California Psychological Inventory (CPI; Gough, 1957, 1988) included a femininity and masculinity measure. The CPI masculinity and femininity measure was used by Baucom (1976) to empirically develop a gender scale that conceptualized masculinity and femininity as separate variables. Using responses from the CPI, Baucom separated variables according to whether more than 70% males had answered in the same direction and less than 10% of females to create a masculine scale (MSC). More than 70% of females and less than 10% of males’ similar responses were used for the creation of the femininity scale (FMN). He then created a circumplex with four quadrants: stereotypical masculinity (high on MSC), stereotypical femininity (high on FMN), undifferentiated gender (low on MSC and FMN) and androgynous (high on MSC and FMN).

In the development and restandardization of the MMPI, gender role masculine (GM) and gender role feminine (GF) scales were developed (Peterson & Dahlstrom,
1992). The developers described the usage of the non-patient sample of 2600 age and racial-ethnic diverse participants to create the GF and GM. Using a similar development strategy as Baucom, the GM and GF was conceptualized as two separate scales with no overlapping items. This was a revision of the initial development of the M-F scale (Scale 5) which was operationalized as a single scale where scores were given to men and women in opposite directions.

*Androgyny and Expanding Notions of Gender/Sex*

Beginning in the 1960’s, coinciding with the advent of the women’s liberation movement, psychologists began to examine how gender was defined and utilized within the discipline. Researchers began to critically analyze the literature in sex differences and conduct meta-analyses to examine what differences, if any, existed. The research in sex differences is based in the conflation of biological sex and the social construction of gender. Sex difference research was used by feminist researchers trying to prove women’s equality to men, (Eagly, 1987, Hyde & Linn, 1986, Maccoby & Jacklin, 1974) and contrarily, evolutionary psychologists used gender measurement to hypothesize that women are naturally better at stereotypical feminine tasks while men are better at stereotypical masculine tasks (Buunk, Angleitner, Oubaid, & Buss, 1996).

The women’s liberation movement began a shift in the conceptualization of gender within psychology. This shift was powerful in challenging psychologists to examine the underlying social constructions, ideologies and mores embedded in gender measurement. To this day, the two most well known and widely used gender measures are the Bem Sex-Role Inventory (BSRI) (Bem, 1974), and the Personality Attributes Questionnaire (PAQ) (Spence, Helmreich & Stapp, 1974). These measures have framed
the research in gender (for a review, see Cook, 1987) and have been used in over 750 articles to date (Hoffman, 2000). Both of these scales began the process of introducing the concept of androgyny, and gender traits as variable within men and women as a way to expand traditional understandings of men, women, the masculine and the feminine (Bem, 1974, Spence, Helmreich and Stapp, 1974).

The BSRI consists of 20 stereotypically feminine traits, 20 stereotypically masculine traits, and 20 neutral traits, with the 20 neutral items serving as a social desirability check. The BSRI was developed through the author and her students generating an initial list of 200 adjectives that were positive and deemed masculine or feminine. Two hundred additional neutral characteristics were generated, half of which were positive and half were negative. Judges were then asked to rate the generated adjectives on a 7 point scale as to whether they were deemed more appropriate for one sex over the other. One hundred undergraduate students (half men, half women) at Stanford University served as judges. Bem then calculated which characteristics were significantly considered more desirable for men than for women and vice versa. She also compared the judges’ ratings of these general ratings with their ratings of what was considered appropriate and desirable for themselves. Through these analyses she found that the male and female judges were nearly equal in their perceptions of what characteristics were sex appropriate or inappropriate (Bem, 1974).

The BSRI uses a 7-point Likert type scale where the respondent indicates how much each adjective applies to themselves. Through these responses, each respondent is given a masculinity score, a femininity score and an androgyny score. A social desirability score can also be computed. The androgyny score was defined in the original
BSRI as the student’s t-ratio for the difference between the respondents masculine and feminine self-endorsement. Bem intentionally used a t-ratio rather than a simple difference score in order to assess significance level in difference of endorsement of feminine versus masculine traits and to allow researchers to compare different populations. Scores are calculated and participants are assigned to one of four gendered groups – masculine, feminine, androgynous or undifferentiated. A highly sex-typed masculine response indicates the rejection of feminine attributes and vice versa, while an androgynous response set indicates high endorsement of both masculine and feminine attributes.

Using undergraduate men and women’s perceptions of sex stereotypes as a basis to develop masculine and feminine traits, Bem created a highly reliable and usable measurement tool. One reason it is so highly reliable is the reification of social sex stereotypes – it is not an internal definition of what constitutes someone’s psychological experience of their masculinity femininity or androgyny. The attributes are socially defined and mirrored in the endorsement of items reflecting sex stereotypes.

The Personality Attributes Questionnaire (PAQ) (Spence, Helmreich, & Stapp, 1974) is another widely used and respected gender measurement tool. The PAQ consists of 55 bipolar items selected from an original pool of 130 that were generated from undergraduate students as representing masculine or feminine characteristics. Students were asked to describe the “ideal” man or woman. The scale has three subscales (masculinity, femininity, and masculinity-femininity) higher scores on the femininity subscale reflect characteristics that are more common in women, but socially desirable for both sexes, the masculine subscale is constructed similarly, while higher scores on the
masculine-feminine subscale reflect more socially desirable traits for men. The scales are conceptually supported as representing stereotypical masculine/instrumental or feminine/expressive traits. Respondents are asked to rate themselves on a 5 point scale on each attribute. Each respondent receives a score on each of the three scales.

As with the Bem, the endorsement of traits on the PAQ represents the degree to which respondents feel stereotypic masculine or feminine traits are consistent with their views of themselves. The major contribution of the BSRI and PAQ was the theoretical advancement of gender existing on a bipolar not unipolar scale where respondents could possess masculine and feminine traits in combination with each other rather than higher masculinity equaling the absence of femininity and vice versa. A key problem with their approaches lies in the construction of their scales and the reliance on mostly white middle class college students to define masculine and feminine traits. These traits end up representing a set of attributes that are reflective of a dominant hegemonic view of gendered traits void of cultural variations within the dominant culture and reifying the dominant paradigm of gendered representations.

In the concluding comments in the original article describing the development of the BSRI, Bem (1974) states that she hopes “the androgynous person will come to define a more human standard of psychological health.” In further writings, Bem expanded upon her ideas about gender, sex typing, and gender norms. Her work serves as the basis for much of the gender measurement development since. She and others have worked to establish the social construction and patriarchal origin of societal gender norms. Many scales focus on establishing the social injunctions of feminine and masculine behavior as well as acknowledging that men and women hold both masculine and feminine qualities.
In her book, *The Lenses of Gender* (1993), Bem explores the past research in psychology and the gender debate. Several of the premises that she explores are integral in understanding how the gender framework brings researchers to reproduce hegemonic notions of gender, sex and sexuality. Bem states, “the institution of male power depends for its survival on the construction of males and females whose gendered personalities mirror the different and unequal roles assigned to them in the social structure.” She is keenly aware of the impact of sociopolitical factors on the psychology of gender.

Spence & Helmreich’s Attitudes Towards Women Scale (ATWS) has become one of the most widely used measures of gender-role attitudes (Spence & Helmreich, 1972, Spence, Helmreich, & Stapp, 1973). The primary goal of the authors was to assess how people perceived the roles of men and women, and to assess the judgment placed on those who might challenge gender norms. As this scale was developed during the height of the women’s liberation movement, the items appear dated. The focus of the scale remains integral to measurement in the area of sex and gender as the ATWS initiated the exploration of how gender is constructed in society. This work is important in establishing that sex and gender are different, and propagating the idea that gender is constructed to some extent, and strictly monitored and enforced by societal norms.

The Attitudes Towards Men Scale (Iazzo, 1983) was developed to assess attitudes that women have regarding men and the male gender-role. A 4-point Likert type scale was used to assess attitudes across four major domains: marriage and parenthood, sexuality, work, and physical and personality attributes. All items are descriptive and use a male anchor. Although the ATMS evidenced adequate reliability and validity, it has not
been utilized often in research and is need of additional construct and discriminant validity (Thompson, Pleck & Ferrera, 1992).

*Men and Masculinity Studies*

The study of men and masculinity arose from the emergence of gender freedom and androgyny models of gender measurement. As women made gains within the domain of psychology, researchers recognized that along with neglecting women’s psychological experiences, psychology had not critically analyzed what constituted masculinity or how the masculinity norms may impact the mental health of men.

The study of men and masculinity has produced a distinguished canon of gender assessment development research. In their review of 11 masculinity measurements, Thompson, Pleck and Ferrera identify two divergent frameworks for assessing gender (1992). Trait based measurements assume that gender is biologically based whereas gender ideological approaches theorize that gender consists of socially constructed sets of characteristics. This distinction is not as clear-cut as it seems, the authors note, because the PAQ has been argued to measure personality attributes rather than gendered traits, while Bem asserts that her BSRI is measuring cognitive gender schemas rather than prescriptive gender role stereotypes.

Early masculinity measures that were developed include The Macho Scale (Villemez & Touhey, 1977), Attitudes Towards the Male Role Scale (Doyle & Moore, 1978), the Brannon Masculinity Scale (Brannon & Juni, 1984), the Attitudes Towards Men Scale (Iazzo, 1983), and the Male Role Norms Scale (Thompson & Pleck, 1986). These scales measured attitudes towards men and the embodiment of masculine norms.
Thompson et al (1992) also point out that many of these scales devised to measure masculinity are limited in the types of masculinity that they measure. Authors of scales tended to use a white middle class masculinity paradigm without naming this framework as such. As a result, many of these early masculinity scales tend to measure patriarchal masculinity ideology and not individual men’s internal sense of masculinity.

Another important distinction between gender measurements is between measures that assess internally experienced gender attitudes, ideology or beliefs, and external, prescriptive ideas about gender. Some measures look at conformity to gender norms, whereas others examine personal gendered expressions. Thompson, Pleck and Ferrera refer to this as “gender orientation” versus “gender ideology” thus articulating a difference between gender prescriptions managed by society and expressed by individuals, and a person’s own internal sense of their gendered experiences. Gender identity is defined as “an individual’s structured set of gender-related personal identities,” personal identities including “interests and abilities, relationships with specific other people, social categories and dimensions of affect and personality, and styles of behavior.” (Ashmore, 1990).

The concept of gender role conflict has been a well-researched area in men and masculinity studies (Good et al, 1994). The Gender Role Conflict Scale, developed by O’Neil, has been used in hundreds of studies and provides a rich framework for the exploration of the restrictiveness of gender norms on men’s gender expressions. The GRCS measures gender role conflict across different contextual domains, and within behavioral, cognitive and affective realms. O’Neill and his colleagues have hypothesized the emotional and psychological effects of experiencing gender role conflict, and have
explored the processes through which conflict occurs within multiple contexts. New scales have been developed since the GRCS, utilizing it as a theoretical foundation.

Good et al explored the psychometric properties and clinical implications. After analyzing the factor structure of the GRCS, Good et al found strong statistical support for the GRCS factors, with the exception of the Conflict Between Work and Family Relations Subscale (CBWFR). The authors explore several reasons for the instability of the CBWFR factor including the possibility that the items on this subscale do not measure conflict between family and work as seen through the lens of gender role but through the lens of stress or other life experiences. The authors did find that gender role conflict was found to be a strong predictor of psychological distress. Restrictive emotionality was found to be the strongest predictor of psychological distress in this sample of college counseling center male clients.

The Conformity to Masculine Norms Inventory (Mahalik et al, 2003) measures the degree to which men adhere to socially prescribed roles for men in western society. The gender role norms that are represented in the CMNI are representative of white middle class heterosexual male dominant cultural norms. The authors took several preliminary steps in identifying dominant masculine ideology. They reviewed relevant literature and prepared two focus groups of diverse men and women that met on a weekly basis for over eight months. The focus groups were to identify masculinity norms that were separate from feminine norms and to discuss these norms with friends and family to assess the applicability of these norms with a referent group. After identifying twelve masculine norms; Winning, Emotional Control, Risk-Taking, Violence, Dominance, Playboy, Self-Reliance, Primacy of Work, Power Over Women, Disdain for
Homosexuals, Physical Toughness, and Pursuit of Status, the authors completed an exploratory factor analysis using principal components analysis and an oblimin rotation. Based on these results, the Physical Toughness factor was removed, as it did not adequately load cohesively as a factor. Additional analyses revealed differences in men and women on 9 of the 11 subscales. Pursuit of Status and Primacy of Work did not produce significant differences between men and women participants, most likely due to the primarily middle class, college age sample who was receiving similar messages regarding success and status. The authors describe several potential uses for the CMNI including assessment, research and practice implications. An interesting possible research application would be the assessment of costs and benefits of gender role conformity.

The Gender Role Journey Measure (GRJM) (O’Neil et al 1993) was developed to assist individuals in exploring sexism and their gender role socialization. The authors describe their intent in developing a scale that would focus on the process of becoming aware of and challenging gender roles instead of the continued empirical focus on how gender roles are learned or socialized. The GRJM conceptually consists of five phases; Acceptance of Traditional Gender Roles, Ambivalence about Gender Roles, Anger, Activism, and Celebration and Integration of Gender Roles. After completing an exploratory factor analysis, O’Neill et al discovered three emergent factors from the five conceptual phases; Acceptance of Traditional Gender Roles, Gender Role Ambivalence Confusion, Anger and Fear, and Personal-Professional Activism. Activism and Acceptance were negatively correlated and shared 21% of the variance. The GRJM, although conceptually conceived to measure the process of internal and external gender role changes, measures three different phases of gender-role belief endorsement. An
important aspect of the GRJM is its incorporation of affect states in measuring gender role beliefs. The negative or positive affect in relation to gender role cognitions and/or actions attempts to incorporate an important dynamic to the construct of gender. The authors discuss conceptualizing sexism as a psychopathological construct that negatively impacts men and women (O’Neill et al, 1993).

Another area of research has been on social consequences for gender role transgressions (Sirin, McCreary, & Mahalik, 2004). Researchers have utilized the Social Status (SS) model to illustrate the social sanctions utilized to pressure men and women to conform to gender norms (Sirin, McCreary, & Mahalik, 2004). If men are perceived as feminine they lose social status, are at risk of being perceived as homosexual (and thus exposed to homophobia), and are viewed more negatively than a women who is perceived as masculine.

New Directions

As the research on sex differences pointed to less clear differences between men and women, researchers began to examine the assumptions that underlay gender. Based in the idea that gender is different from sex, and is a socially constructed variable, researchers began to examine attitudes towards gender.

As gender was initially defined uni-dimensionally, the current emphasis on recognizing the complexity of gender has left some researchers and theorists casting for appropriate models and terms. Deaux (1999) in her review of gender research, identifies a context specific paradigm for understanding the broadness and complexity of gender. She points psychologists towards seeing people as having gendered identities and that these identities exist within a person-environment specific context. She cites several theorists
who challenge psychologists to see gender as one axis of an interlocking matrix of
identities within a social context that distributes power unequally to different constructed

The Conformity to Feminine Norms Inventory (CFNI) (Mahalik et al, 2005) was
developed to assess the full range of socio-cultural injunctions communicated for women
in Western society. The authors held four separate focus groups consisting of female
identified graduate students, community members and undergraduate students which
identified 13 feminine gender role norms. The authors then held two focus groups
consisting of graduate students from counseling psychology to generate items for the
proposed scale. Following factor analysis, the authors identified 8 factors with a final
item set of 87 items. The 8 factors were; (1) develop friendly and supportive relationships
with others (Nice in Relationships), (2) pursuit of thinness (Thinness), (3) refrain from
calling attention to oneself (Modesty), (4) maintain the home (Domestic), (5) take care
and be with children (Care for Children), (6) invest self in romantic relationship
(Romantic Relationship), (7) keep sexual intimacy contained within one committed
relationship (Sexual Fidelity), and (8) commit resources to maintaining appearance
(Invest in Appearance). The authors hypothesized that the CFNI would measure
conformity to traditional gender role norms in the dominant culture of western society.
The authors found that the CFNI measured these norms and was that the CFNI was
negatively related to stages of increasing feminist identity development. Specifically, the
authors found that conformity to the norms of homemaking, childcare, and being married
were the variables related to passive acceptance of traditional gender roles. The authors
found a negative relationship between conformity to the Modesty norm and higher levels
of feminist identity. Other subscales were not associated with levels of feminist identity. The CFNI had large sample sizes of mostly white heterosexual college age students. As feminine norms are variable within age, culture and ethnicity groups, the study is limited in its applicability to different populations. Although the authors make the point that dominant hegemonic cultural norms are defined by the white heterosexual middle class majority, the impact of subculture norms may adjust levels of conformity or expression.

Hoffman (2000) makes the point that masculinity or femininity is really a measure of a sense of one’s maleness or one’s femaleness, their gender identity. Hoffman argues that it is more important to measure gender identity, gender self-confidence, and gender self-concept through gender self-acceptance and gender self-definition. Although based in a critique of the existing restrictive gender role identity measurement already in place, Hoffman’s argument that an individual’s gender identity can be internally determined seems short sighted in that it fails to acknowledge how powerful social norming in gender rigidity impacts an individual’s ability to develop their own standards of masculinity-femininity. The idea of gender identity representing security in one’s self as a male or female tends to reinstate a normative binary of gender appropriateness.

Gender dimorphism is a theme that continues as researchers use gender or sex as a variable in studies. Using gender as an independent variable reinforces that there are inherent differences between men and women, and that there are only two categories of gender and sex – men and women/ masculine and feminine. This excludes other gender identifications such as genderqueer or transgender while also reifying stereotypical gender/sex roles and norms.
In examining the relationship between gender role orientation/identity and attitudes towards gender equality/feminism, (Toller, Suter, & Trautman, 2004) researchers found support for the hypothesis that individuals on either end of the gender binary (highly masculine men and highly feminine women) will not identify with feminism or gender equality ideals in a sample of college students (Toller, Suter, & Trautman, 2004). This finding, in combination with the other findings described above, brings into focus the differing beliefs, psychological adjustments, distress levels, ideologies and other variables that are contained within the study of gender role orientation.

*Gender as a Social Construct*

Kessler and McKenna’s landmark study “Gender: An Ethnomethodological Approach” challenged psychologists to reappraise how the terms gender and sex were used as a conceptual framework for analyzing men and women. The authors posit that sex is not a biological reality or fact, but in actuality is just as much a social construction as gender. They used cross cultural data and the case of intersex infants to support their conceptualization of sex as well as gender as a continuum. This theme is also restated in the feminist biologist Anne Fausto-Sterling’s work on sex identity.

Theorists from feminist and queer studies are engaging a paradigmatic shift in how the constructs of gender and sex are conceptualized. As discussed, gender has been conflated with sex in the past, and both gender and sex have been defined as uni-dimensional binary constructs. Psychologists such as Milton Diamond and feminist theorists, such as biologists Anne Fausto-Sterling, argue gender is multifaceted and complex; it is a separate construct from sex and exists on a continuum of experience.
There are gray areas that exist within sex (Fausto-Sterling, 2000) including intersex individuals, individuals with various combinations of chromosomes, transsexual people, and other sexes besides men and women. As for gender, instead of being the socially constructed physical, social, and emotional corollary of sex, gender is a set of diverse attributes that are both inherently determined and socially shaped. This is an important distinction given the debates within queer, feminist and Transgender communities about the nature of gender identity. Some theorists, grounding themselves in the epistemological challenges of Foucault and Lacan (Foucault, 1978), see gender as a frontier of post-modernist debate. If gender is wholly constructed, the analysis of power as maintained through the construction of language and hegemonic discourse deconstructs the entire field of gender research in psychology. Some theorists, such as Judith Butler, and Foucoulidian thinkers, go as far as to state that gender is wholly a construct that individuals falsely believe is an internally felt phenomenon, but in reality is an external concept that we perform, mimicking norms of femininity or masculinity that are created through socialization (Butler, 1990).

