STEED, TENEKA C., Ph.D. Superwoman Schema: Using Structural Equation Modeling To Investigate Measurement Invariance In A Questionnaire (2013).
Directed by Dr. John T. Willse. 125 pp.

Evaluating the psychometric properties of a newly developed instrument is critical to understanding how well an instrument measures what it intends to measure, and ensuring proposed use and interpretation of questionnaire scores are valid. The current study uses Structural Equation Modeling (SEM) techniques to examine the factorial structure and invariance properties of a newly developed construct called Superwoman Schema (SWS). The SWS instrument describes the characteristics of a superwoman (strong woman) which consists of 35 items representing five subscales: obligation to present an image of strength, obligation to suppress emotions, resistance to being vulnerable, intense motivation to succeed, and obligation to help others. Multigroup confirmatory factor analysis (CFA) and a multiple indicators multiple causes (MIMIC) model were the SEM approaches used to examine measurement invariance in the SWS instrument. Specifically in the multigroup CFA analyses, configural invariance, metric invariance, intercept invariance, residual variance invariance, and latent mean invariance are examined between a group of young (18-39 years old) women and middle-aged (40-65 years old) women. In the MIMIC model, the hypothesized model of the SWS was used to investigate the group differences in the young and middle-aged women. Both SEM techniques provided a didactic discussion about the findings of the study, which confirmed that the SWS instrument could be broadly used (i.e., invariance held) to compare young and middle-aged African American women on superwoman characteristics. Further research is needed to better understand the possible contextual

factors (i.e., racial or gender stereotyping, oppression, spiritual values, etc.) that may contribute to group differences on the SWS subscales and minor violation to invariance.

SUPERWOMAN SCHEMA: USING STRUCTURAL EQUATION MODELING TO INVESTIGATE MEASUREMENT INVARIANCE IN A QUESTIONNAIRE

by

Teneka C. Steed

A Dissertation Submitted to
the Faculty of The Graduate School at
The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Greensboro
2013

Approved by

_____
Committee Chair

To my circle

APPROVAL PAGE

This dissertation has been approved by the following committee of the Faculty of The Graduate School at The University of North Carolina at Greensboro.

Committee Chair _____

Committee Members _____

_____

_____

_____
Date of Acceptance by Committee

_____
Date of Final Oral Examination

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my committee for their contributions and guidance in getting this dissertation done!

To Dr. John Willse, huge thank you for all of your support, dedication, and motivation for getting me to the finish line! Words cannot really express how grateful I am to have had you as my chairperson and advisor over the last several years. Our countless meetings and phone conversations provided me with what I needed to keep moving forward and I am forever thankful for your contributions.

Thanks to Dr. Micheline Chalhoub-Deville for your mentoring, guidance, and thorough feedback on making my dissertation more polished. I really appreciate your thoughtfulness and words of encouragement throughout this process. Thank you for giving me advice on my future career and believing in me.

Thanks to Dr. Ric Luecht for all of your support and guidance over the years. You have been one of the best professors I have ever had. I appreciate your willingness and drive to teach with such a big heart and enthusiasm in Educational Research Methodology (ERM). I hope one day that I can be like you…big shoes to fill!

And to Dr. Cheryl Giscombé, you have impacted my life for quite some time now and you were the first to introduce me to applied research! Thank you for allowing me to use the Superwoman Schema (SWS) for this project. I look forward to working on future studies with you and making a major impact in the field on SWS.

Last but not least, I would like to thank all of my ERM fellow colleagues, friends, and family who made a major impact on getting this project done. We did it!

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER I

INTRODUCTION

The Superwoman Schema (SWS) framework was developed to better understand the relationship between stress and health in women (Woods-Giscombé, 2010). The SWS framework proposes to measure the superwoman role as it relates to how women present an image of strength, suppress their emotions, resist being vulnerable to others, take on multiple roles and responsibilities while neglecting their own self-care, and despite all of these characteristics of a superwoman, they still have an intense motivation to succeed. This study provides the foundational work of evaluating the factorial structure and measurement invariance of the SWS framework. This chapter provides an overview of this study including statement of problem, purpose of study, and research questions.