Diamond (2000) makes a clear distinction between sex identity and gender identity. He defines sex identity as referring to a person’s sense of their maleness or femaleness, whereas gender identity is an individual’s sense of their masculinity, femininity, androgyny, transgender identity or a range of possible gender identities (Diamond, 2000). Gender is the socially prescribed attributes deemed appropriate for the two recognized sexes. Gender is the lens through which we interpret sex differences/similarities. Most people embody a complex range of gendered qualities and gender identity refers to the identity that is ascribed to the gender presentation people
present. People can be aware of their gender identity or if they are highly gender conformant, may not think about their gender expression as a gendered identity at all. Although there may be physiological or characteralogical differences between men and women, these are moderated by social and cultural schemas for interpreting these differences. Thus they are ascribed meaning through the cultural lens (i.e. men are seen as unemotional or rational whereas women are seen as more emotionally adept, but these are given different values within our society) (Bem, 1993).

In this paradigmatic shift, the concepts of gender expression comfort and conformity help make visible the processes through which gender hegemony is upheld in dominant society. Gender expression is impacted by contextual and internal factors. Some individuals have high gender norm internalization and thus constrict their behaviors and grooming (i.e. external markers of gender) in order to adhere to cultural mandates. Others are perceived through their external behaviors/grooming/identifiers as gender non-conformant. These people may have low or high levels of gender norm internalization and thus various levels of comfort or conformity of their gender expression. Gender expression differs from gender internalization in that gender internalization is the internal process, whereas gender expressions are the external markers of gender conformity/nonconformity. Gender expression consists of the expression (physical, social interactions, grooming, and behaviors) of social and cultural roles associated with biological sex. Expression may be congruent with social mandates or incongruent, or a mix. Gender expression is moderated by socialization, internalization and context.
Measurement of Gender

The current scales that propose to measure gender, for the most part assess a variety of variables that most likely approximate gender norms and stereotypes, rather than an internal sense of gendered expression. For example, the Macho Scale (Villemez & Touhey, 1977) is composed of items that are based on anti-feminine attitudes and traditional notions of male dominance. In essence, the Macho Scale does not measure gender itself but instead measures patriarchal masculinity ideology.

As early as the 1970’s psychologists began to question the unidimensionality of gender. Many self report measures prior to the development of the BSRI and PAQ were based in the methodology of a unidimensional scale where presence of more masculine traits equaled the absence of feminine traits and vice versa. “Item analysis shows that separate masculinity and femininity dimensions show only moderate correlation with each other and that both possibly together with a bipolar masc-fem scale are necessary to create a conceptually accurate inventory.” (Vestewig & Stericker, 1978).

In summary, many of early attempts at operationalizing gender as a construct have served to either reify existing cultural expectations of gendered expressions, or have measured discrete aspects of gender in the form of gender roles, sex roles, or gender conformity. There is a clear need for continued exploration of gender as a construct and further inquiry into how individuals experience their gender on cognitive, behavioral and affective levels.

Research Questions and Hypotheses

The development and validation of the Comfort and Conformity of Gender Expression Scale (CAGES) included three studies. Study 1 consisted of the scale
construction and exploratory factor analysis. Study 2 provided initial estimates of convergent and discriminant validity of the instrument using two existing gender conformity scales. Study 3 assessed the test-retest reliability of the instrument.

In Study 1, using the previously described existing literature as a foundation, items were generated and tested to assess the level of comfort/discomfort with conformity/nonconformity within the hypothesized domains of gender expression: cognition, behavior, and affect. It was hypothesized that the items would measure the levels of conformity or comfort of gender expression on two continua. The researcher hypothesized that there would be two continua; one assessing comfort with conformity and another assessing comfort with nonconformity in gender expression (see figure 1). These two overlapping continua would assess individuals’ different levels of comfort and discomfort with conformity and nonconformity in gender expression through thoughts, behaviors and feelings.

Thus, the CAGES was hypothesized to assess the internal cognitive, affective and behavioral processes in comfort or discomfort with conformity or nonconformity in expressing gender, specifically, comfort and discomfort with conformity/nonconformity to gender norms. Given the pressures of societal gender role mandates (Mahalik et al, 2003, 2005), it was hypothesized that participant responses would range on a continuum of perceived comfort/discomfort with non-conforming freedom of gender expression to comfort/discomfort with conformant gender expression. Gender expression was hypothesized to be impacted by level of awareness of gender and gender role norms. Gender expression was defined as the expression of social and cultural roles associated with biological sex, how we communicate and experience gender through behavior,
thoughts and feelings. The items were hypothesized to fit a four factor structure with the following hypothesized factors; Comfort-Conformity-Public, Comfort-Nonconformity-Public, Discomfort-Nonconformity-Public, and Discomfort-Conforming-Public.

In Study 2, convergent and discriminant validity estimates were explored. In an attempt to demonstrate convergent validity of the CAGES, it was hypothesized that the CAGES would moderately correlate to the Conformity to Masculine Norms Inventory (Mahalik, 2003) and the Conformity to Feminine Norms Inventory (CFNI) (Mahalik, 2005). Since the CFNI and CMNI assess conformity, it was hypothesized that there will be a moderately high relationship between these scales and the CAGES.

Study 3 attempted to establish test-retest reliability of the CAGES. It was hypothesized that test-retest reliability would be high across a two-week period.
CHAPTER 3: METHOD

The preceding review and critique of the literature provided the rationale for the development of a scale that measures levels of gender expression within a sociocultural context, the Comfort and Conformity of Gender Expression Scale (CAGES). This chapter provides a description of the methods that were used to psychometrically validate the CAGES. The chapter will be divided into three sections outlining three separate studies. Study 1 will focus on the development and factor structure for the CAGES. Study 2 will focus on establishing convergent and divergent validity for the completed CAGES from study 1. Convergent validity was established using two previously validated gender scales that measure expression of gendered attributes, the Conformity to Masculine Norms Inventory (CMNI) and the Conformity to Feminine Norms Inventory (CFNI). Divergent validity was established using the Balanced Inventory of Desirable Responding (Paulhus, 1991). Study 3 attempted to establish test-retest reliability using the final scale developed from Studies 1 and 2.

Study 1: Scale Construction and Exploratory Factor Analysis

Item Development

Items were generated for each area after a thorough review of both the literature and measures related to gender norms internalization, gender expression and gender conformity. Following the guidelines for rational item development as described by DeVellis (1991), several preliminary steps were taken to ensure a fully realized scale that assesses comfort/discomfort with conformity/nonconformity in gender expression. First, an initial pool of items was developed based upon the complex gender theoretical
frameworks within psychology and Women’s and Gender Studies. The item pool included items that examine gender expression through behavioral, contextual, cognitive and affective components. The author generated items based in existing theory that address these components. After identifying the main factors which composed the subscales of the scale, the author generated a maximum number of items for each factor, ending with 41 initial items across four hypothesized factors.

*Expert Evaluation of Items*

Experts in the field of gender theory and research were asked to assess the content validity of the items. The experts were interdisciplinary, including faculty and doctoral students in counseling psychology and sociology, master’s level students from counseling psychology and social work, and undergraduates and community members who are active in gender advocacy and gender issues. An initial focus group of 6 graduate students in counseling psychology, sociology, and social work, along with 3 undergraduate women’s and gender studies students, and 3 community members active in LGBT and gender advocacy work explored different understandings of the term gender, and how individuals used visible cues from appearance or behavior to determine gender expressions. The experts also participated in a group discussion and brainstorming session designed to reach unarticulated processes of gender identification, conformity, and the processes of gender judgment. The researcher also met individually with acknowledged 2 different experts in the fields of gender in psychology to receive feedback on item construction, as well as utilized feedback from the dissertation proposal meeting process involving five faculty advisors. Additionally, the researcher distributed the items to 1 PhD level counseling psychology practicing psychologist working in the
filed of gender and sexuality, 3 interdisciplinary graduate students, and 1 sociology and women’s gender studies faculty member, for additional feedback.

The experts was asked to examine the items to (a) determine whether they are reflective of the critical issues that have been hypothesized, (b) ensure coverage of the content domains, (c) eliminate unnecessary items, (d) revise any confusing items, (e) provide general feedback that would assist in developing items, (f) ensure that the generated items are representative of comfort and conformity of gender expression and (g) develop any additional items that may be needed. The experts then examined the items in order to match the items to the proposed content domains of gender expression comfort and conformity in behaviors, cognitions and affect. The experts was asked to evaluate these dimensions to determine if they fully articulate the proposed content areas. This data was compiled and utilized to make revisions to the CAGES items.

*Exploratory Factor Analysis (EFA)*

After revisions to improve clarity, the final item pool was administered to a large national sample diverse in class, race/ethnicity, age, gender identity, sexual orientation and ability status. The sample was recruited via the Internet to conduct the exploratory factor analysis and obtain initial reliability and validity estimates. It was hypothesized that the CAGES would demonstrate adequate reliability and validity, and would measure several distinct (but correlated) areas involved in the process comfort and conformity of gender expression. Discriminant validity was hypothesized to be evidenced by low correlations between social desirability and CAGES total and subscale scores.

*Participants*
Recruitment of participants was identical for Study 1 and Study 2 and was targeted at including as diverse and generalizeable a sample as possible. A minimum of 200 participants was sought for Study 1. In order to determine the adequacy of the items for exploratory factor analysis (EFA) a minimum of 5-10 participants per item is recommended to be recruited (Worthington & Whittaker, 2006). New guidelines for sample size in scale construction suggest that the salient criteria for determining sample size should be the hypothesized number of factors and sizes of correlations among the set of items. If there are smaller numbers of factors theorized, smaller participant pools are acceptable (Worthington & Whittaker, 2006). As the CAGES was developed for use with a wide range of populations, target characteristics of the sample are designed to be representative of the diversity of the US population. In regards to gendered characteristics, individuals who are gender conformant, gender non-conformant, gender unaware and gender hyper-aware were sought out as possible participants. Efforts were made to represent a diversity of racial-ethnic minorities, sexual orientations, cultural affiliations, spiritual orientations, ability statuses and other demographic factors consistent with U.S. census population records through pursuing diverse outlets for dissemination of the survey, and oversampling underrepresented groups. As gender expression was hypothesized to be a construct that exists on two continua, it was hypothesized that the CAGES will be able to be utilized with a general population who embody a range of gendered expressions and identities.

Participants for this study were recruited through email, flyers, and word-of-mouth. The online data collection software SurveyMonkey was utilized for data collection. A total of 469 persons began the survey. There were several blank (n = 11)
and missing data (n = 88) entries that were removed. After data cleaning for malicious or duplicate responses, and invalid responses, (n = 14), 356 usable cases were identified. Of these 356 cases, 59.4% identified as female (n = 211), 28.7% identified as male (n = 102), 7.3% identified as transgender (n = 26), 2.8% identified as genderqueer (n = 10), 1.7% identified as other (n = 6), and one person declined to respond. 61.8% (n = 220) of the participants identified as straight/heterosexual, 9.3% (n = 33) identified as lesbian, 5.6% (n = 20) identified as gay, 8.1% (n = 29) identified as queer, 14.0% (n = 50) identified as bisexual, 0.8% (3) identified as other, and one person declined to answer. As for racial/ethnic representation, 83.7% (n = 298) of the sample identified as white/European-American, 2.5% (n = 9) identified as biracial/multiracial, 2.5% (n = 9) identified as Latino/a/Hispanic, 7.9% (n = 28) identified as African-American/Black, 0.6% (n = 2) identified as South Asian/Indian, 1.1% (n = 4) identified as Asian-American, 1.1% (n = 4) identified as Native American, and 0.3% (n = 1) identified as other.

**Instruments**

Demographics. Demographic questions were created and included to measure participant characteristics. Gender/Sex identity, Race/Ethnicity, Sexual Orientation Identity, Age, and Educational level were used to assess demographics (See Appendix A for the questionnaire and Table 1 for demographic data for Study 1).

The Comfort and Conformity of Gender Expression Scale (CAGES). The initial version of the CAGES included 41 items rated on a 4-point Likert scale ranging from 1 to 4 (see Appendix A for scale item exemplars). Participants indicated their current level of agreement with statements about their own personal level of comfort or discomfort with
transgression or conformity in gender expression (1 = strongly disagree, 2 = disagree, 3 = agree, 4 = strongly agree). The scoring was reversed in subsequent analyses to be (1 = strongly agree, 2 = agree, 3 = disagree, 4 = strongly disagree).

*Balanced Inventory of Desirable Responding* (BIDR; Paulhus, 1991). The BIDR (Appendix C) is a 40-item scale that measures self-deceptive positivity (SDE) and impression management (IM). These two constructs measure the levels to which participants attempt to present themselves in a favorable light. The BIDR-IM subscale was used in the current study to provide an estimate of the validity of the CAGES. A 7-point Likert scale (1 = not true, 4= somewhat true, 7= very true), is used, and participants rate their level of agreement to each item, with a total test range of 0-20. Participants who score high may be exaggerating desirable responses on the subscales, thus only items with scores of 6 or 7 are counted, with 20 points possible. Higher scores indicate higher social desirability. The scoring key is balanced, with 10 items negatively scored. A Cronbach’s alpha reliability of .83 was reported for the total BIDR score, and ranged from .75 to .86 for the IM scale. For Study 1 the BIDR-IM internal consistency was .74. Test-retest correlations over a 5-week period have been reported at .65 for the IM scale. Both subscales have shown consistent convergent reliability estimates with other lie and validity measures. Examples of the items are “I don’t care to know what other people really think of me,” and “When I hear people talking privately, I avoid listening.”

Procedure

The study was announced in email notices that were sent to university staff, student and faculty listservs, to community and national listservs and advertised within the community. It was indicated in the e-mail that participation in the research project
was voluntary and anonymous. Upon choosing to participate, respondents were asked to click on the URL within the text of the e-mail to link to the online survey. Reminder e-mails were sent one week following the original request. (See Appendix D for email recruitment notice examples)

Participants were recruited through university student, faculty and staff listservs, university and community organizational listservs, and national organizational listservs such as the American Psychological Association of Graduate Students listserv. The primary researcher also utilized contacts in private business to gain access to general employment listservs. Specifically, the researcher sent the recruitment email to persons within three large Midwestern based private companies and asked the contacts to forward the email to personal contacts and listservs.

For internet recruitment of participants, the announcement message included a request for voluntary participation from a diverse national sample to complete the CAGES through clicking or pasting a webpage address into the participant’s web browser. The message described the study as an investigation of gender expressions. Participants were directed to a webpage containing an informed consent form (Appendix E) that explained the transmission of survey data via the Internet. Agreement was indicated by clicking “Yes” to a text box reading, "I have read this page, and I would like to take the survey." Participants were then directed to the survey page, which included the CAGES items and a demographic information form (see Appendix F). After completing the survey, participants were directed to a debriefing form (see Appendix G) that explained the hypotheses of the study. They were given the contact information of the primary investigator for any additional comments or feedback. Participants were
informed of a lottery in which two participants were randomly selected from all voluntarily submitted emails at the end of data collection. The selected participants were contacted via e-mail, and were awarded $50 gift certificates to Amazon.com each.

Internet data collection, although it provides the researcher with access to an extensive participant pool, also decreases the amount of control the researcher has over assessment environments. Participants could possibly submit their completed survey more than once, be less attentive, or engage in malicious responding. Keller and Lee (2003) discuss some of the ethical issues to be considered when engaging in internet data collection. As suggested by Birnbaum (2004), several strategies were employed to defend against multiple or malicious responses, and to ensure anonymity. Instructions were clear in asking people to participate only once, and included a statement indicating that participants may only be considered for the incentive once; multiple responses would not increase chances in the $50 lottery. In addition, the IP addresses and browser characteristics of the submitted data were retained and analyzed to assess for multiple responses. In addressing anonymity, along with informed consent detailing the data collection procedures and storage of data, no identifying information was collected, and the email that was sent to enter the lottery was not associated with an IP address or participant data.

Study 2: Reliability and Validity Tests of the CAGES

The purpose of the second study was to obtain reliability and validity estimates for the CAGES. It was hypothesized that the CAGES would demonstrate convergent validity with the Conformity to Masculine Norms Inventory (CMNI) and the Conformity to Feminine Norms Inventory (CFNI). The CFNI and CMNI both measure an
individual’s level of conformity to masculine or feminine norms as defined by western societal gender role norms. The CAGES was hypothesized to correlate moderately with these two empirically validated scales, but to differ in respect to measuring the process of gender expression regulation rather than conformity to gender role norms.

Participants

A minimum of 150 participants were sought for Study 2. As with Study 1, recruitment of participants attempted to include as diverse and generalizeable a sample as possible. As the CAGES was developed for use with a wide range of populations, efforts were made to ensure that the sample was representative of the diversity of the US population. Specifically, with regards to gendered characteristics, individuals who are gender conformant, gender non-conformant and people who are very aware and very unaware of gender were be sought out as potential participants. Similar efforts were made to represent a diversity of racial-ethnic minorities, sexual orientations, cultural affiliations, spiritual orientations, ability statuses and other demographic factors as in Study 1.

Participants for this study were recruited through email, flyers, and word-of-mouth. The online data collection software SurveyMonkey was utilized for data collection. A total of 197 persons began the survey. There were several blank (n = 5) and missing data (n = 13) entries that were removed. After data cleaning for malicious or duplicate responses, and invalid responses, (n = 3), 176 usable cases were identified. Of these 176 cases, 58.0% identified as female (n = 102), 36.4% identified as male (n = 64), 1.7% identified as transgender (n = 3), 2.3% identified as genderqueer (n = 4), and three persons declined to respond. 64. 8% (n = 114) of the participants identified as
straight/heterosexual, 9.1% (n = 16) identified as lesbian, 8.5% (n = 15) identified as gay, 9.1% (n = 16) identified as queer, 6.3% (n = 11) identified as bisexual, and four persons declined to answer. As for racial/ethnic representation, 78.4% (n = 138) of the sample identified as white/European-American, 6.3% (n = 11) identified as biracial/multiracial, 2.8% (n = 5) identified as Latino/a/Hispanic, 5.7% (n = 10) identified as African-American/Black, 1.7% (n = 3) identified as Asian-American, 2.8% (n = 5) identified as Native American, and four persons declined to answer.

**Instruments**

*Demographics.* Demographic questions were created and included to measure participant characteristics. Gender/Sex identity, Race/Ethnicity, Sexual Orientation Identity, Age, and Educational Level were used to assess demographics (See Appendix A for the questionnaire and Table 2 for demographic data for Study 2).

*Comfort and Conformity of Gender Expression Scale (CAGES).* The revised version of the CAGES, with 24 total items across four factors based on the results from the EFA in Study 1, was given to participants. Participants were given the same instructions used in Study 1. Internal consistency was found to be .86 for the total scale for the current study, and .88 for Factor 1, .84 for Factor 2, .80 for Factor 3, and .86 for Factor 4.

*The Conformity to Male Norms Inventory (CMNI)* (Mahalik et al, 2003). The CMNI measures the degree to which men adhere to socially prescribed masculine gender roles in western society. The gender role norms utilized in the CMNI are representative of white middle class heterosexual male dominant cultural norms. The CMNI was developed through several intensive focus groups and targeted item generation in order to
attempt to fully capture the range of masculine gender-role norms. The CMNI consists of 11 masculine gender role norms: Winning (WIN), Emotional Control (EMOC), Risk-Taking (RISK), Violence (VIOL), Dominance (DOM), Playboy (PLAY), Self-Reliance (SELF), Primacy of Work (WORK), Power Over Women (POW), Disdain for Homosexuals (HOMO), and Pursuit of Status (POS). The items consist of cognitive, behavioral, and affective components of these norms, such as, “In general, I will do anything to win,” “I feel comfortable trying to get my way,” and “I would be furious if someone thought I was gay.” Items are scored on a 4-point Likert scale ranging from Strongly Agree to Strongly Disagree. Men were found to score significantly higher than women on nine of eleven subscales (WIN, EMOC, RISK, VIOL, POW, DOM, PLAY, SELF, and HOMO). Coefficient alpha for the CMNI was found to be .92 for the total score, while individual subscale were found to be the following; WIN = .88, EMOC = .91, RISK = .82, VIOL = .84, POW = .87, DOM = .73, PLAY = .88, SELF = .85, WORK = .76, HOMO = .90, and POS = .72. Internal consistency for the CMNI total scale was .90 for the current study. Individual subscale internal consistencies for the current study were as follows; WIN = .82, EMOC = .91, RISK = .86, VIOL = .87, POW = .77, DOM = .66, PLAY = .89, SELF = .87, WORK = .81, HOMO = .94, and POS = .71

*The Conformity to Feminine Norms Inventory (CFNI)* (Mahalik et al, 2005). The CFNI measures the degree to which women adhere to socially prescribed feminine gender role norms in western society. The gender role norms utilized in the CFNI are representative of white middle class heterosexual female dominant cultural norms. The CFNI was developed through several intensive focus groups and targeted item generation in order to attempt to fully capture the range of feminine gender-role norms. The CFNI
consists of 12 identified feminine norms; Have Nice Relationships (HNR), Thinness (THIN), Modesty (MOD), Domestic (DOMS), Involvement with Children (IWC), Involvement in Romantic Relationships (IRR), Sexual Fidelity (SEXFID), and Invest in Appearance (IIA). Examples of items include, “I feel good when others know I care,” and “I would be happier if I were thinner.” Items are scored on a 4-point Likert scale ranging from Strongly Agree to Strongly Disagree. Women were found to score significantly higher on the CFNI total scale score and on six of the eight subscales (HNR, SEXFID, DOMS, IIA, IWC, and THIN). Coefficient alpha for the total CFNI score was found to be .88, while individual subscales reported the following internal consistencies: HNR = .84, THIN = .90, MOD = .82, DOMS = .84, IWC = .92, IRR = .77, SEXFID = .85, and IIA = .82. In the current study CFNI, total scale internal consistency was found to be .85. Individual subscales (with both men and women completing the scale) were found to have the following internal consistencies: HNR = .76, THIN = .86, MOD = .69, DOMS = .81, IWC = .94, IRR = .81, SEXFID = .85, and IIA = .81.

Procedure

As with Study 1, data was collected via the Internet through email solicitation. Identical efforts to those in Study 1 were made in Study 2 to maintain anonymity of participants, to prevent malicious responding, and to inform participants of any potential risks in the study. In addition to the CAGES, participants were asked to complete one of two separate gender measurement scales, the CMNI or the CFNI, to evaluate convergent validity. The CFNI and CMNI were administered to all sexes. All participants completed the CAGES, but participants were randomly assigned to either the CFNI or CMNI.
Participants were again recruited through university student, faculty and staff listservs, university and community organizational listservs, and national organizational listservs such as the American Psychological Association of Graduate Students listserv. The primary researcher also utilized contacts in private business to gain access to general employment listservs. Additionally, advertisements at community events and flyers advertising the availability of the study online were placed in several cities/towns across the Midwest. Efforts were made to gather a representative sample of men, women, and transgender participants, reflecting the current sex demographics of the U.S.

Study 3: Test—Retest Reliability Estimates

The purpose of Study 3 was to provide additional reliability estimates for the CAGES. It was hypothesized that the CAGES would have adequate test-retest reliability.

Participants

In order to gain access to participants that could be retested in similar conditions, participants were recruited from undergraduate courses in psychology and sociology. Participants were informed of the nature of the study, the voluntary nature of their participation, and the following re-test in two weeks.

A total of 63 usable paired surveys were completed. Several surveys could not be included, as 16 surveys did not have matched pairs, due to absences or non-completion of the second data collection. Seven surveys were not used due to incomplete or malicious responding. Surveys with both validity items incorrect were not used. This left 56 paired, complete surveys. Of these 56 surveys, 23 were missing demographic data due to an error in data collection. These 23 participants did not receive the demographics page of the packet. Of the remaining 33 cases that did receive a demographics form, 54.5% identified
as female (n = 18), 45.5% identified as male (n = 15). A total of 100.00% (n = 33) of the participants identified as straight/heterosexual. As for racial/ethnic representation, 78.8% (n = 26) of the sample identified as white/European-American, 3.0% (n = 1) identified as biracial/multiracial, 9.1% (n = 3) identified as African-American/Black, 6.1% (n = 2) and 3.0% (n = 1) identified as other.

**Instruments**

*Demographics.* Demographic questions were created and included to measure participant characteristics. Gender/Sex identity, Race/Ethnicity, Sexual Orientation Identity, Age, and Educational level were used to assess demographics (See Appendix A for the questionnaire and Table 3 for demographic data for Study 3).

*Comfort and Conformity of Gender Expression Scale (CAGES).* The revised version, with 24 items and four factors, was given to participants. Participants were given the same instructions to complete the revised CAGES as were used in Studies 1 and 2.

**Procedure**

Study 3 investigated the test-retest reliability of the CAGES by administering the CAGES to a sample twice over a 2-week period. Undergraduate students in two humanities courses at a large Midwestern public university were recruited to fill out a survey during class. Students were informed that their participation was voluntary and that their information would be kept anonymous. The survey packets included an informed consent form detailing the risks of participation and the contact information of the primary researcher. Participants were instructed to mark the first page of their survey packet with a code that would be used to match their first set of responses with the second retest responses in two weeks. Additionally, potential participants were informed
that the code was used in order to maintain their anonymity. The primary researcher and research assistants used a research protocol (see Appendix J for text of protocol). The primary researcher or a research assistant returned to the classroom in two weeks to collect the retest data and repeated the same instructions.
CHAPTER 4: RESULTS

The goal of the current study was to construct a scale that would measure levels of comfort with conformity or nonconformity in individual’s gender expression in the public sphere. Additionally, the researcher sought to construct a scale that was inclusive of a continuum of gender/sex identities, and inclusive of the impact of social context on gender expression. This chapter describes in detail the statistical analyses used to complete the process of developing a new gender expression scale. First, the processes of data cleaning and screening are reported. The chapter will then detail the results of the exploratory factor analyses conducted. The factor structures, the names of the factors, and internal consistency estimates are provided. Second, the results from the validity studies are presented, including correlation coefficients between the CAGES and the CMNI and CFNI for women and men on all three measures, for men on the CMNI and for women on the CFNI separately. Lastly, the results of the reliability study are given, specifically the correlation coefficients for the CAGES given twice over a two-week period.

Study 1: Scale Construction and Exploratory Factor Analysis

This section will describe in detail the process of factor analysis for Study 1. Decisions for limiting the factors, removal and retention of items are discussed in order to provide a thorough explanation of all decisions.

Using a total of 356 cases, the responses from the final sample were analyzed to test the factor structure of the CAGES. The responses were analyzed after examining the items for any flaws in the data collection including assessing the range of possible
answers given and reverse scoring the appropriate items. The data was screened for
duplicate cases, malicious responding, and for data sets with missing data. The data was
checked for accuracy in scoring and data entry clerical errors. Reverse scored items were
checked to maintain accuracy of results.

A total of 41 items were used in the EFA through a principal-axis factor analysis. Before proceeding with the principal-axis factor analysis, the Bartlett’s test of sphericity (1954) and the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy (Kaiser, 1974) were examined. An Exploratory Factor Analysis (EFA) was conducted in order to
identify underlying constructs of the CAGES. The Kaiser-Meyer-Olkin (KMO) measure
or the Bartlett’s test of sphericity was used to assess the adequacy of sample size of the
data. It is recommended that in sample sizes with a ratio of 5:1 the Bartlett’s be used to
assess the correlation matrix (Worthington & Whittaker, 2006). The KMO assesses the
existence of true factors over random correlations between subsets of the variables. It is
recommended that KMO values larger than or equal to .60 be utilized to determine a
good factor analysis (Tabachnick & Fidell, 2001 as cited in Worthington & Whittaker,
2006). Bartlett’s test of sphericity was significant (p <.0001), which indicates that the
factor matrix is adequate for analysis. The KMO yielded a value of .911 indicating that
the sample size was large enough to evaluate the factor structure. Decisions regarding the
number of factors to retain for the final scale were based on examination of the
eigenvalues of each factor, variance accounted for by each factor, and the number of
items loading on each factor. Review of the correlation between factors was also used to
determine the number of factors with high correlations (> .5) suggesting that the factors
might measure similar constructs. Additionally, retention of items were based on
communalities (> .30), cross-loadings (when the factor loading differences were < .15), and conceptual cohesion. Because the factors for this scale are likely to be correlated, both an orthogonal and an oblique rotation were conducted prior to the analysis. In order to assess the factor structure of the CAGES, Scree plot, eigenvalues (variance in a set of variables explained by the factor) and communalities (amount of variable’s variance explained by the factor structure) were assessed. Eigenvalues greater than 1 were explored as possible factor solutions, however if multiple factors emerged, the Scree plot was used to identify possible factor structure.

After completing the rotation, output was assessed for additional low loading items. Items with low factor loadings (below .30) and items that loaded on multiple factors (factor loadings closer than .15) were removed. Factors that consisted of less than three items were eliminated. Extraneous items that accounted for little variance were eliminated. Any items that had communalities below .30 were considered for removal from the data set. Correlations among the factors were examined to assess for shared or individual variance. The factor analysis process is discussed below.

A principal axis factor analysis was performed on the 41 items of the preliminary CAGES (see Appendix A). Eigenvalues were set at 1.0 to assess the initial factor structure of the data. The original examination of the Scree plot and the amount of variance accounted for by the items suggested the possibility of five factors. The eigenvalues for each of the factors were 11.94 (factor 1), 3.60 (factor 2), 2.43 (factor 3), 1.89 (factor 4), and 1.53 (factor 5). The amount of variance accounted for by each factor was 29.13% (factor 1), 8.79% (factor 2), 4.94% (factor 3), 1.89% (factor 4), and 1.53% (factor 5). The eigenvalues, communalities and amount of variance accounted for by each
factor was examined to determine the next step in the EFA. The pattern matrix failed to converge without a limited number of factors. After examining the scree plot and initial eigenvalues, it was decided to limit the EFA to 5 factors. Following this, a new EFA was performed specifying a five-factor solution.

The data was reanalyzed specifying a 5-factor solution with an oblique rotation. This rotation method was chosen because it was believed that the factors would be correlated due to the common underlying construct. The eigenvalues for each of the factors were 11.94 (factor 1), 3.60 (factor 2), 2.43 (factor 3), 1.89 (factor 4), and 1.52 (factor 5). The amount of variance accounted for by each factor was 29.13% (factor 1), 8.89% (factor 2), 5.94% (factor 3), 4.61(factor 4), and 3.72% (factor 5). The eigenvalues, communalities and amount of variance accounted for by each factor were examined to determine the next step in the EFA. Additionally, the pattern matrix was examined for item cross-loadings and the number items loading on each factor. In an effort to approximate simple structure, the pattern matrix was examined to identify the strength of the loadings on each factor and the absence of inter-factor correlations. The pattern matrix indicated multiple cross-loadings on several factors. After examining the items for conceptual clarity, the communalities to identify low communality items, and the initial eigenvalues, it was decided a four-factor scale might more closely approximate simple structure. Following this analysis a new EFA was performed specifying a 4-factor solution.

Following the examination of a 5-factor solution a third EFA was performed specifying a 4-factor solution. Again, an oblique rotation was specified because it was believed that the factors would be correlated due to the common underlying construct.
The eigenvalues for each of the factors were 11.94 (factor 1), 3.60 (factor 2), 2.43 (factor 3), and 1.89 (factor 4). The amount of variance accounted for by each factor was 29.13% (factor 1), 8.89% (factor 2), 5.94% (factor 3), and 4.61% (factor 4). The pattern matrix failed to converge in under 25 iterations. It was decided to explore a 3-factor solution given the failure of the four-factor solution to converge and the indication of the scree plot of the presence of three to four strong factors.

A fourth EFA was conducted, specifying a 3-factor solution. Eigenvalues for each of the factors were 11.94 (factor 1), 3.60 (factor 2), and 2.43 (factor 3). The amount of variance accounted for by each factor was 29.13% (factor 1), 8.79% (factor 2), and 5.94% (factor 3). After examining the five, four, and three factor solutions, percentage of variance explained, eigenvalues, Scree plot, and conceptual interpretation, it was decided to pursue the possibility of a four-factor solution. Results indicated several items needed to be removed due to low communalities (below .30). Items 9, 10, 13, 14, 25, 29, 32, and 40 were removed.

With 33 items remaining, an EFA was performed using oblique rotation and specifying 4 factors. The eigenvalues, communalities and amount of variance accounted for by each factor was examined to determine the next step in the EFA. The examination of the Scree plot and the amount of variance accounted for by the items suggested the possibility of four factors. The Bartlett’s test of sphericity (1954) and the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy (Kaiser, 1974) were examined again and were found to be adequate to continue with the analysis (p < .0001, and .907 respectively). The eigenvalues for each of the factors were 11.42 (factor 1), 2.98 (factor 2), 2.25 (factor 3), and 1.63 (factor 4). The amount of variance accounted for by each factor was 34.62%
(factor 1), 8.72% (factor 2), 6.51% (factor 3), and 4.94 (factor 4). After reviewing the pattern matrix it was decided to remove items 16, 31, and 37 due to cross-loadings meeting criteria for removal (the difference between factor loadings was <.15). Following the removal of these items the EFA specifying a four-factor solution was performed again.

With 30 items remaining an EFA was performed using an oblique rotation and specifying 4 factors. The eigenvalues, communalities and amount of variance accounted for by each factor was examined to determine the next step in the EFA. The eigenvalues for each of the factors were 10.01 (factor 1), 2.85 (factor 2), 2.13 (factor 3), and 1.59 (factor 4). The amount of variance accounted for by each factor was 33.52% (factor 1), 9.49% (factor 2), 7.09% (factor 3), and 5.23% (factor 4). After reviewing the pattern matrix and the items that loaded on each factor it was decided to remove item number 6 due to cross-loading meeting criteria for removal (the difference between factor loadings was <.15).

An EFA was performed with the remaining 29 items using an oblique rotation and specifying four factors. The eigenvalues, communalities and amount of variance accounted for by each factor was examined to determine the next step in the EFA. The eigenvalues for each of the factors were 9.75 (factor 1), 2.82 (factor 2), 2.12 (factor 3), and 1.54 (factor 4). The amount of variance accounted for by each factor was 33.62% (factor 1), 9.74% (factor 2), 7.31% (factor 3), and 5.31 (factor 4). After reviewing the pattern matrix, communalities, and items, it was decided to remove items 4, 5, 22, and 30. These items had lower-marginal factor loadings and did not fit conceptually with the other items on the factor. Specifically, on factor 1, item 30 (I feel uncomfortable using
diet/nutrition in order to have my body appear to others as conforming to my gender) assessed conformity through body modification while all other items measured discomfort with conformity in public. Item 22 on factor 2 was cross-loaded on both factor 2 and 4, but did not appear to fit conceptually well with either factor. Item 22 measured altering one’s appearance, while other items measured levels of comfort or discomfort. The language was inconsistent with other items. Item 5 on factor 3 measured discomfort with nonconforming whereas all other items on factor 3 measure comfort with conforming through diet/exercise. Item 4 on factor 4 contained language that was inconstant with the other items for measuring comfort with conforming through appearance.

After the removal of four items an EFA with the remaining 25 items was performed specifying a 4-factor structure and oblique rotation. The eigenvalues for each of the factors were 8.88 (factor 1), 2.51 (factor 2), 1.80 (factor 3), and 1.40 (factor 4). The amount of variance accounted for by each factor was 35.11% (factor 1), 10.06% (factor 2), 7.23% (factor 3), and 5.61% (factor 4). Eigenvalues were set at one to assess the initial factor structure of the remaining items. After reviewing the pattern matrix and factor correlation matrix it was determined that a 4-factor solution would be the most effective solution. Specifically, the high amount of variance accounted for by the four factors, the clear and strong loading of times onto a 4-factor solution, and the conceptual cohesiveness of items suggested the 4 factor solution

Item 6 on Factor 1 was left off subsequent data collection for the CAGES in Studies 2 and 3 due to researcher error. Therefore, an additional EFA was conducted in order to affirm stability of factor structure without the missing item. An EFA with 24
items (minus item 6 from the final factor solution) was performed specifying a 4-factor structure and oblique rotation. The eigenvalues for each of the factors were 8.43 (factor 1), 2.38 (factor 2), 1.80 (factor 3), and 1.38 (factor 4). The amount of variance accounted for by each factor was 35.12% (factor 1), 9.92% (factor 2), 7.51% (factor 3), and 5.74% (factor 4). The final factor structure accounted for 58.29% of the total variance and was chosen over the other solutions because it resulted in the most conceptually sound and statistically significant factor structure including stronger factor loadings and less cross loadings than the other solutions. The 4-factor solution had factor loadings ranging from .37 to .85 for all 24 items that remained in the CAGES. For the final step of Study 1, the scale’s internal consistency was assessed using Cronbach’s alpha. The reliability coefficient for the final total CAGES was .85, and the individual subscales were as follows; Factor 1, \( \alpha = .88 \), Factor 2, \( \alpha = .84 \), Factor 3, \( \alpha = .80 \), and Factor 4, \( \alpha = .86 \).

**Naming the Factors**

In the final CAGES Factor 1 was named Discomfort with Nonconformity. This scale consists of 11 items, four that are reverse-scored. The items deal with levels of discomfort with nonconformity across several domains of gender expression. Higher scores indicate higher levels of discomfort with nonconforming expressions of gender in the social context. Examples of items include; “I feel uncomfortable if I do not conform to gender expectation in social situations,” “I feel uncomfortable when my appearance does not meet others’ expectations of my gender,” and “I would be upset if my personal environment (e.g. home/office) did not conform to gender expectations.” Examples of reverse scored items include: “I am comfortable interacting with others in ways that defy gender norms,” “I feel comfortable talking with others about ways that gender stereotypes
can be overcome,” and “I feel comfortable purchasing items that are inconsistent with social expectations of my gender.” Discomfort with Nonconformity accounted for 35.12% of the variance (eigenvalue = 8.43) and demonstrated internal consistency of $\alpha = .88$.

The second factor was named Resentful Conformity This subscale consists of three items. Higher scores indicate higher levels of discomfort with conformity, as well as higher levels of distress at societal pressure to conform. Items included in this scale are: “I feel uneasy when gender norms inhibit the way I can express myself in public,” “It upsets me that gender norms influence my behaviors in public,” and “I feel upset when I conform to gender stereotypes when meeting new people.” The Resentful Conformity subscale accounted for 9.92% of the variance (eigenvalue = 2.38) and evidenced an internal consistency of $\alpha = .84$.