Statement of the Problem

Researchers are requesting to use the SWS without unknown properties of reliability or validity. The instrument should not be released into the marketplace with intentions of measuring superwoman characteristics in women when it's not known if the SWS instrument is measuring what it intends to measure based on the SWS framework. Therefore, there is a need for further research on the operational use of the SWS instrument. Helping evaluating the psychometric properties of the SWS instrument allows for more effective research to be conducted on stress and health in women across the country. Furthermore, evaluating how groups are similar in endorsing questions on

the SWS instrument adds to the ability to compare groups using the SWS instrument. Once the psychometric properties of the instrument is assessed, the SWS instrument can be used with confidence in knowing that the instrument measures what it intends to measure and that distinctive groups can be compared using this instrument based on empirical research. This study uses empirical data from the SWS instrument to examine the group differences among women using Structural Equation Modeling (SEM) approaches, which are discussed later.

There have been a couple instruments developed to measure the superwoman phenomenon; however, few studies appear to have considered if the superwoman phenomenon is equivalently valid across different subpopulations, and no published studies have used two SEM approaches (multigroup Confirmatory Factor Analysis and multiple indicators and multiple causes model) to measure the characteristics of a superwoman. Thompson (2003) and Hamin (2008) have both evaluated a similar instrument related to SWS called the Strong Black Woman Scale (SBWS) and the Strong Black Women Cultural Construct Scale (SBWCCS), respectively. The SBWS was revised by Hamin (2008) and renamed the SBWCCS. The latest version of this instrument comprises 22 items defining three factors: self-reliance – belief of independence and control; affect regulation – control over emotions (sadness or fear is a sign of weakness) and suppressing emotional needs; and caretaking – taking care of others and neglecting own needs. Self-reliance, affect regulation and caretaking are all interrelated characteristics of the SBWCCS. Cronbach's α for the subscales of SBW were 0.62 for self-reliance; 0.69 for affect regulation; and 0.75 for caretaking (Hamin, 2008).

Thompson (2003) conducted a series of analyses to examine the construct on the SBWS which primarily used one group (i.e. African American women). A focus group study was used in Thompson's study to establish content validity, and confirmatory factor analysis was used to examine the psychometric properties of the scale.

There are limited research studies (Rivera & Satorra, 2002; Reynolds, Keith, Ridley, & Patel, 2008; Rosén, 1995; Thompson & Green, 2006) that compare multigroup CFA and MIMIC model in understanding if the same measurement models are applicable across different groups, and whether the relationship among latent variables and observed variables are the same for each group. These two SEM approaches in examining group differences supplement each other in providing alternatives to researchers in studying measurement invariance in heterogeneous populations (Reynolds et al., 2008). This study intends to add to the existing literature on the superwoman and measurement invariance by comparing two SEM approaches to examine group differences using the SWS instrument between young and middle aged women.

Overview of Measurement Invariance

Establishing measurement invariance in an instrument is one aspect of validity, which is the evaluation of the usefulness and appropriateness of a test for a particular purpose. The evaluation of validity is not a one-time event; it's an on-going process (Sireci, 2007). Traditionally, notions of validity were composed of several aspects such as evidence related to criteria, content, and construct validities (American Educational Research Association, American Psychological Association, & National Council on

Measurement in Education, 1990). One of the ways validity can be a threat is if there is a lack of measurement invariance in assessments and questionnaires (Messick, 1989).

Measurement invariance is established when an instrument is shown to be operating in the same way across groups. If measurement invariance cannot be established, then findings between groups cannot be interpreted with much confidence. The more psychometrically sound (e.g. reliable and valid) an instrument is, the more confident a practitioner or researcher can be in making decisions based on results from the instrument. In order for practitioners to make accurate comparisons among heterogeneous populations, the instrument being used should provide evidence of measurement invariance. Because the instrument being investigated in this study is a newly developed framework, one purpose of this study is to begin the process of evaluating aspects of validity and measurement invariance with the SWS instrument to ensure that the instrument is can be used as a research tool to better understanding the superwoman role.