Factor 3 consists of three items and was named Active Physical Comfort-Conformity. Higher scores on Active-Physical Comfort Conformity indicate higher levels of use of diet/nutrition or exercise to bodily conform to gender expectations in physical gender expression. Items on Active-Physical Comfort Conformity include: “I feel comfortable using diet/nutrition in order to have my body appear to others as conforming to my gender,” “I enjoy using exercise/weight training that makes my appearance more consistent with gender expectations,” and “I feel most at ease using diet/nutrition to make my body appear more consistent with gender expectations.” The Active-Physical Comfort Conformity subscale accounted for 7.51% of the variance (eigenvalue = 1.80), and had an internal consistency of .80
Factor 4 was named Comfort with Conformity Appearance. This factor consists of six items, with three of these items reverse scored. Higher scores on Comfort with Conformity-Appearance indicate higher levels of comfort with conforming to gender norms in gender expression. Items include: “I am happy when I express myself in ways that are consistent with gender expectations,” “I feel most comfortable making clothing choices that fit expectations of my gender,” and “I feel most comfortable getting my hair cut in a way that most people perceive to clearly match my gender.” Reverse scored items include: “I feel resentful buying clothing that meets others’ expectations of my gender,” and “I feel most attractive when I challenge expectations of my gender.” Comfort with Conformity-Appearance accounted for 5.74% of the variance (eigenvalue = 1.38), with an internal consistency of .86.

**Between group Analyses**

In order to determine if significant differences in participants’ comfort levels with conformity of gender expression differed across sex or gender identity, sexual orientation, age, and race and ethnicity, a between-subjects multiple analyses of variance (MANOVA) or general linear model was performed (see Table 4). Given the sample size and characteristics, demographic categories were collapsed to allow for analysis. The race/ethnicity category was collapsed into people of color (n = 56) and whites (n = 298). Two cases were dropped from the analysis because of refusal to answer or uncategorizeable answers (e.g. “there is no scientific basis to race classification”). Sex/Gender was collapsed into three categories; female identified persons (n = 211), male identified persons (n = 102), and Transgender/Genderqueer/Transsexual/Other (n = 42). One case was dropped from sex/gender analysis due to refusal to answer. Sexual
Orientation was collapsed into two categories, lesbian/gay/bisexual/queer/ other (n = 132), and heterosexual/straight (n = 220). Four cases were dropped due to ambiguous answers or refusal to answer (e.g. “sex is not important, sanity is”). Given the unequal sample sizes in these collapsed categories, Levene’s test of unequal variances was used to assess reliability of comparisons. Additionally the Wilks–Lambda, Partial Eta squared, and power was assessed, with a partial eta squared value of .14 indicating a large effect size, .06 indicating a medium effect size, and .01 indicating a small effect size. In the case of sex/gender with three groups, a post-hoc comparison was used to identify variance between multiple groups. Given the unequal variances, the Games-Howell post-hoc comparison test was used for analysis.

An analysis of variance showed that scores on the CAGES were not significantly different across the demographic variables of race and ethnicity or age. Significant differences were found on the CAGES for sex/gender and sexual orientation. In order to examine the relationships among gender, sexual orientation identities and CAGES subscale scores, a 2 (sexual orientation identity) by 3 (sex/gender) multivariate analysis of variance (MANOVA) was conducted with the CAGES subscales (Discomfort with Nonconformity, Resentful Conformity, Active Physical Comfort-Conformity, Comfort Conformity-Appearance) as the dependent variables.

There were significant main effects for sex/gender on all four factors, F(2, 349) = 12.08, p = .000, partial $\eta^2 = .07$ (Discomfort with Nonconformity), F(2, 349) = 10.94, p = .000, partial $\eta^2 = .106$ (Resentful Conformity), F(2, 349) = 6.35, p = .002, partial $\eta^2 = .04$ (Active Physical Comfort-Conformity), and F(2, 349) = 13.87, p = .000, partial $\eta^2 = .08$ (Comfort with Conformity-Appearance). Levene’s test for homogeneity showed that the
assumption of equality of variance among the three groups was not violated. The analysis results can therefore be considered valid.

Post hoc analyses using the Games-Howell post hoc criterion for significance indicated significant differences between females and males (MD = .268, p = .000), and males and transgender/genderqueer persons (MD = -.350, p = .000) on Discomfort with Nonconformity, as well as significant differences between females and males (MD = -.359, p = .000) males and transgender/genderqueer persons (MD = .680, p = .000), and females and transgender/genderqueer persons (MD = -.321, p = .014) on Resentful Conformity. Additionally significant differences were found on Active-Physical Comfort Conformity between females and males (MD = -.226, p = .004), and on Comfort with Conformity-Appearance between females and males (MD = .171, p = .012), and males and transgender/genderqueer persons (MD = -.740, p = .000), and females and transgender/genderqueer persons (MD = .569, p = .000).

There were significant main effects for sexual orientation identity on all four factors F(1, 347) = 15.49, p = .000, partial η² = .04 (Discomfort with Nonconformity), F(1, 347) = 17.22, p = .000, partial η² = .05 (Resentful Conformity), F(1, 347) = 4.94, p = .027, partial η² = .01 (Active-Physical Comfort Conformity), and F(1, 347) = 34.63, p = .000, partial η² = .09 (Comfort with Conformity-Appearance).

No significant interaction effects were found between sex/gender and sexual orientation identity for any of the four CAGEs subscales in Study 1.

Descriptive Statistics

Internal consistency estimates were obtained from the final four-factor CAGES scale and the scale demonstrated high reliability (α = .85). Individual factors were found
to have high levels of internal consistency and reliability, Discomfort with Nonconformity ($\alpha = .90$), Resentful Conformity ($\alpha = .84$), Active Physical Comfort-Conformity ($\alpha = .80$), and Comfort with Conformity ($\alpha = .86$).

**Discriminant Validity**

Bivariate correlations were also calculated between the CAGES and a scale hypothesized to have little or no correlation with gender expression comfort and conformity. Specifically, scores on the BIDR-IM were correlated with scores on the CAGES. No significant correlations were found, thus supporting the CAGES as statistically different from the BIDR-IM.

**Study 2: Reliability and Validity Tests of the CAGES**

This study presents the results from the validity studies including correlation coefficients between the CAGES and the CMNI and CFNI for women and men on all three measures, as well as for men on the CMNI and for women on the CFNI.

**Convergent Validity**

To investigate the convergent validity of the CAGES, bivariate correlations were calculated between the CAGES and two gender scales used to assess conformity to societal gender norm conformity, the Conformity to Masculine Norms Inventory (Mahalik, 2003) and the Conformity to Feminine Norms Inventory (Mahalik, 2005). Although the measures were constructed to be used with a population of women (CFNI) and men (CMNI), participants were assigned to the CFNI or the CMNI, not based on gender/sex identification, but randomly. This method was chosen in order to explore and maintain the structural integrity of the CAGES, as the CAGES was developed to be given to anyone no matter what their gender/sex identification. It was hypothesized that the
salient construct in the subsequent analyses is conformity, not the specific gender norm. Between groups analyses were conducted to measure if men differed from women and women from men on both the CMNI and CFNI. In order to fully explore the convergent validity of the CAGES with the CFNI and CMNI, correlation coefficients for the CAGES with women participants on the CFNI, men on the CMNI, men on the CFNI and women on the CMNI are reported. Transgender and genderqueer participants were not included in analysis due to low sample size (n = 7).

*Preliminary Analyses*

The researcher hypothesized the CMNI and CFNI could be applied to both men and women in convergent validity analyses given the underlying construct in the CFNI, CMNI, and CAGEs is gender-role conformity. Mahalik and colleagues, in developing these scales, used focus groups of men and women to generate and articulate these specific gender-role norms. The CFNI and CMNI were created to measure the participants’ levels of conformity with these generated gender role norms. In the development of the CFNI, using a traditionally aged college student population who was primarily white, middle-class, and heterosexual, Mahalik and colleagues found that women scored significantly higher on the CFNI total scale score and on six of the eight subscales (HNR, SEXFID, DOMS, IIA, IWC, and THIN) (2005). Men were found to score higher on nine of eleven subscales (WIN, EMOC, RISK, VIOL, POW, DOM, PLAY, SELF, and HOMO) on the CMNI in Mahalik et al’s original study (2003). Mahalik, in his original studies concluded that the CFNI and CMNI measured specific gender-role norms for men on the CMNI and women on the CFNI. Since the researcher in the current study hypothesized that both men and women could complete either of the
scales and the underlying construct of conformity would still be measureable, between-
groups analyses were conducted to test for significant differences between men and 
women on the CMNI and CFNI subscales.

Between-groups analyses were conducted to assess if men and women differed in 
their responses on the CFNI and CMNI. A one-way MANOVA was conducted and found 
differences between men and women on four of eight CFNI subscales, and four of eleven 
CMNI subscales. For the CFNI, men differed from women significantly on; Have Nice 
Relationships (HNR) F(1, 87) = 4.62, p = .034, Involvement With Children (IWC) F(1, 
87) = 3.49, p = .065, Involvement in Romantic Relationships (IRR), F(1, 87) = 11.15, p = 
.001, and Investment in Appearance (IIA), F(1, 87) = 28.48, p = .000. Women differed 
from men on the CMNI on four of eleven subscales; Risk Taking (RISK), F(1, 87) = 
4.10, p = .046, Violence (VIOL), F(1, 87) = 8.47, p = .005, Power Over Women (POW), 
F(1,87) = 20.65, p = .000, and Disdain for Homosexuality (HOMO), F(1, 87) = 11.09, p 
= .001.

These findings indicate that men and women respond significantly differently to 
the CMNI and CFNI. The wording of items and applicability of certain subscales 
specifically to men or women on the CMNI and CFNI respectively, led the researcher to 
conduct separate analyses for men on the CMNI, women on the CFNI, men on the CFNI, 
and women on the CMNI. Analyses including both men and women on the CMNI and 
CFNI were not utilized given the findings of the preliminary analyses indicating 
significant differences in men and women’s responses on the subscales of both scales.

Convergent Validity
For women taking the CFNI (n = 55), significant correlations were found between all four CAGES subscales and four of eight CFNI subscales. Discomfort with Nonconformity was correlated with Investment in Children (IWC) \((r = .28, p = .04)\), Thinness (THIN) \((r = .28, p = .006)\), and Sexual Fidelity (SEXFID) \((r = .51, p = .000)\). The CAGES Resentful Conformity subscale was negatively correlated with the CFNI subscale IWC, \((r = .37, p = .006)\), and positively correlated with the Modesty (MOD), \((r = .34, p = .011)\). Active Physical Comfort-Conformity was strongly positively correlated with THIN, \((r = .52, p = .000)\). Comfort with Conformity-Appearance was negatively correlated with MOD, \((r = -.28, p = .041)\), and positively correlated with Investment in Appearance (IIA) \((r = .37, p = .005)\). With women on the total CAGES and total CFNI a significant positive correlation of \((r = .49, p = .000)\) was found.

For men taking the CMNI (n = 33), significant correlations were found between all four CAGES subscales and five of eleven CMNI subscales. Discomfort with Nonconformity was positively correlated with Emotional Control (EMOC), \((r = .49, p = .004)\), Power Over Women (POW), \((r = .56, p = .001)\), Disdain for Homosexuality (HOMO), \((r = .78, p = .000)\), and negatively correlated with Pursuit of Status (POS), \((r = -.44, p = .011)\). Resentful Conformity was negatively correlated with two subscales, POW, \((r = -.43, p = .013)\), and HOMO \((r = -.41, p = .017)\). Active Physical Comfort-Conformity was positively correlated with EMOC, \((r = .43, p = .012)\), Violence (VIOL), \((r = .45, p = .008)\), POW, \((r = .41, p = .019)\), and HOMO, \((r = .51, p = .002)\). Comfort with Conformity-Appearance was correlated with EMOC, \((r = .36, p = .039)\), and HOMO, \((r = .53, p = .002)\). With men on the total CAGES and total CMNI a significant positive correlation of \((r = .43, p = .012)\) was found.
For women taking the CMNI (n = 47), significant correlations were found between two CAGES subscales and four of eleven CMNI subscales. Discomfort with Nonconformity was positively correlated with Winning (WIN), \( r = 0.32, p = 0.029 \), POW, \( r = 0.38, p = 0.008 \), and HOMO, \( r = 0.53, p = 0.000 \). Resentful Conformity was negatively correlated with two subscales, POW, \( r = -0.43, p = 0.013 \), and HOMO \( r = -0.41, p = 0.017 \). Comfort with Conformity-Appearance was correlated with Dominance (DOM), \( r = 0.35, p = 0.015 \), and HOMO, \( r = 0.60, p = 0.000 \). With women on the total CAGES and total CMNI a significant positive correlation of \( r = 0.34, p = 0.021 \) was found.

For men taking the CFNI (n = 32), significant correlations were found between three CAGES subscales and four of eight CFNI subscales. Discomfort with Nonconformity was correlated with SEXFID \( r = 0.35, p = 0.047 \), Investment in Romantic Relationships (IRR) \( r = 0.40, p = 0.022 \), and negatively correlated with Domestic (DOMS) \( r = -0.35, p = 0.050 \). The CAGES Resentful Conformity subscale was positively correlated with DOMS, \( r = 0.39, p = 0.029 \), and with Investment in Appearance (IIA) \( r = 0.45, p = 0.009 \). Comfort with Conformity-Appearance (CCA) was positively correlated with IRR, \( r = -0.39, p = 0.028 \), and negatively correlated with IIA \( r = -0.37, p = 0.037 \). See Table 7 for all correlations calculated. With men on the total CAGES and total CFNI no significant correlation was found.

**Between-Groups Analyses**

In order to determine if significant differences in participants’ comfort levels with conformity of gender expression differed across sex or gender identity, sexual orientation, age, and race and ethnicity, a between-subjects multivariate analysis of variance (MANOVA) was performed (see Table 4). Given the sample size and
characteristics, demographic categories were collapsed to allow for analysis. The race/ethnicity category was collapsed into people of color (n = 34) and whites (n = 138). Sex/Gender was collapsed into two categories; female identified persons (n = 102), and male identified persons (n = 64). The transgender/genderqueer/transsexual/Other (n = 7) category was not included in the analyses because the sample size was too small for adequate statistical comparison. Sexual Orientation was collapsed into two categories, lesbian/gay/bisexual/queer/other (n = 58), and heterosexual/straight (n = 114). The Wilks–Lambda, Patial Eta squared and power were assessed.

A one-way analysis of variance showed that scores on the CAGES were not significantly different across the demographic variables of race and ethnicity or age. Significant differences were found on the CAGES for sex/gender and sexual orientation.

In order to examine the relationships among gender, sexual orientation identities and CAGES subscale scores, a 2 (sexual orientation identity) by 2 (sex/gender) multivariate analysis of variance (MANOVA) was conducted with the CAGES subscales (Discomfort with Nonconformity, Resentful Conformity, Active Physical Comfort-Conformity, Comfort Conformity-Appearance) as the dependent variables. Significant main effects were found for sex/gender on all four factors, F(1, 165) = 14.69, p = .000, partial \( \eta^2 = .08 \) (Discomfort with Nonconformity), F(1, 165) = 17.47, p = .000, partial \( \eta^2 = .10 \) (Resentful Conformity), F(1, 165) = 15.55, p = .000, partial \( \eta^2 = .09 \) (Active-Physical Comfort Conformity), and F(1, 166) = 18.18, p = .000, partial \( \eta^2 = .10 \) (Comfort with Conformity-Appearance).

Significant main effects were found for sexual orientation identity on three of four factors F(1, 165) = 6.29, p = .013, partial \( \eta^2 = .04 \) (Discomfort with Nonconformity), F(1,
A significant interaction effect between sex/gender and sexual orientation identity was found on one CAGES subscale, Comfort with Conformity-Appearance, F(1, 165) = 6.86, p = .010, partial $\eta^2 = .04$. As no significant interactions were found on any other factors, this finding is noninterpretable.

**Descriptive Statistics**

The means for the CAGES subscales indicated that the participants reported moderate levels of comfort with gender expression conformity; Discomfort with Nonconformity ($M = 2.07, SD = .50$), Resentful Conformity ($M = 2.69, SD = .66$), Active-Physical Comfort Conformity ($M = 2.48, SD = .72$), and Comfort with Conformity-Appearance ($M = 2.78, SD = .60$).

**Study 3: Test-Retest Reliability**

The purpose of this study was to explore the test-retest reliability of the final CAGES. Participants were recruited from undergraduate classrooms at a large Midwestern university and given a survey packet including the final CAGES and demographic questions twice over a two-week period. The analyses described below include the correlations found between the first and second administration of the CAGES.

**Descriptive statistics.** The means and standard deviations for the four factors of the CAGES at the first test administration were as follows; Discomfort with Nonconformity ($M = 2.6, SD = .40$), Resentful Conformity ($M = 2.70, SD = .50$), Active-Physical Comfort Conformity ($M = 2.16, SD = .54$), and Comfort with Conformity-Appearance ($M = 1.86, SD = .39$). The means and standard deviations for the four factors
at the second test administration were as follows: Discomfort with Nonconformity ($M = 2.58, SD = .41$), Resentful Conformity ($M = 2.79, SD = .51$), Active-Physical Comfort Conformity ($M = 2.15, SD = .59$), and Comfort with Conformity-Appearance ($M = 1.91, SD = .33$).

Reliability statistics. The two-week test-retest reliability estimates for the CAGES subscales were as follows: Discomfort with Nonconformity ($r = .86$), Resentful Conformity ($r = .82$), Active-Physical Comfort Conformity ($r = .49$), and Comfort with Conformity-Appearance ($r = .79$). These results suggest that three of the four CAGES subscales are acceptably stable over a two-week period, with the exception of Active-Physical Comfort Conformity which was below the established cut-off ($r = .70$). This finding is explored in the discussion section below. Given the strong correlations found for all three other subscales, the CAGES appears to evidence strong test-retest reliability.
CHAPTER 5: DISCUSSION

The purpose of the proposed study was to construct a new gender scale that measures levels of comfort with conformity and nonconformity in gender expression. The proposed scale is grounded in psychological and gender theory, and is intended to be non-binary based in its framework. It was developed to be taken by men, women, and transgender individuals using an inclusive language. The CAGES was developed to assess the interaction of comfort and conformity across cognitive, affective, and behavioral aspects of gender expression. This chapter will discuss the implications of the results presented in Chapter 4. First, the relationship of the results to previous research and theory will be presented. The findings of the three studies will be discussed including the exploratory factor analysis, convergent and discriminant validity, and test-retest reliability, within the framework of the proposed research hypotheses grounded in gender and psychological theories. Lastly, limitations of the current study, directions for future research, and conclusions will be discussed.

Relationship of the Results to Previous Research and Theory

The development and validation of the CAGES is intended to further gender measurement in the field of psychology. In Chapter 2, the existing research in gender measurement is reviewed. The literature review identified three main areas in gender scale development: early explorations and definitions of gender, feminist and social theoretical challenges to gender, and new directions including transgender and men and masculinity studies.
Early gender measurement, including Terman and Miles (1938) work, strived to quantify gender as an inherent trait that could be measured in men and women. Gender measurement was primarily aimed at identifying those who were “deviants” or “inverts”, e.g. homosexuals. With the advent of the women’s movement and the growth of psychology, challenges to this definition of gender and its subsequent use in measurement were made. Bem (1974), and Spence and Helmreich (1974), made tremendous advances in the scientific quantification of gender as a measurable phenomena with the BSRI and the PAQ. These two scales presented gender as a construct that included a range of gendered traits. The concept of severing gender norms from personality attributes was introduced through these two landmark scales, and other studies of the time.

The rise of men and masculinity studies furthered scale development in the area of gender, and introduced the concept of gender role conflict (O’Neill, 1986). Gender as a construct, and separate from gender norms or stereotypes, began to be explored more fully in the last 10 years, with several studies questioning the definition of femininity and masculinity themselves (Hoffman, 2000, Mahalik, 2003, 2005).

Throughout the process of developing the current scale, the existing literature within the domains of psychology and gender studies were utilized as foundations to understanding the construct to be measured. Inherent in these different but overlapping disciplines is a tension between different ontological understandings of gender. Psychology, based in logical positivist, empirical scientific processes of furthering knowledge, and the post-modern challenge represented by qualitative and inductive methods of gender studies, presents the challenge of integrating these divergent modes of
knowledge production. Throughout the process of developing the items for the CAGES, this tension was intertwined in the project of making measurable the construct of gender expression.

Given this history of gender measurement and the fluid and amorphous nature of gender as a quantifiable construct, it was necessary for the current study to be clearly grounded in gender measurement theory. The literature on gender-role conflict, self-concept theory, transgender and feminist challenges to sex and gender, and gender norm conformity were integral to the development of the current scale. The results of the three studies will be discussed within this framework of theoretical underpinnings in the following sections.

Gender-Role Conflict and Gender Expression

The literature defines gender-role conflict as “a psychological state in which socialized gender roles have negative consequence on the person and others” (O’Neill, 1986). The idea of gender roles as a phenomena that occurs outside of the individual and becomes internalized and manifested through three domains: cognitions, affect, and behaviors, is foundational for the development of the items on the CAGES. Gender expression was hypothesized to be a measurable phenomena manifested through behaviors, cognition, and affect. A primary goal of the development and validation of the CAGES was to develop a scale that could be used for men, women and transgender persons, and would measure levels of comfort and conformity across these three domains of cognition, affect, and behaviors. A main hypothesis of the CAGES development was the manifestation of these domains in the items to be supported as empirically sound constructs. This goal and associated hypotheses were predominantly supported by the
results. Of the 41 initial items developed for the scale and included in the initial exploratory factor analysis, 24 items were found to contribute to the final four-factor solution. Of the four factors, Discomfort with Nonconformity, Resentful Conformity, Active Physical Comfort-Conformity, and Comfort with Conformity-Appearance, all the items are based in an affective or behaviorally based manifestation of gender expression, (e.g., “It upsets me that gender norms influence my behaviors in public.”) This item from the Resentful Conformity subscale clearly illustrates the basis in behavioral component of gender expression, as well as the range of levels of comfort with conformity the CAGES was proposed to measure.

The original factors were hypothesized to include each of the components of gender expression – behaviors, cognitions, and affect. The final CAGES items each contain an affective (“e.g. “I am happy” or “I am upset”) and a behavioral component (e.g. “Talk with others” or “purchasing things”), but do not appear to contain the cognitive aspect of gender expression. Items intended to capture the cognitive aspect of gender expression were not kept in the final CAGES because of indecipherability and lack of fit with the other CAGES items. Many of the items including a cognitive aspect (e.g. “I like to think about ways gender stereotypes can be challenged”) were revised upon feedback from experts to items that included an affective and behavioral component (e.g. “I feel comfortable talking with others about ways that gender-role stereotypes can be challenged.”).

Continua of Comfort and Conformity

A main component of gender-role conflict and gender expression is the embedded idea that gender roles are an external phenomena and that individual’s experience internal
conflict with meeting these gendered expectations. Many items on the CAGES elucidate this theoretical underpinning, as evidenced by the four-factor structure of items exploring discomfort or comfort with conformity or nonconformity on affective and behavioral levels, and measuring incongruence behaviorally and affectively, (e.g. the CAGES item “I feel upset when I conform to gender stereotypes when meeting new people.”) In the early stages of the development of the CAGES, after exploring the existing literature measuring gendered phenomena, it became clear that an area deserving additional development in gender measurement was gender expression. The researcher initially hypothesized gender expression would exist on a continuum, from those who are highly gender conformant to those who are highly gender nonconformant, with many shades of conformity in between.

As items were being developed it became clear that conformity in gender expression also must include the affective component of comfort or discomfort, as the meaning of gender expression that is conformant or nonconformant changes psychologically with the added layer of comfort or discomfort with this nonconformity. Several researchers have documented the management of affect as a central component of negotiating gender expression in the social context (O’Neill, 1986, Bem, 1993, Mahalik, 2003, 2005). O’Neill (1986), in the development of the gender-role conflict scale, speaks to the affective component through the naming of the internal experience of conflict in gender-role endorsement or rejection. Additionally, Bem (1993), and Mahalik (2003, 2005), also address the role of affect in gender-role conformity. The researcher hypothesized that there would be two continua: one assessing comfort with conformity and another assessing comfort with nonconformity in gender expression (see figure 1).
These two overlapping continua would assess individuals’ different levels of comfort and discomfort with conformity and nonconformity in gender expression through thoughts, behaviors and feelings. The items were developed with this hypothesis in mind (see Appendix A for full list of initial items).

Before the exploratory factor analysis (EFA), the items were hypothesized to fit a 4-factor structure with the following hypothesized factors: Comfort-Conformity-Public, Comfort-Nonconformity-Public, Discomfort-Nonconformity-Public, and Discomfort-Conforming-Public. After completion of the EFA, the 24 remaining items were found to fit a four-factor structure of Discomfort with Nonconformity, Resentful Conformity, Active Physical Comfort-Conformity and Comfort with Conformity-Appearance. The hypothesized factor structure was not supported in the exact format, but the salient constructs of comfort-conformity and discomfort-nonconformity in gender expression were supported. The EFA found support for items that appeared to be grouped differently than originally hypothesized. The Active-Physical Comfort Conformity factor contains items that measure comfort and conformity but with the physical maintenance of gender expression being the salient construct. The Resentful Conformity factor contains items that grouped around the affective component of feeling resentment or upset at pressure to conform, while the other two scales, Discomfort with Nonconformity and Comfort with Conformity-Appearance, contain items focused on the hypothesized groupings of discomfort and comfort with conformity or nonconformity. Overall, the hypothesized constructs were partially supported, but the hypothesized continua of gender expression were not supported.
The means and standard deviations for all four factors seem to indicate a more skewed, on some factors, to normal range of responses on other factors (see figures 2-5). Differences were found between groups, but instead of finding a clear strong continuum of gender expressions, the CAGES appears to measure levels of comfort or discomfort with levels of conformity or nonconformity in gender expression. The results from the EFA seem to point towards a quadrant conceptual understanding of the CAGES. In this quadrant representation, respondents to the CAGES can be placed in one of four quadrants based on their scores on different subscales. The comfort with conformity quadrant includes respondents who score highly on the Comfort with Conformity-Appearance and Active-Physical Comfort Conformity subscales. The discomfort with conformity quadrant includes respondents who feel uncomfortable conforming to gender norms and score higher on the Resentful Conformity subscale. The comfort with nonconformity quadrant includes respondents who score higher on Resentful Conformity and lower on Discomfort with Nonconformity, and the last quadrant, discomfort with nonconformity includes respondents who score high on Discomfort with Nonconformity. Further research with the CAGES would assist in clarifying this conceptual issue.

Between Group Differences

The participant pools for the exploratory factor analysis and the convergent validity studies were targeted at representing a normal population as well as oversampling for lesbian, gay, bisexual, transgender, transsexual, and genderqueer participants. The range of responses indicate significant differences by gender/sex identity and for sexual orientation, meaning that LGBTQ participants tended to score significantly higher on the Resentful Conformity subscale, while heterosexually
identified participants scored higher on the Discomfort with Nonconformity and Comfort with Conformity-Appearance subscales. Significant differences were found on the Active-Physical Comfort Conformity subscale, but these results were smaller and possibly mediated by gender/sex identity. Men tended to score significantly higher than women and transgender individuals on Discomfort with Nonconformity and Comfort with Conformity-Appearance subscales, whereas women and transgender individuals scored higher on Resentful Conformity and Active-Physical Comfort Conformity. Women scored lower on Resentful Conformity than transgender individuals, with men scoring the lowest. These differences indicate the importance of contextual and identity factors in the expression of gender within the social context, and point to possible research and practice implications of the CAGES. A benefit of further research would be to continue the exploration of the interaction of sex/gender identity and sexual orientation identity in gender expression.

**Gender Norm Conformity**

A main premise in the gender literature, specifically men and masculinity studies, is the existence of conflict and strain with societal gender norms. Several measures have been developed in an attempt to explore the internal psychological process of managing gender role conflict with conformity, most notably O’Neill’s gender role conflict scale (discussed above) and Mahalik’s work in the development and validation of the Conformity to Feminine Norms Inventory (CFNI) (2005) and the Conformity to Masculine Norms Inventory (CMNI) (2003). The underlying premise of these two scales is that there exist in society implicit and explicit gender role norms that men and women are expected to fulfill. Mahalik and colleagues, in developing these scales, used focus
groups of men and women to generate and articulate these gender-role norms. The CFNI and CMNI were created to measure the participants’ levels of conformity with these generated gender role norms. In the development of the CFNI, Mahalik and colleagues found that women scored significantly higher on the CFNI total scale score and on six of the eight subscales (HNR, SEXFID, DOMS, IIA, IWC, and THIN) (2005). Men were found to score higher on nine of eleven subscales (WIN, EMOC, RISK, VIOL, POW, DOM, PLAY, SELF, and HOMO) on the CMNI in Mahalik et al.’s original study (2003). Mahalik concluded that these findings indicated that the CMNI (for men) and the CFNI (for women) measured distinct gender-role norms for men and women, and men and women would consistently respond differently on these two scales.

The CAGES was developed to measure comfort and discomfort with conformity to gender role norms, but does not name the specific gender role norm, instead focusing on the process of and feeling of comfort or discomfort in conformity or nonconformity in gender expression. The researcher specifically developed the CAGES as an instrument that used non sex-specific language in order to avoid reinforcing the sex/gender binary embedded in previous gender measurement instruments. It was hypothesized that comfort and conformity to gender-role norms would be the significant constructs, and that utilizing sex-specific instruments only serves to conflate measurement of endorsement of the specific gender norms with the psychological processes of conformity to socially constructed gender-role norms. In choosing the CFNI and CMNI, the researcher decided to have both men and women complete either the CMNI or CFNI based on these underlying assumptions; the salient construct would be conformity and to avoid reifying the gender binary in separating responses by sex.
The CAGES was hypothesized to moderately correlate with the CFNI and CMNI to evidence convergent validity. This hypothesis was supported by the results, as discussed below. Given the hypotheses guiding the choice to administer the CFNI and CMNI to both men and women, and given Mahalik’s original findings of significant differences between men and women on these scales, between-groups analyses were conducted to explore differences in responses by men and women. Significant differences between men and women were found on four of eight subscales on the CFNI (HNR, IWC, IRR, and IIA). These findings are primarily consistent with Mahalik et al’s findings of significant differences on six of eight subscales (HNR, SEXFID, DOMS, IIA, IWC, and THIN) in their 2005 study. The current study did not find differences between men and women on the SEXFID or DOMS subscale.

Significant differences were found between men and women on four of eleven subscales on the CMNI including RISK, VIOL, POW, and HOMO. Men were found to score higher on nine of eleven subscales (WIN, EMOC, RISK, VIOL, POW, DOM, PLAY, SELF, and HOMO) on the CMNI in Mahalik et al’s original study (2003).

There are several possible interpretations for the similarities and differences in the between-groups differences for men and women on the CMNI and CFNI in the current study. One possible interpretation is that the samples were different enough to produce different results. Mahalik’s original samples included traditionally aged college students who were primarily middle class, white, and heterosexual. The current study included a more diverse sample in age, education level, and most importantly in sexual orientation and sex/gender identity. The current study oversampled for LGBQ populations, and included a larger than normal LGBQ population (38%). The inclusion of a more diverse
sample may have caused more individuals who are less likely to rigidly endorse gender-role stereotypes. By including people from a wider range of ages, experiences, and sexual orientation identities, a wider range of gender-norm conformity or nonconformity was represented.

After examining the items on the CMNI and CFNI and the CAGES, it became clear that several items on the CFNI and CMNI were difficult for a substantial proportion of male (CFNI) and female (CMNI) participants to answer. Although significant correlations were found on the scales when all participants were included, because of the significant differences found between groups on a majority of the subscales, it was decided to analyze the CFNI with female participants, and the CMNI with male participants in order to accurately assess and interpret results. Transgender/genderqueer participants were not included because of small sample size (n = 7). The results of the intercorrelations between the CFNI, CMNI, and CAGES are discussed below. Again, results from female participants on the CMNI and male participants on the CFNI can only be interpreted tentatively given the non-applicability of many of the items on these scales to the separate populations.

Women Completing the CFNI

For women on the CFNI, several significant correlations were found to support the CAGES as measuring a similar yet different construct. Involvement with Children (IWC) was found to be positively correlated with Discomfort with Nonconformity, and negatively correlated with Resentful Conformity. Thus, women who are more uncomfortable with gender nonconformity reported higher levels of agreement with the importance of the involvement with children, whereas women who were more resentful
of gender conformity reported lower levels of agreement with this value. The Thinness (THIN) subscale was moderately correlated with Discomfort with Nonconformity, and highly correlated with Active-Physical Comfort Conformity, indicating that women who were more uncomfortable with nonconformity and more actively physically conforming were supportive of thinness ideals for women. These findings seem to fit with a body of research and writings on femininity ideals for women, and thus were expected, given that two main components of the femininity gender-role schema for women is a focus on the role of motherhood and on women’s physical appearance (Mahalik, 2005).

*Women Completing the CMNI*

For female participants completing the CMNI, significant correlations were found on three CAGES subscales and five of eleven CMNI subscales. Discomfort with Nonconformity was found to have significant positive correlations with WIN, POW, and HOMO, meaning women who were more uncomfortable with gender nonconformity endorsed attitudes consistent with winning and competition, patriarchal domination, and homonegativity. Women who are uncomfortable with gender nonconformity appear to endorse more traditional, conservative, patriarchal values. The POW subscale includes several items that are worded awkwardly when applied to women (e.g. “I treat women as equals.”), but also contains items that appear to endorse more traditionally patriarchal values (e.g. “Women should be subservient to men.”).

The CAGES Resentful Conformity subscale was found to negatively correlate with POW and HOMO, meaning that women who were more resentful of pressures to conform to gender norms scored lower in patriarchal and homonegative attitudes. Conversely, women who scored highly on the Comfort with Conformity-Appearance
subscale and who endorsed higher levels of comfort with conformity in their appearance, were more likely to score higher on the DOM and HOMO subscale. Thus, women who scored higher on the Comfort with Conformity-Appearance subscale, endorsed items indicating support for traditional notions of dominance and homonegative attitudes.

Although interpretations of these findings can only be tentative at best, the results seem to support an underlying hypothesis in much research on gender norms: that gender expression, conformity, and conflict are rooted in sociopolitical frameworks of patriarchy, male dominance, and homonegativity (hooks, 2000).

*Men Completing the CMNI*

Several significant correlations were found on the CMNI subscales. All four CAGES subscales correlated strongly with the Disdain for Homosexuals (HOMO) subscale on the CMNI. These items on the CMNI include statements indicating fear of and dislike for homosexuals. Given the documented connections between fear of gender nonconformity and homosexuality, these correlations are supported by theoretical, practical, and empirical literature. Interestingly, the Resentful Conformity subscale was negatively correlated with the HOMO subscale, meaning that men who were more highly resentful of pressures to conform to gender norms, had lower scores on HOMO, indicating lower levels of homophobia.

Additionally, the Power Over Women (POW) subscale on the CMNI was correlated with three of four CAGES subscales. The POW subscale includes items indicating men’s preference and comfort with patriarchal values and women’s inequality. POW was negatively correlated with Resentful Conformity, meaning that men who were more resentful of gender conformity, were less supportive of patriarchal values.
Emotional control (EMOC) was also correlated with three of the four CAGES subscales, indicating men who were more highly comfortable with gender conformity scored higher in controlling their emotions. These findings seem to support the expected outcomes for men conforming to gender role norms. A body of research, specifically within men and masculinity studies, has long hypothesized about the connection between patriarchal oppression and inequality with conformity to hegemonic masculinity norms. Additionally, the restrictiveness of emotions associated with higher levels of conforming to masculinity norms fits with much research on men and emotional expressiveness.

*Men Completing the CFNI*

With male participants who completed the CFNI, significant correlations were found between three of the CAGES subscales with four CFNI subscales. Discomfort with Nonconformity was found to have moderate significant correlations with IRR and SEXFID, and a negative correlation with DOMS, meaning that men who scored higher as being uncomfortable with not conforming to gender norms scored higher on subscales measuring investment in romantic relationships and sexual fidelity, and lower on domestic matters. The first two findings seem to be at odds with dominant thought on gender norms as one would expect men who are more highly uncomfortable with gender nonconformity would not score high on these subscales. One possible interpretation is that these subscales and their items may measure more conservative values, rather than gendered norms. Thus, men who are more uncomfortable with gender nonconformity may be more conservative and it follows would be more supportive of traditional, conservative values such as sexual fidelity and monogamy. The IIR subscale was also strongly correlated with the CAGES Comfort with Conformity-Appearance subscale,
which supports this possible interpretation: men who are more comfortable with conformity are more highly invested in romantic relationships.

On the CAGES Resentful Conformity subscale, positive correlations were found between DOMS and IIA, meaning men who are more resentful of pressure to conform to gender roles score higher on domestic matters and investment in appearance. This is an interesting finding as these two subscales are two of the strongest in representing feminine gender norms, as well as include many items that include typically female behaviors in dominant culture (e.g. wearing make-up). The IIA subscale was negatively correlated with the Comfort with Conformity-Appearance subscale on the CAGES, meaning that men who were more comfortable with conforming to gender norms in their appearance scored lower on the CFNI’s measurement of investment in appearance. This finding illuminates the difference in the wording of the CAGES and CFNI, as the CFNI measures endorsement of behaviors of feminine appearance management (e.g. “I regularly wear make-up.”), whereas the CAGES measures the general endorsement of gender conformant grooming behaviors (e.g. “I feel most comfortable getting my hair cut in a way that most people perceive to clearly match my gender.”).

Test-Retest Reliability

It was hypothesized that the CAGES would evidence adequate test-retest reliability over a two week period. The results mainly support the hypothesis. Three of the four CAGES subscales evidenced high significant correlations over a two week period. One subscale, the Active-Physical Comfort Conformity subscale evidenced a moderate correlation (r = .49). This subscale includes three items that measure levels of comfort with active physical conformity to gender norms (e.g. “I feel most at ease using
diet/nutrition to make my body appear more consistent with gender expectations.”). The sample may have been too small for an adequate analysis of this subscale, or the particular dimensions of the sample (primarily heterosexual) may have had an impact. Additionally, this subscale may be impacted by variables other than gender norm comfort and conformity, possibly levels of physical activity, body image and dieting beliefs, or exercise and sports participation. This subscale may not be measuring a stable construct of comfort with actively using physical means to conform to gender norms. It would be important for further research to explore the validity and reliability of this subscale, and possibly consider it for removal as a subscale from the CAGES. In the current study, it was decided to keep the Active-Physical Comfort Conformity subscale pending further research with the caution that results be interpreted tentatively.

**Implications for Future Research and Practice**

The findings of the current study have a number of implications for future research and practice in the area of gender and gender expression. For future research, confirmatory factor analyses need to be conducted on the CAGES. A new sample of participants, as diverse in gender/sex identity, but more diverse in racial-ethnic identity is needed to increase generalizability of the results.

An important area of future research would be confirming the validity of the factor structure of the CAGES through confirmatory factor analysis (CFA) and structural equation modeling (SEM). The analyses utilized in the current study were chosen based on the existing literature recommending best practices for scale development (Worthington & Whittaker, 2006, Devellis, 2003). Exploratory factor analysis (EFA) is often used when there are untested theoretical relationships between hypothesized
factors. Because the researcher hypothesized the items would be correlated, a principal axis factor (PAF) method with an oblique rotation was chosen. It has been shown in the literature that PAF as a factor analysis technique, rather than a principal components analysis (PCA), generalizes more easily to CFA and SEM (Worthington & Whittaker, 2006). Thus, the CAGES, with the factors already having been explored through EFA, would be a good candidate for SEM and specifically, CFA. The use of SEM to confirm tested hypothetical relationships between variables would be useful in further delineating the relationship of the constructs of comfort, conformity, and gender expression in the CAGES. With the complexity of gender expression as a quantative construct, SEM would be a highly useful tool in assessing latent variables (social and contextual factors) impact on the identified measured variables of comfort, conformity, and gender expression (as manifested through gendered behaviors).

Additionally, the use of CFA to test if the CAGES is applicable to different populations, given the significant differences found between men, women, and transgender individuals as well as between LGBQ and heterosexual individuals, would be a possibility for further research. CFA can be used to test the applicability of the already tested scale to different populations. Although the CAGES was developed in the wording and intent of the scale to be inclusive of men, women, transgender and differently sexed individuals, results from the between-groups analyses indicate that men, women, and transgender persons respond significantly differently on all four factors. Heterosexually identified participants responded significantly differently on all four subscales than LGBQ individuals, and additional analyses could be useful in exploring how gender identity and sexual orientation identity interact to impact comfort or discomfort with
conformity or nonconformity in gender expression. Additional analyses could explore the reliability and validity of the CAGES as a scale intended to be used with an inclusive population of men, women, transgender, and LGBQ individuals.

The development and validation of the CAGES introduces to gender measurement the construct of gender expression as a phenomenon experienced within the social context and impacted by social gender norms. Additional ideas for future research could be guided by this framework. A major contribution to the literature is the inclusivity of the items of the CAGES, which serves the purpose of separating the measurement of gender expression from the measurement of gender stereotypes. Deepening this line of research could further illuminate how men, women and transgender person experience their gender expression within the socially constructed meanings of gender norms.

Additionally, the CAGES could be used in conjunction with other established scales to explore the relationship between comfort and conformity of gender expression with psychological correlates such as self-esteem, self-concept, depression, anxiety, and well-being. Further exploring the connections between gender expression flexibility and social contextual factors that contribute to pressure for conformity could provide further empirical evidence of the complexity of gender in context in psychology.

Another important area of research would be to validate the CAGES for different populations. The current study sought a representative sample, with oversampling for LGBTQ persons, but did not approximate current demographics based on U.S. census data. It would be an important area of research inquiry to validate the CAGES with an LGBTQ sample and separately a heterosexual sample, as well as a male sample, female sample, and transgender sample. Significant differences were found in responses from
participants by sexual orientation and gender and as such, it would be important o explore if the CAGES measures gender expression adequately across these identity variables. Additionally, generalizability to persons of color is limited given the small sample of persons of color in the three studies. Increasing the sample size of persons of color in future research would improve the ability of the CAGES to be effectively utilized.

The current study also provides additional insight into the connections between homophobic attitudes, patriarchal attitudes, and negative attitudes towards gender nonconforming. Given the strength of the correlation coefficients between the HOMO and POW subscale and all four of the CAGES subscales, it invites additional research questions on the nature of the relationship between these attitudes and behaviors. Researchers could create new measures to further explore this relationship, or develop studies to further understanding of the meaning of these relationships.

The results of this study provide several implications for practice. First, for practitioners working with genderqueer, gender nonconformant, and transgender populations, the CAGES could be a useful tool in assessing levels of comfort and discomfort with conformity or nonconformity. The CAGES could also be useful in working with nontransgender populations who have concerns about gender identity, for example, men struggling with the pressure of gender role norms influencing their emotional selves or ability to relate to others. Lastly, the CAGES could be useful in practice as a way to explore the rigidity or flexibility of someone in their gender expression. Higher levels of conformity in gender expression could be related to higher levels of conformity in other areas of life and this could be an important area of exploration for a psychologist in practice.
Finally, the current scale construction project negotiated the tensions between the restrictions of logical positivist research and the ambiguity and complexity of construct being explored – gender expression. Future research utilizing a mixed methods approach wherein qualitative experiences of comfort and conformity of gender expression in the social context would be combined with quantitative methods to validate an empirically based scale would serve to deepen understandings of the proposed area of inquiry.

Limitations of the Current Research

Although the results of the current studies provide promising preliminary reliability and validity information on the CAGES, it is important to note the limitations of the study. The scale and its four subscales evidence strong factor loadings and good inter-item reliability, but two of the subscales are at the minimum number of items for a subscale. Resentful Conformity and Active-Physical Comfort Conformity are three items each, and this may play a factor in subsequent reliability and validity analyses, such as the Active-Physical Comfort Conformity subscale evidencing low test-retest reliability as compared to the other three scales. This may continue to be a concern in further research, specifically in a confirmatory factor analysis of the CAGES.

The current study primarily relied upon internet data collection for the first two studies, and classroom data collection for the third study. Although the participant demographics for the first and second study includes a more diverse sample than typically found in most social science research, there was an overrepresentation of women, LGBQ and transgender persons in the sample. There was also an underrepresentation of racial/ethnic minorities. In the third study, the classroom data collection, aside from the researcher error of not including demographics for
approximately half the sample, the demographics that were found were limited to a traditionally college aged, primarily heterosexual and white sample. In further studies, it would be important to increase the diversity of the participants along all of these demographic variables.

Another possible limitation in the interpretability of the validity results from the current study lies in the choice of the CFNI and CMNI as comparison scales. Although strong correlations were found on several of the CFNI and CMNI subscales with all four of the CAGES subscales, the complexity of the analyses given the choice to have women, men and transgender persons fill out both surveys, limits the interpretability of some of the results. There was not a large enough sample of transgender or genderqueer individuals on the CFNI or CMNI to include these participants in the analysis. Interpretations from men on the CFNI and women on the CMNI are tentative at best, given the inapplicability of the wording of many of the items. This wording issue impacts the interpretability of the results to all men and women’s scores on the CFNI and CMNI as well, which left a much smaller sample to use for the analyses of these scales with the CAGES.

Additionally, although the hypothesized salient factor of conformity appears to be supported by the strength of the correlations, a different scale may have been simpler and clearer in connection with the CAGES. Although, this may have proved difficult as the CFNI and CMNI were chosen as best bets because of the similarity to the proposed scale and most other gender scales are not written in a inclusive manner for men and women, and do not consider transgender persons in their population target.
Conclusions

There are several conclusions that can be drawn from the results discussed above from the three studies conducted. First, empirical support was found for the CAGES, a scale constructed to measure comfort or discomfort with conformity or nonconformity for men, women, and transgender persons in the public sphere in their gender expression as manifested through affect, cognitions, and behaviors. This is the first study in the existing literature that measures gender expression, and measures this expression for men, women and transgender persons. It is also the first study to assess comfort and discomfort with conformity or nonconformity with gender norms by separating gender stereotypes from the act of conforming. Second, empirical support was found for the convergent and divergent validity of the CAGES. The CAGES evidenced moderate to strong correlations with two existing conformity measures, the CFNI and CMNI, and evidenced no correlation to another existing scale that does not measure gender expression (the BIDR-IM). Third, strong test re-test reliability was found for three of the four CAGES subscales, with one scale evidencing significant, yet smaller test re-test reliability.

Thus, the CAGES contributes to the existing literature in gender measurement by introducing a scale that measure comfort and conformity of gender expression for men, women, and transgender persons in an inclusive format. It is hoped that the CAGES will be used and will inspire additional research examining conformity, gender expression, and will remain inclusive in its development and execution.
REFERENCES


A measure of sex-role stereotypes and masculinity-femininity, *JSAS Catalog of
Selected Documents in Psychology, 4*, 43-44.

Psychology, 55*, 519-544.


Unpublished manuscript.

Terman, L.M. & Miles, C.C. (1936). *Sex and Personality: Studies in Masculinity and

towards feminism. *Sex Roles. 51*(1/2), 85-90.

masculinity ideology and masculinity-related constructs. *Sex Roles. 27* (11/12),
573-607.


APPENDICES

APPENDIX A:

Comfort and Conformity of Gender Expression Scale (CAGES)

Below are some statements related to gender and how you might feel, act, and think about how you experience gender in the world. Keeping in mind how you think, feel and behave in regards to your gendered self, please indicate how much you personally agree or disagree with each statement by circling SD for "Strongly Disagree", D for "Disagree", A for "Agree", or SA for "Strongly agree" to the left of the statement. There are no right or wrong responses to the statements. You should give the responses that most accurately describe your personal actions, feelings and beliefs.

The following definitions may help you in interpreting the items:

**Gender** is a psychosocial construct most people use to classify a person as male or female, or as gender different. It includes the social actions, expectations and roles associated with biological sex. For example: Men are expected to be masculine and women are expected to be feminine.

**Gender expression** is the expression through actions, thoughts and feelings of the social and cultural roles associated with biological sex.

**Remember to use the following scale when answering the items:**
- Strongly Disagree (SD)
- Disagree (D)
- Agree (A)
- Strongly agree (SA)

1. I am happy when I express myself in ways that are consistent with gender expectations.
2. I feel comfortable talking with others about ways that gender-role stereotypes can be challenged.
3. I feel most comfortable when I make clothing choices that fit expectations of my gender.
4. I would be upset if some people were confused about my gender.
5. I feel uncomfortable about making unconventional clothing choices.
6. I feel uncomfortable expressing gender nonconformity in public.

7. I feel upset when my appearance does not conform to expectations of my gender.

8. I feel uncomfortable if I do not conform to gender expectations in social situations.

9. I feel comfortable engaging in sports that most people perceive to be inconsistent with my gender.

10. I would not be upset if people were confused about my gender.

11. I enjoy challenging societal gender norms in creating my personal spaces (e.g. home/office).

12. I feel most attractive when my appearance matches others’ expectations of my gender.

13. I feel uncomfortable when I express an emotion that is not considered appropriate for my gender.

14. I would feel very uncomfortable using cosmetic surgery to defy gender expectations with my appearance.

15. I am uncomfortable interacting with others in ways that do not conform to gender expectations.

16. I feel most at ease in social environments when I conform to others’ expectations of my gender.

17. I feel most comfortable getting my hair cut in a way that most people perceive to clearly match my gender.

18. I feel most attractive when I challenge expectations of my gender.

19. I feel resentful buying clothing that meets others’ expectations of my gender.

20. I feel uneasy when gender norms inhibit the way I can express myself in public.

21. It upsets me that gender norms influence my behaviors in public.

22. I get upset when I alter my appearance to fit others’ expectations of my gender.

23. I feel upset when I conform to gender stereotypes when meeting new people.
24. I feel most comfortable wearing clothes that do not conform to expectations of my gender.

25. I would feel comfortable obtaining cosmetic surgery that would increase the degree to which my appearance matches gender expectations.

26. I feel comfortable using diet/nutrition in order to have my body appear to others as conforming to my gender.

27. I am comfortable interacting with others in ways that defy gender norms.

28. I feel comfortable purchasing items that are inconsistent with social expectations of my gender.

29. Even if others were to judge me, I feel comfortable expressing my gender in public.

30. I feel uncomfortable using diet/nutrition in order to have my body appear to others as conforming to my gender.

31. Having a haircut/style that defies gender roles would make me feel uncomfortable.

32. Having gender-stereotyped physical features would make me feel uncomfortable.

33. I enjoy using exercise/weight training that makes my appearance more consistent with gender expectations.

34. I would feel upset if my personal environment (e.g. home/office) did not conform to gender expectations.

35. I feel uncomfortable when my appearance does not meet others’ expectations of my gender.

36. I feel uneasy engaging in certain hobbies or activities that are not considered appropriate for my gender.

37. I feel most at ease interacting with others in ways that are perceived to be consistent with my gender.

38. I feel most at ease when my personal environment (e.g. home, office) reflects what most people expect of my gender.

39. I feel uncomfortable purchasing things that most people might consider inappropriate for my gender.
40. Engaging in certain hobbies or activities that are considered stereotypical for my gender makes me feel uncomfortable.

41. I feel most at ease using diet/nutrition to make my body appear more consistent with gender expectations.
APPENDIX B

Expert Feedback Questions

1. Initial Reactions to taking the CAGES:

2. In what ways do you experience your expression of gender?

3. Feedback on the language and phrasing of the items of the CAGES:

4. Is it clear what the CAGES is measuring from the items included on the scale? If so, suggestions for additional items or domains?

5. Suggestions for improvements to the CAGES?
APPENDIX C

Balanced Inventory of Desirable Responding (BIDR)

Balanced Inventory of Desirable Responding
Paulhus (1991)

Instructions: Using the scale below as a guide, write a number beside each statement to indicate how much you agree with it.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Not True</td>
<td></td>
<td></td>
<td>Somewhat True</td>
<td></td>
<td></td>
<td>Very True</td>
</tr>
</tbody>
</table>

*1. I sometimes tell lies if I have to.
2. I never cover up my mistakes.
*3. There have been occasions when I have taken advantage of someone.
4. I never swear.
*5. I sometimes try to get even rather than forgive and forget.
6. I always obey laws, even if I'm unlikely to get caught.
*7. I have said something bad about a friend behind his or her back.
8. When I hear people talking privately, I avoid listening.
*9. I have received too much change from a salesperson without telling him or her.
10. I always declare everything at customs.
*11. When I was young I sometimes stole things.
12. I have never dropped litter on the street.
*13. I sometimes drive faster than the speed limit.
14. I never read sexy books or magazines.
*15. I have done things that don't I don't tell other people about.
16. I never take things that don't belong to me.

*17. I have taken sick-leave from work or school even though I wasn't really sick.

18. I have never damaged a library book or store merchandise without reporting it.

*19. I have some pretty awful habits.

20. I don't gossip about other people's business.
Solicitation email:

Hello,

My name is Katie Spencer, and I am a doctoral student in Counseling Psychology at the University of Missouri-Columbia. I am currently working on my dissertation under the direction of Dr. Roger Worthington. The goal of my dissertation project is to explore gender identification and expression. We are seeking to examine the role gender plays in individuals' lives.

I am emailing to request your participation in this study. If you would like to participate, please click on the link below. You may also forward this message to others who might be interested in this research. The surveys take approximately 10 minutes to complete. Please remember to fill out each item fully and to close out your internet browser after completing the survey.

If you are willing to do this, you can cut this part of the email out and forward what follows.

Thank you for your time and consideration,

Katie Spencer, M.A.
Doctoral Student
Department of Educational, School and Counseling Psychology
16 Hill Hall
Columbia, MO 65201
Appendix E

Project Information and Informed Consent

Project Information and Informed Consent
(web version)

The goal of this project is to explore how individuals view different thoughts, behaviors and emotions related to their gender. In order to study this relationship we will be asking participants to provide some demographic information and respond to a series of questions. This survey should take you approximately 5-15 minutes to complete.

As a participant in this research, you should read and understand the following statements:

- Your participation in this research is VOLUNTARY. You are not required to answer every question that might be asked. This means that you are free to stop participating at any point without penalty or loss of privilege, except for benefits directly related to your participation in this study.

- All participant responses will be completely ANONYMOUS. In order to assure anonymity, please do not put your name (or any other identifying information) anywhere on the accompanying questionnaires.

- Because this research is ANONYMOUS, you will not be identified in any presentation or publication of this research. All information you provide will be combined with the data from other respondents and reported as grouped data.

- In order to assure ANONYMITY, while at the same time facilitating our efforts to obtain a high quality data set, we have developed the following procedure:

  1. There are no codes or any other information contained on the questionnaire or any other materials associated with it that identifies you as an individual respondent to this survey.

  2. However, in order to ensure that our data does not include duplications or multiple submissions from the same individual, we will retrieve and record the IP address of each computer from which data is submitted, along with a time/date stamp that records when the data was submitted. The IP address and time/date stamp information will serve only to identify duplicate or multiple submissions. Although it is conceivable that the IP address could be used to gain access to the location of the computer used to submit data, the information WILL NOT be used in this way. Further, it is nearly impossible to ascertain the identity of the individual using any
particular computer. Finally, this is a highly unlikely scenario, and one that is not intended by the research investigators.

- You have a right to be informed of all potential risks associated with your participation in this research. There is no more than minimal risk associated with participation in this survey. Possible psychological risks are likely to be small and unlikely to occur. You may at any time discontinue participation.

**NOTE:** Because the research questionnaire requests you to provide information about yourself that you may not want other people to know, there is a risk associated with the unlikely chance that somebody else might view the information you provide. For example, you should protect yourself from the types of occurrences identified below:

1. There is a possibility that your responses could be viewed by an outside party if you do not **EXIT/CLOSE** your Internet browser (e.g., Netscape Navigator, Internet Explorer, etc.) as soon as you finish responding to the questionnaire because your responses might be visible if you (or someone else) click the **BACK** button on the browser. In order to **ELIMINATE** this possibility, you should **EXIT/CLOSE** the browser as soon as you finish responding to the survey and have submitted your responses.

2. There is a possibility that your responses could be viewed by an outside party if you leave your browser on and leave the computer terminal before finishing the questionnaire (e.g., answer the phone, leave the computer unattended, etc.). In order to avoid inadvertent access to your responses by a third party, do not leave the terminal or stop responding to the questionnaire until you have completely finished and closed the browser.

- If you have questions or concerns about this research or your participation, please contact one of the following:

<table>
<thead>
<tr>
<th>Katherine Spencer</th>
<th>Roger L. Worthington, Ph.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Graduate Student</td>
<td>Associate Professor</td>
</tr>
<tr>
<td>16 Hill Hall</td>
<td>16 Hill Hall</td>
</tr>
<tr>
<td>Educational, School and Counseling Psychology</td>
<td>Educational, School and Counseling Psychology</td>
</tr>
<tr>
<td>University of Missouri</td>
<td>University of Missouri</td>
</tr>
<tr>
<td>Columbia, MO 65211</td>
<td>Columbia, MO 65211</td>
</tr>
<tr>
<td><a href="mailto:kgs91d@mizzou.edu">kgs91d@mizzou.edu</a></td>
<td><a href="mailto:Worthingtonr@missouri.edu">Worthingtonr@missouri.edu</a></td>
</tr>
</tbody>
</table>

For additional information regarding human participation in research, please feel free to contact the UMC Campus IRB Office at 573-882-9585.
Project Information and Informed Consent

(paper version)

The goal of this project is to explore how individuals view different thoughts, behaviors and emotions related to their gender. In order to study this relationship we will be asking participants to provide some demographic information and respond to a series of questions. This survey should take you approximately 5-15 minutes to complete.

As a participant in this research, you should read and understand the following statements:

- Your participation in this research is VOLUNTARY. You are not required to answer every question that might be asked. This means that you are free to stop participating at any point without penalty or loss of privilege, except for benefits directly related to your participation in this study.

- All participant responses will be completely ANONYMOUS. In order to assure anonymity, please do not put your name (or any other identifying information) anywhere on the accompanying questionnaires.

- Because this research is ANONYMOUS, you will not be identified in any presentation or publication of this research. All information you provide will be combined with the data from other respondents and reported as grouped data.

- You have a right to be informed of all potential risks associated with your participation in this research. There is no more than minimal risk associated with participation in this survey. Possible psychological risks are likely to be small and unlikely to occur. You may at any time discontinue participation.

- If you have questions or concerns about this research or your participation, please contact one of the following:

<table>
<thead>
<tr>
<th>Katherine Spencer</th>
<th>Roger L. Worthington, Ph.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Graduate Student</td>
<td>Associate Professor</td>
</tr>
<tr>
<td>16 Hill Hall</td>
<td>16 Hill Hall</td>
</tr>
<tr>
<td>Educational, School and Counseling Psychology</td>
<td>Educational, School and Counseling Psychology</td>
</tr>
<tr>
<td>University of Missouri</td>
<td>University of Missouri</td>
</tr>
<tr>
<td>Columbia, MO 65211</td>
<td>Columbia, MO 65211</td>
</tr>
<tr>
<td><a href="mailto:kgs91d@mizzou.edu">kgs91d@mizzou.edu</a></td>
<td><a href="mailto:Worthingtonr@missouri.edu">Worthingtonr@missouri.edu</a></td>
</tr>
</tbody>
</table>

For additional information regarding human participation in research, please feel free to contact the UMC Campus IRB Office at 573-882-9585.
APPENDIX F
Demographics Form

Demographics Form

Directions: Please answer the following questions about yourself.

1. Gender/Sex: (circle as many as apply)
   - Male
   - Female
   - Transgender
   - Intersex
   - Genderqueer
   - Other_______

2. Age_____

3. Race/Ethnicity:
   - African American/Black
   - Latino/a/Hispanic/Chicano/a
   - Asian American
   - Native American
   - European American/White
   - Biracial/Multiracial
   - Other_____________

4. Sexual Orientation:
   - Lesbian
   - Gay
   - Queer
   - Heterosexual/Straight
   - Bisexual
   - Other

5. Highest degree:
   a. Less than high school diploma
   b. High school diploma
   c. Some college
   d. 4-year college degree
   e. Post-baccalaureate study
   f. Master’s degree
   g. Doctoral degree
   h. Other professional degree (MD, JD, DO, etc.)
APPENDIX G

Debriefing Form
(web version)

Thank you participating in this study. The purpose of this study is to examine individuals' levels of comfort with expressing the range of their gender within the context of the social environment. Please remember that there is a possibility that your responses could be viewed by an outside party if you do not EXIT/CLOSE your Internet browser (e.g., Netscape Navigator, Internet Explorer, etc.) as soon as you finish responding to the questionnaire because your responses might be visible if you (or someone else) click the BACK button on the browser. In order to ELIMINATE this possibility, you should EXIT/CLOSE the browser as soon as you finish responding to the survey and have submitted your responses. Again, if you have any questions/concerns regarding this study please contact one of the following:

<table>
<thead>
<tr>
<th>Katherine Spencer</th>
<th>Roger L. Worthington, Ph.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Graduate Student</td>
<td>Associate Professor</td>
</tr>
<tr>
<td>16 Hill Hall</td>
<td>16 Hill Hall</td>
</tr>
<tr>
<td>Educational, School and Counseling Psychology</td>
<td>Educational, School and Counseling Psychology</td>
</tr>
<tr>
<td>University of Missouri</td>
<td>University of Missouri</td>
</tr>
<tr>
<td>Columbia, MO 65211</td>
<td>Columbia, MO 65211</td>
</tr>
<tr>
<td><a href="mailto:kgs91d@mizzou.edu">kgs91d@mizzou.edu</a></td>
<td><a href="mailto:Worthingtonr@missouri.edu">Worthingtonr@missouri.edu</a></td>
</tr>
</tbody>
</table>

For additional information regarding human participation in research, please feel free to contact the UMC Campus IRB Office at 573-882-9585.
Thank you participating in this study. The purpose of this study is to examine individuals’ levels of comfort with expressing the range of their gender within the context of the social environment. Again, if you have any questions/concerns regarding this study please contact one of the following:

<table>
<thead>
<tr>
<th>Katherine Spencer</th>
<th>Roger L. Worthington, Ph.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Graduate Student</td>
<td>Associate Professor</td>
</tr>
<tr>
<td>16 Hill Hall</td>
<td>16 Hill Hall</td>
</tr>
<tr>
<td>Educational, School and Counseling Psychology</td>
<td>Educational, School and Counseling Psychology</td>
</tr>
<tr>
<td>University of Missouri</td>
<td>University of Missouri</td>
</tr>
<tr>
<td>Columbia, MO 65211</td>
<td>Columbia, MO 65211</td>
</tr>
<tr>
<td><a href="mailto:kgs91d@mizzou.edu">kgs91d@mizzou.edu</a></td>
<td><a href="mailto:Worthingtonr@missouri.edu">Worthingtonr@missouri.edu</a></td>
</tr>
</tbody>
</table>

For additional information regarding human participation in research, please feel free to contact the UMC Campus IRB Office at 573-882-9585.
APPENDIX H

Conformity to Feminine Norms Inventory (CFNI)

CFNI Instructions

Please read carefully

The following pages contain a series of statements about how people might think, feel or behave. The statements are designed to measure attitudes, beliefs, and behaviors associated with both traditional and non-traditional feminine gender roles. For example, the statements are about issues such as appearance, taking care of others, sexuality, and relationships.

Thinking about your own actions, feelings and beliefs, please indicate how much you personally agree or disagree with each statement by circling SD for "Strongly Disagree", D for "Disagree", A for "Agree", or SA for "Strongly agree" to the right of the statement.

EXAMPLE ITEM:

Being nice to others is extremely important

Circle SD if you strongly disagree with the statement.
Circle D if you disagree with the statement.
Circle A if you agree with the statement, or
Circle SA if you strongly agree with the statement.

There are no right or wrong responses to the statements. You should give the responses that most accurately describe your personal actions, feelings and beliefs. It is best if you respond with your first impression when answering.

1. It is important to let people know they are special
2. I would baby-sit for fun
3. I would be happier if I was thinner
4. I would feel extremely ashamed if I had many sexual partners
5. I feel uncomfortable being singled out for praise
6. When I am in a romantic relationship, I give it all my energy
7. It is important to keep your living space clean
8. I spend more than 30 minutes a day doing my hair and make-up
9. Putting energy into friendships is a waste of time
10. I participate in activities that include kids
11. I don’t tend to worry about gaining weight
12. If I was single, I would want to date a lot of people
13. Being mean gets you ahead in life
14. I tell everyone about my accomplishments
15. Whether I’m in one or not, romantic relationships are often on my mind.
16. I clean my home on a regular basis
17. I feel attractive without makeup
18. I believe that my friendships should be maintained at all costs
19. I find children annoying
20. Being thin is important
21. I prefer long-term relationships to casual sexual ones
22. There is nothing wrong with bragging
23. I pity people who are single
24. I am comfortable when my living space is a little cluttered
25. I’d feel superficial if I wore make-up
26. I feel good about myself when others know that I care about them
27. Taking care of kids is just not for me
28. I would only diet if a doctor ordered me to do so
29. I would feel guilty if I had a one-night stand
30. When I succeed, I tell my friends about it
31. Having a romantic relationship is essential in life
32. I enjoy spending time making my living space look nice
33. Being nice to others is extremely important
34. I regularly wear makeup
35. I don’t go out of my way to keep in touch with friends
36. Most people enjoy children more than I do
37. I would like to lose a few pounds
38. It is impossible to always be nice to others
39. It is not necessary to be in a committed relationship to have sex
40. I hate telling people about my accomplishments
41. I can be happy without being in a romantic relationship
42. I haven’t cleaned my living space in the past week SD D A SA
43. I get ready in the morning without looking in the mirror very much SD D A SA
44. I would feel burdened if I had to maintain a lot of friendships SD D A SA
45. When I want to relax, I don’t want to be around kids SD D A SA
46. I tend to watch what I eat in order to stay thin SD D A SA
47. I would feel comfortable having casual sex SD D A SA
48. I make it a point to get together with my friends regularly SD D A SA
49. I always downplay my achievements SD D A SA
50. Being in a romantic relationship is important SD D A SA
51. I don’t care if my living space looks messy SD D A SA
52. I never wear make-up SD D A SA
53. I always try to make people feel special SD D A SA
54. Caring for children adds meaning to one’s life SD D A SA
55. I’d look better if I put on a few pounds SD D A SA
56. I frequently change sexual partners SD D A SA
57. I am not afraid to tell people about my achievements SD D A SA
58. My life plans do not rely on my having a romantic relationship SD D A SA
59. I do all of the cleaning, cooking and decorating where I live SD D A SA
60. It is important to look physically attractive in public SD D A SA
61. If a friendship isn’t working, I’ll end it SD D A SA
62. I would feel empty if my life did not involve children SD D A SA
63. I try to be sweet and nice SD D A SA
64. I am always trying to lose weight SD D A SA
65. I would only have sex with the person I love SD D A SA
66. I don’t seek recognition for my efforts SD D A SA
67. When I have a romantic relationship, I enjoy focusing my energies on it SD D A SA
68. There is no point to cleaning because things will get dirty again SD D A SA
69. I am not afraid to hurt people’s feelings to get what I want SD D A SA
70. Taking care of children is extremely fulfilling SD D A SA
71. I would be perfectly happy with myself even if I gained weight SD D A SA
72. It would be enjoyable to date more than one person at a time SD D A SA
73. I enjoy being in the spotlight SD D A SA
74. If I were single, my life would be complete without a partner SD D A SA
75. I rarely go out of my way to act nice  
76. I actively avoid children
77. I am terrified of gaining weight
78. I would only have sex if I was in a committed relationship like marriage
79. I am only nice to people I like
80. I like being around children
81. I tend to eat whatever I want
82. I don’t feel guilty if I lose contact with a friend
83. I feel uneasy around children
84. I would be ashamed if someone thought I was mean

*Please check to make sure you have answered all the items*
APPENDIX I

Conformity to Masculine Norms Inventory (CMNI)

CMNI Instructions

*Please read carefully*

The following pages contain a series of statements about how men might think, feel or behave. The statements are designed to measure attitudes, beliefs, and behaviors associated with both traditional and non-traditional masculine gender roles.

**Thinking about your own actions, feelings and beliefs**, please indicate how much you personally agree or disagree with each statement by circling SD for "Strongly Disagree", D for "Disagree", A for "Agree", or SA for "Strongly agree" to the left of the statement. There are no right or wrong responses to the statements. You should give the responses that most accurately describe your personal actions, feelings and beliefs. It is best if you respond with your first impression when answering.

<table>
<thead>
<tr>
<th>Statement</th>
<th>SD</th>
<th>D</th>
<th>A</th>
<th>SA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. It is best to keep your emotions hidden</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. In general, I will do anything to win</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. If I could, I would frequently change sexual partners</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. If there is going to be violence, I find a way to avoid it</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. It is important to me that people think I am heterosexual</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. In general, I must get my way</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Trying to be important is the greatest waste of time</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. I am often absorbed in my work</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. I will only be satisfied when women are equal to men</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. I hate asking for help</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Taking dangerous risks helps me to prove myself</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. In general, I do not expend a lot of energy trying to win at things</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. An emotional bond with a partner is the best part of sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14. I should take every opportunity to show my feelings</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. I believe that violence is never justified</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16. Being thought of as gay is not a bad thing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17. In general, I do not like risky situations</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18. I should be in charge</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19. Feelings are important to show</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
20. I feel miserable when work occupies all my attention
21. I feel best about my relationships with women when we are equals
22. Winning is not my first priority
23. I make sure that people think I am heterosexual
24. I enjoy taking risks
25. I am disgusted by any kind of violence
26. I would hate to be important
27. I love to explore my feelings with others
28. If I could, I would date a lot of different people
29. I ask for help when I need it
30. My work is the most important part of my life
31. Winning isn’t everything, it’s the only thing
32. I never take chances
33. I would only have sex if I was in a committed relationship
34. I like fighting
35. I treat women as equals
36. I bring up my feelings when talking to others
37. I would be furious if someone thought I was gay
38. I only get romantically involved with one person
39. I don't mind losing
40. I take risks
41. I never do things to be an important person
42. It would not bother me at all if someone thought I was gay
43. I never share my feelings
44. Sometimes violent action is necessary
45. Asking for help is a sign of failure
46. In general, I control the women in my life
47. I would feel good if I had many sexual partners
48. It is important for me to win
49. I don't like giving all my attention to work
50. I feel uncomfortable when others see me as important
51. It would be awful if people thought I was gay
52. I like to talk about my feelings
53. I never ask for help
54. More often than not, losing does not bother me
55. It is foolish to take risks
56. Work is not the most important thing in my life
57. Men and women should respect each other as equals
58. Long term relationships are better than casual sexual encounters
59. Having status is not very important to me
60. I frequently put myself in risky situations
61. Women should be subservient to men
62. I am willing to get into a physical fight if necessary
63. I like having gay friends
64. I feel good when work is my first priority
65. I tend to keep my feelings to myself
66. Emotional involvement should be avoided when having sex
67. Winning is not important to me
68. Violence is almost never justified
69. I am comfortable trying to get my way
70. I am happiest when I'm risking danger
71. Men should not have power over women
72. It would be enjoyable to date more than one person at a time
73. I would feel uncomfortable if someone thought I was gay
74. I am not ashamed to ask for help
75. The best feeling in the world comes from winning
76. Work comes first
77. I tend to share my feelings
78. I like emotional involvement in a romantic relationship
79. No matter what the situation I would never act violently
80. If someone thought I was gay, I would not argue with them about it
81. Things tend to be better when men are in charge
82. I prefer to be safe and careful
83. A person shouldn't get tied down to dating just one person
84. I tend to invest my energy in things other than work
85. It bothers me when I have to ask for help
86. I love it when men are in charge of women  
87. It feels good to be important  
88. I hate it when people ask me to talk about my feelings  
89. I work hard to win  
90. I would only be satisfied with sex if there was an emotional bond  
91. I try to avoid being perceived as gay  
92. I hate any kind of risk  
93. I prefer to stay unemotional  
94. I make sure people do as I say

*Please check to make sure you have answered all the items*
APPENDIX J

Study Three Data Collection Protocol

Data Collection Protocol: CAGES Test-Retest Study 3

Researcher:

“Hello, my name is _______, and I am here collecting data for a dissertation project that examines gender issues. If you would like to participate in this research, please take a survey. It will take about 5-10 minutes to complete. Your participation is completely VOLUNTARY, and is in no way associated with this class or your grade. You do not have to participate if you do not want to.

If you choose to participate, you will see on the inside sheet there is a space for an ID code. This code is composed of the first two letters of your mother’s maiden name, and the last four numbers of your phone number. This code will be used to match your responses, but will not be able to used to identify you directly in any way. This code is used so that your responses cannot be traced back to any individual, thus keeping your responses ANONYMOUS.

If you have any questions you may contact the researcher, whose contact information is located on the first page of this packet. Thank you for participating in this important research.”
APPENDIX K

Glossary of Terms

**Gender expression**: The expression of social and cultural roles associated with biological sex, how we communicate and experience gender through behavior, thoughts and feelings.

**Sex**: either of the two major forms of individuals that occur in many species and that are distinguished respectively as female or male; the anatomy and biology that determines whether one is male, female, or intersexed (formerly called *hermaphroditic*).

**Gender**: is a psychosocial construct that includes the behavioral, cultural, or psychological traits typically associated with one sex.

**Gender Identity**: a person's sense of their own gender, which is communicated to others by their Gender Expression. Since most people conform to societal gender norms, they have a Gender Identity congruent with their Gender Expression. However, Gender, like sexuality, is fluid and can change over time, in individuals and in human society. For some people, Gender Identity, Gender Expression and sex do not correspond with each other. Those who cannot or choose not to conform to societal gender norms associated with their physical sex are *gender diverse/genderqueer*.

**Transgender**: an umbrella term used to describe gender diverse people who have gender identities, expressions or behaviors not traditionally associated with their birth sex. Transgender also can mean anyone who *transcends* the conventional definitions of 'man' and 'woman'. Thus transgender also can include butch lesbians, radical faeries, drag queens, drag kings, femmes, and many other kinds of gender diverse people who use a variety of terms to self-identify.

**Queer**: To say that someone is "queer" indicates an indeterminacy or indecipherability about their sexuality and gender, a sense that they cannot be categorized without a careful contextual examination. Queer is also used by individuals within LGBT communities to indicate an inherent questioning of gender and sexuality binaries.

**Genderqueer**: Individual who views the two-gender (masculine and feminine) system as limiting. Genderqueers usually place themselves outside of such a system. May also identify as Transgender, and may alternatively describe themselves as lacking gender entirely.
APPENDIX L

Final Comfort and Conformity of Gender Expression Scale (CAGES) Items by Factor

**Discomfort With Nonconformity (DWN)**

1. I feel comfortable talking with others about ways that gender-role stereotypes can be challenged. (R)

2. I feel uncomfortable if I do not conform to gender expectations in social situations.

3. I enjoy challenging societal gender norms in creating my personal spaces (e.g. home/office). (R)

4. I am uncomfortable interacting with others in ways that do not conform to gender expectations.

5. I am comfortable interacting with others in ways that defy gender norms. (R)

6. I feel comfortable purchasing items that are inconsistent with social expectations of my gender. (R)

7. I would feel upset if my personal environment (e.g. home/office) did not conform to gender expectations.

8. I feel uncomfortable when my appearance does not meet others' expectations of my gender.

9. I feel uneasy engaging in certain hobbies or activities that are not considered appropriate for my gender.

10. I feel most at ease when my personal environment (e.g. home, office) reflects what most people expect of my gender.

11. I feel uncomfortable purchasing things that most people might consider inappropriate for my gender.

**Resentful Conformity (RC)**

1. I feel uneasy when gender norms inhibit the way I can express myself in public.

2. It upsets me that gender norms influence my behaviors in public.

3. I feel upset when I conform to gender stereotypes when meeting new people.
Active-Physical Comfort Conformity (APCC)

1. I feel comfortable using diet/nutrition in order to have my body appear to others as conforming to my gender.

2. I enjoy using exercise/weight training that makes my appearance more consistent with gender expectations.

3. I feel most at ease using diet/nutrition to make my body appear more consistent with gender expectations.

Comfort with Conformity-Appearance (CCA)

1. I feel most comfortable wearing clothes that do not conform to expectations of my gender.

2. I feel most comfortable getting my hair cut in a way that most people perceive to clearly match my gender.

3. I feel most comfortable when I make clothing choices that fit expectations of my gender.

4. I feel most attractive when my appearance matches others' expectations of my gender.

5. I am happy when I express myself in ways that are consistent with gender expectations.

6. I feel most attractive when I challenge expectations of my gender.

7. I feel resentful buying clothing that meets others' expectations of my gender.
Table 1
*Item-Total Correlations, Factor Matrix, Communalities, and Item Means and Standard Deviations for the CAGES*

<table>
<thead>
<tr>
<th>Item</th>
<th>Item Total Correlations</th>
<th>Factor Loading</th>
<th>h²</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>It upsets me that gender norms influence my behaviors in public.</td>
<td>-.31</td>
<td>.85</td>
<td>.76</td>
<td>2.22</td>
<td>.80</td>
</tr>
<tr>
<td>I feel comfortable using diet/nutrition in order to have my body</td>
<td>.31</td>
<td>.83</td>
<td>.65</td>
<td>2.59</td>
<td>.78</td>
</tr>
<tr>
<td>appear to others as conforming to my gender.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I feel most at ease using diet/nutrition to make my body appear</td>
<td>.44</td>
<td>.82</td>
<td>.68</td>
<td>2.71</td>
<td>.70</td>
</tr>
<tr>
<td>more consistent with gender expectations.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I feel most comfortable wearing clothes that do not conform to</td>
<td>.52</td>
<td>.75</td>
<td>.59</td>
<td>2.25</td>
<td>.79</td>
</tr>
<tr>
<td>expectations of my gender. (R)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I feel comfortable purchasing items that are inconsistent with</td>
<td>.59</td>
<td>.72</td>
<td>.56</td>
<td>2.98</td>
<td>.69</td>
</tr>
<tr>
<td>social expectations of my gender. (R)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I feel uneasy when gender norms inhibit the way I can express</td>
<td>-.42</td>
<td>.72</td>
<td>.68</td>
<td>2.10</td>
<td>.76</td>
</tr>
<tr>
<td>myself in public.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I am comfortable interacting with others in ways that defy</td>
<td>.60</td>
<td>.71</td>
<td>.59</td>
<td>2.94</td>
<td>.66</td>
</tr>
<tr>
<td>gender norms. (R)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I feel most comfortable when I make clothing choices that fit</td>
<td>.49</td>
<td>.71</td>
<td>.51</td>
<td>2.27</td>
<td>.80</td>
</tr>
<tr>
<td>expectations of my gender.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I feel uncomfortable purchasing things that most people might</td>
<td>.54</td>
<td>.69</td>
<td>.49</td>
<td>2.98</td>
<td>.66</td>
</tr>
<tr>
<td>consider inappropriate for my gender.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I feel most attractive when my appearance matches others’ expectations of my gender.</td>
<td>.68</td>
<td>.65</td>
<td>.64</td>
<td>2.35</td>
<td>.79</td>
</tr>
<tr>
<td>I feel uncomfortable if I do not conform to gender expectations in social situations.</td>
<td>.53</td>
<td>.63</td>
<td>.46</td>
<td>2.73</td>
<td>.73</td>
</tr>
<tr>
<td>I enjoy using exercise/ weight training that makes my appearance</td>
<td>.40</td>
<td>.63</td>
<td>.44</td>
<td>2.51</td>
<td>.73</td>
</tr>
<tr>
<td>more consistent with gender expectations.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I feel uneasy engaging in certain hobbies or activities that are not</td>
<td>.48</td>
<td>.60</td>
<td>.35</td>
<td>3.10</td>
<td>.65</td>
</tr>
<tr>
<td>considered appropriate for my gender.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I feel comfortable talking with others about ways that gender-</td>
<td>.33</td>
<td>.59</td>
<td>.32</td>
<td>3.38</td>
<td>.67</td>
</tr>
<tr>
<td>role stereotypes can be challenged. (R)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note: N = 356.*
Table 1 (continued)
*Item-Total Correlations, Factor Matrix, Communalities, and Item Means and Standard Deviations for the CAGES*

<table>
<thead>
<tr>
<th>Item</th>
<th>Item Total Correlations</th>
<th>Factor Loading</th>
<th>h²</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>I am uncomfortable interacting with others in ways that do not conform to gender expectations.</td>
<td>.55</td>
<td>.59</td>
<td>.41</td>
<td>3.02</td>
<td>.64</td>
</tr>
<tr>
<td>I feel resentful buying clothing that meets others’ expectations of my gender. (R)</td>
<td>.41</td>
<td>.58</td>
<td>.50</td>
<td>2.24</td>
<td>.84</td>
</tr>
<tr>
<td>I feel upset when I conform to gender stereotypes when meeting new people.</td>
<td>-.43</td>
<td>.56</td>
<td>.54</td>
<td>2.38</td>
<td>.81</td>
</tr>
<tr>
<td>I feel most attractive when I challenge expectations of my gender. (R)</td>
<td>.55</td>
<td>.56</td>
<td>.46</td>
<td>2.23</td>
<td>.72</td>
</tr>
<tr>
<td>I am happy when I express myself in ways that are consistent with gender expectations</td>
<td>.47</td>
<td>.52</td>
<td>.37</td>
<td>2.34</td>
<td>.71</td>
</tr>
<tr>
<td>I would feel upset if my personal environment (e.g. home/office) did not conform to gender expectations.</td>
<td>.54</td>
<td>.50</td>
<td>.41</td>
<td>2.95</td>
<td>.62</td>
</tr>
<tr>
<td>I feel uncomfortable when my appearance does not meet others’ expectations of my gender.</td>
<td>.59</td>
<td>.50</td>
<td>.46</td>
<td>2.73</td>
<td>.67</td>
</tr>
<tr>
<td>I feel most comfortable getting my hair cut in a way that most people perceive to clearly match my gender.</td>
<td>.57</td>
<td>.49</td>
<td>.43</td>
<td>2.44</td>
<td>.84</td>
</tr>
<tr>
<td>I feel most at ease when my personal environment (e.g. home, office) reflects what most people expect of my gender.</td>
<td>.59</td>
<td>.37</td>
<td>.52</td>
<td>2.65</td>
<td>.67</td>
</tr>
<tr>
<td>I enjoy challenging societal gender norms in creating my personal spaces (e.g. home/office). (R)</td>
<td>.50</td>
<td>.37</td>
<td>.40</td>
<td>2.77</td>
<td>.74</td>
</tr>
</tbody>
</table>

*Note: N = 356.*
Table 2
*Intercorrelations Between Subscales of CAGES and Validity Instruments: Women on the CFNI*

<table>
<thead>
<tr>
<th>Measures</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Discomfort with Nonconformity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Resentful Conformity</td>
<td>-.46**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Active-Physical Comfort Conformity</td>
<td>.38**</td>
<td>-.38**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Comfort with Conformity-Appearance</td>
<td>.55**</td>
<td>-.61**</td>
<td>.43**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. CFNI HNR</td>
<td>-.09</td>
<td>.08</td>
<td>-.11</td>
<td>.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. CFNI IWC</td>
<td>.28*</td>
<td>-.37**</td>
<td>-.10</td>
<td>.03</td>
<td>.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. CFNI THIN</td>
<td>.28*</td>
<td>-.23</td>
<td>.52**</td>
<td>.22</td>
<td>-.10</td>
<td>-.19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. CFNI SEXFID</td>
<td>.51**</td>
<td>-.16</td>
<td>.16</td>
<td>.14</td>
<td>.14</td>
<td>.33*</td>
<td>-.06</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. CFNI MOD</td>
<td>-.04</td>
<td>.34*</td>
<td>-.13</td>
<td>-.28*</td>
<td>.11</td>
<td>.01</td>
<td>-.12</td>
<td>.26</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. CFNI IRR</td>
<td>.22</td>
<td>-.21</td>
<td>.05</td>
<td>.13</td>
<td>-.33*</td>
<td>.08</td>
<td>.09</td>
<td>.34*</td>
<td>-.16</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. CFNI DOMS</td>
<td>.16</td>
<td>-.07</td>
<td>.00</td>
<td>-.15</td>
<td>.25</td>
<td>.15</td>
<td>-.23</td>
<td>.15</td>
<td>.09</td>
<td>.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. CFNI IIA</td>
<td>.16</td>
<td>.08</td>
<td>.08</td>
<td>.37**</td>
<td>.35*</td>
<td>-.20</td>
<td>.12</td>
<td>.06</td>
<td>-.02</td>
<td>-.12</td>
<td>.10</td>
<td></td>
</tr>
</tbody>
</table>

*Note: N = 55. CFNI = Conformity to Feminine Norms Inventory. HNR = Have Nice Relationships, IWC = Involvement with Children, THIN = Thinness, SEXFID = Sexual Fidelity, MOD = Modesty, IRR = Investment in Romantic Relationships, DOMS = Domestic, IIA = Investment in Appearance. *p < .05, **p < .01*
Table 2 (continued)
Intercorrelations Between Subscales of CAGES and Validity Instruments: Men on the CMNI

<table>
<thead>
<tr>
<th>Measures</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discomfort with Nonconformity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resentful Conformity</td>
<td>-36**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Active-Physical Comfort Conformity</td>
<td></td>
<td>.33**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comfort with Conformity-Appearance</td>
<td></td>
<td></td>
<td>.66**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CMNI WIN</td>
<td>.26</td>
<td>-.16</td>
<td>.16</td>
<td>.30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CMNI EMOC</td>
<td>.49**</td>
<td>-.30</td>
<td>.43*</td>
<td>.36*</td>
<td>.15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CMNI RISK</td>
<td>-.34</td>
<td>.30</td>
<td>-.02</td>
<td>-.15</td>
<td>-.21</td>
<td>-.39*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CMNI VIOL</td>
<td>.22</td>
<td>.17</td>
<td>.45*</td>
<td>.06</td>
<td>.10</td>
<td>.08</td>
<td>.32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CMNI POW</td>
<td>.56**</td>
<td>-.43*</td>
<td>.41*</td>
<td>.32</td>
<td>.25</td>
<td>.23</td>
<td>-.24</td>
<td>.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CMNI DOM</td>
<td>.05</td>
<td>-.13</td>
<td>.10</td>
<td>.13</td>
<td>.39*</td>
<td>.35*</td>
<td>.25</td>
<td>.20</td>
<td>.27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CMNI PLAY</td>
<td>.08</td>
<td>.18</td>
<td>-.20</td>
<td>-.20</td>
<td>-.16</td>
<td>-.06</td>
<td>.22</td>
<td>.06</td>
<td>.23</td>
<td>.20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CMNI SELF</td>
<td>.25</td>
<td>-.25</td>
<td>.34</td>
<td>.26</td>
<td>.31</td>
<td>.47**</td>
<td>-.10</td>
<td>-.09</td>
<td>.20</td>
<td>.48**</td>
<td>-.13</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CMNI WORK</td>
<td>-.25</td>
<td>.13</td>
<td>-.19</td>
<td>-.29</td>
<td>.33</td>
<td>.04</td>
<td>-.23</td>
<td>-.08</td>
<td>-.06</td>
<td>.15</td>
<td>-.29</td>
<td>.17</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CMNI HOMO</td>
<td>.78**</td>
<td>-.41</td>
<td>.51**</td>
<td>.53**</td>
<td>.29</td>
<td>.54</td>
<td>-.26</td>
<td>.30</td>
<td>.46**</td>
<td>.22</td>
<td>-.03</td>
<td>.25</td>
<td>-.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CMNI POS</td>
<td>-.44*</td>
<td>.10</td>
<td>.29</td>
<td>-.30</td>
<td>.07</td>
<td>.03</td>
<td>.22</td>
<td>-.19</td>
<td>.02</td>
<td>.34</td>
<td>.18</td>
<td>.10</td>
<td>.12</td>
<td>-.46**</td>
<td></td>
</tr>
</tbody>
</table>

Note. N = 33 CMNI = Conformity to Masculine Norms Inventory, WIN = Winning, EMOC = Emotional Control, RISK = Risk-taking, VIOL = Violence, POW = Power Over Women, DOM = Dominance, PLAY = Playboy, SELF = Self-Reliant, WORK = Primacy of Work, HOMO = Disdain for Homosexuals, POS = Pursuit of Status. *p < .05, **p < .01.
Table 2 (continued)

Intercorrelations Between Subscales of CAGES and Validity Instruments: Men on the CFNI

<table>
<thead>
<tr>
<th>Measures</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discomfort with Nonconformity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resentful Conformity</td>
<td>-.36**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Active-Physical Comfort Conformity</td>
<td>.33**</td>
<td>-.27*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comfort with Conformity-Appearance</td>
<td>.66**</td>
<td>-.58**</td>
<td>.28*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CFNI HNR</td>
<td>-.09</td>
<td>.25</td>
<td>.12</td>
<td>-.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CFNI IWC</td>
<td>.04</td>
<td>-.26</td>
<td>.29</td>
<td>-.08</td>
<td>.07</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CFNI THIN</td>
<td>-.22</td>
<td>-.20</td>
<td>-.08</td>
<td>-.19</td>
<td>-.33</td>
<td>.16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CFNI SEXFID</td>
<td>.35*</td>
<td>-.20</td>
<td>-.08</td>
<td>-.05</td>
<td>.11</td>
<td>.30</td>
<td>.05</td>
<td>.20</td>
<td>.06</td>
<td>.45*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CFNI MOD</td>
<td>.02</td>
<td>-.29</td>
<td>-.04</td>
<td>-.04</td>
<td>.05</td>
<td>.20</td>
<td>.06</td>
<td>.45*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CFNI IRR</td>
<td>.40*</td>
<td>-.24</td>
<td>.03</td>
<td>.39*</td>
<td>.13</td>
<td>-.09</td>
<td>.01</td>
<td>.23</td>
<td>-.13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CFNI DOMS</td>
<td>-.05</td>
<td>.39*</td>
<td>.06</td>
<td>-.26</td>
<td>.45**</td>
<td>-.12</td>
<td>.14</td>
<td>-.09</td>
<td>-.14</td>
<td>.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CFNI IIA</td>
<td>-.31</td>
<td>.45**</td>
<td>-.17</td>
<td>-.37*</td>
<td>.06</td>
<td>-.29</td>
<td>.15</td>
<td>-.22</td>
<td>-.43</td>
<td>-.07</td>
<td>.48**</td>
<td></td>
</tr>
</tbody>
</table>

Note: N = 32. CFNI = Conformity to Feminine Norms Inventory. HNR = Have Nice Relationships, IWC = Involvement with Children, THIN = Thinness, SEXFID = Sexual Fidelity, MOD = Modesty, IRR = Investment in Romantic Relationships, DOMS = Domestic, IIA = Investment in Appearance. * p < .05, ** p < .01
Table 2 (continued)
Intercorrelations Between Subscales of CAGES and Validity Instruments: Women on the CMNI

<table>
<thead>
<tr>
<th>Measures</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Discomfort with Nonconformity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Resentful Conformity</td>
<td>- .39**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Active-Physical Comfort Conformity</td>
<td>.26**</td>
<td>- .27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Comfort with Conformity-Appearance</td>
<td>.42**</td>
<td>- .51**</td>
<td>.37**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. CMNI WIN</td>
<td>.32*</td>
<td>- .21</td>
<td>.05</td>
<td>.22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. CMNI EMOC</td>
<td>.02</td>
<td>.13</td>
<td>- .08</td>
<td>- .09*</td>
<td>.11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. CMNI RISK</td>
<td>- .13</td>
<td>.19</td>
<td>.12</td>
<td>- .15</td>
<td>.23</td>
<td>- .24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. CMNI VIOL</td>
<td>.03</td>
<td>.02</td>
<td>.19</td>
<td>- .02</td>
<td>.28</td>
<td>- .27</td>
<td>.30*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. CMNI POW</td>
<td>.38**</td>
<td>- .20</td>
<td>.18</td>
<td>.15</td>
<td>.00</td>
<td>- .09</td>
<td>.08</td>
<td>.07</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. CMNI DOM</td>
<td>.18</td>
<td>- .13</td>
<td>.17</td>
<td>.35*</td>
<td>.40**</td>
<td>- .07</td>
<td>.21</td>
<td>.42**</td>
<td>- .12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. CMNI PLAY</td>
<td>- .06</td>
<td>.09</td>
<td>.19</td>
<td>.11</td>
<td>- .18</td>
<td>- .16</td>
<td>.04</td>
<td>.04</td>
<td>- .10</td>
<td>.09</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. CMNI SELF</td>
<td>.15</td>
<td>.03</td>
<td>-.10</td>
<td>.08</td>
<td>.31*</td>
<td>.47**</td>
<td>- .06</td>
<td>- .08</td>
<td>.02</td>
<td>- .10</td>
<td>- .30*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. CMNI WORK</td>
<td>-.03</td>
<td>.10</td>
<td>.09</td>
<td>-.29</td>
<td>.33</td>
<td>.04</td>
<td>- .23</td>
<td>- .08</td>
<td>- .06</td>
<td>.15</td>
<td>- .29</td>
<td>.17</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14. CMNI HOMO</td>
<td>.53**</td>
<td>-.28</td>
<td>.28</td>
<td>.59**</td>
<td>.15</td>
<td>.05</td>
<td>- .15</td>
<td>.03</td>
<td>.31*</td>
<td>.19</td>
<td>- .09</td>
<td>.02</td>
<td>-.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. CMNI POS</td>
<td>.27</td>
<td>.06</td>
<td>.07</td>
<td>.27</td>
<td>.33*</td>
<td>-.15</td>
<td>.19</td>
<td>.30*</td>
<td>-.08</td>
<td>.53*</td>
<td>.17</td>
<td>- .18</td>
<td>.27</td>
<td>.09</td>
<td></td>
</tr>
</tbody>
</table>

Note. N = 47. CMNI = Conformity to Masculine Norms Inventory, WIN = Winning, EMOC = Emotional Control, RISK = Risk-taking, VIOL = Violence, POW = Power Over Women, DOM = Dominance, PLAY = Playboy, SELF = Self-Reliant, WORK = Primacy of Work, HOMO = Disdain for Homosexuals, POS = Pursuit of Status. *p < .05, **p < .01.
<table>
<thead>
<tr>
<th></th>
<th>Study 1</th>
<th></th>
<th>Study 2</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>Skew</td>
<td>Kurtosis</td>
</tr>
<tr>
<td>DWN</td>
<td>2.91</td>
<td>.46</td>
<td>-.15</td>
<td>.53</td>
</tr>
<tr>
<td>RC</td>
<td>2.22</td>
<td>.69</td>
<td>.19</td>
<td>-.28</td>
</tr>
<tr>
<td>APCC</td>
<td>2.60</td>
<td>.62</td>
<td>.14</td>
<td>-.32</td>
</tr>
<tr>
<td>CCA</td>
<td>2.30</td>
<td>.57</td>
<td>.13</td>
<td>.65</td>
</tr>
</tbody>
</table>

**Note:** DWN = Discomfort with Nonconformity, RC = Resentful Conformity, APCC = Active-Physical comfort Conformity, CCA = Comfort Conformity-Appearance. Study 3a = Test, Study 3b = Retest. Study 1, N = 356, Study 2, N = 176, Study 3a, N = 56, Study 3b, N = 56.
### Table 3.1
*Means, Standard Deviations, Skewness, and Kurtosis Studies 1-3, continued*

<table>
<thead>
<tr>
<th></th>
<th>Study 3a</th>
<th></th>
<th>Study 3b</th>
<th></th>
<th></th>
<th></th>
<th>Study 3b</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>Skew</td>
<td>Kurtosis</td>
<td>M</td>
<td>SD</td>
<td>Skew</td>
</tr>
<tr>
<td>DWN</td>
<td>2.61</td>
<td>.40</td>
<td>-.06</td>
<td>.31</td>
<td>2.58</td>
<td>.41</td>
<td>-.71</td>
</tr>
<tr>
<td>RC</td>
<td>2.70</td>
<td>.50</td>
<td>-.10</td>
<td>.58</td>
<td>2.79</td>
<td>.51</td>
<td>.24</td>
</tr>
<tr>
<td>APCC</td>
<td>2.16</td>
<td>.54</td>
<td>.25</td>
<td>.60</td>
<td>2.15</td>
<td>.59</td>
<td>-.17</td>
</tr>
<tr>
<td>CCA</td>
<td>1.87</td>
<td>.39</td>
<td>-.69</td>
<td>-.28</td>
<td>1.91</td>
<td>.33</td>
<td>-.67</td>
</tr>
</tbody>
</table>

*Note:* DWN = Discomfort with Nonconformity, RC = Resentful Conformity, APCC = Active-Physical comfort Conformity, CCA = Comfort Conformity- Appearance. Study 3a = Test, Study 3b = Retest. Study 1, N = 356, Study 2, N = 176, Study 3a, N = 56, Study 3b, N = 56.
Figure 1. Continua of Comfort and Conformity of Gender Expression
Figure 2. Frequencies of Responses for CAGES subscale: Discomfort with Nonconformity
Figure 3. Frequencies of Responses for CAGES subscale: Resentful Conformity
Figure 4. Frequencies of Responses for CAGES subscale: Active-Physical Comfort Conformity
Figure 5. Frequencies of Responses for CAGES subscale: Comfort Conformity-Appearance
Katherine Grace Spencer was born in Iowa City, Iowa on September 17, 1977. She graduated from the University of Wisconsin-Madison in 1999 with a Bachelor’s Degree in Women’s Studies and Psychology. She earned a Master’s Degree in Counseling Psychology from the University of Missouri-Columbia in 2003. She completed her pre-doctoral internship at the University of Illinois-Chicago Counseling Center in August 2007, and began a two-year post-doctoral fellowship at the Program in Human Sexuality in the Department of Family Medicine and Community Health at the University of Minnesota Medical School in October 2007. She plans to pursue a career in clinical practice and advocacy, focusing on privilege/oppression and the intersections of race, gender, sexual orientation, and class.