SEM techniques are used to examine the factor structure, and how the structure compares across groups using the SWS theoretical framework. The SEM techniques used include multigroup confirmatory factor analysis (multigroup CFA), and multiple indicators and multiple causes (MIMIC) model. The multigroup CFA analysis will test a priori hypotheses about the structure of the SWS and its invariant functioning across different user groups. The SWS hypothesized model includes the following five subscales: (1) obligation to present an image of strength; (2) obligation to suppress emotions; (3) obligation to helping others; (4) resistance to being vulnerable; and (5)

intense motivation to succeed. More specifically, the CFA examines the relationship between the observed variables (i.e. the items of the SWS instrument) and the latent variables (i.e. the SWS five subscales). The CFA analysis provides evidence of the multidimensional SWS instrument to support that it measures what it intends to measure.

The SWS instrument was initially developed using three age groups to measure superwoman characteristics in African American women: (1) 18-25 years old, (2) 26-45 years old, and (3) 45 years old and older. This study examines the superwoman characteristics between two age groups according to Erick Erikson's stages of development (Schickendanz, Schickendanz, Forsyth, & Forsyth, 2001): 1) young women ages 18-39, and 2) middle aged women ages 40-65. By examining both age groups of women, multiple group comparisons using the SWS instrument are made. The multiple group comparison tests the invariance of construct measurement between young and middle aged women. If invariance holds between the two groups of women, then the SWS instrument assesses similar characteristics of the superwoman concept between both young and middle-aged women. To assess this cross-group comparison, a multigroup CFA model examines the measurement invariance in the SWS instrument.

Multigroup CFA is a commonly known method used to investigate measurement invariance (Byrne, 1998; Cheung, 2007; Vandenberg, 2011). Multigroup CFA examines measurement invariance using a series of increasingly restrictive tests in a SEM framework. Specifically, equality constraints are imposed on the hypothesized measurement models of two or more groups (e.g. young adulthood and middle adulthood women). The equality constraints ensure the aspect of the model is functioning in the

same way for each group. Next, model fit is examined. By evaluating model fit, the level of measurement invariance can be determined. This is, if the model function well with the imposed equality constraints, the assumption of the level of invariance is supported.

There are two main concepts associated with invariance: measurement invariance and structural invariance (Byrne, 2008; Cheung & Rensvold, 2002). Cheung and Rensvold (2002) and Little (1997) have described these two major components of invariance as category 1 invariance (measurement level) and category 2 invariance (structural level). Category 1 invariance comprises the psychometric properties of the measurement scale that includes configural invariance, construct-level metric invariance, item-level metric invariance, residual variance invariance, and intercepts invariance. Category 2 invariance is concerned with the equality of relations among the factors that includes tests of construct variance invariance, construct covariance invariance, and latent mean invariance. In this study, the two major components of measurement invariance are used to describe the various levels of invariance tests: measurement level invariance and structural level invariance.

The first component of invariance, measurement level invariance, addresses the issue of an instrument being equivalent across groups. Specifically, measurement invariance focuses on the invariant operation of items on an instrument (e.g. factor loadings) when researchers are most concerned with the extent that the content of each item is equivalent across groups (Byrne, 2008). Measurement invariance ensures that the content of the instrument and/or items are perceived and are interpreted the same across different groups. The observed differences between the groups should only reflect true

differences in the variability of the construct. For example, if the SWS instrument is invariant between young women and middle aged women, then the assumption holds that the SWS instrument is measuring the same trait, in the same way, in both groups of women. Furthermore, if the SWS instrument is invariant for both groups of women, then the comparisons and analyses of scores would be acceptable and yield meaningful interpretations.

Structural level invariance is most concerned with the equivalence of relations among factors (i.e. factor covariance) (Byrne, 2008). In particular, structural level invariance can answer the following questions (Byrne, 2008): (1) does the dimensionality of the construct holds across groups and, (2) does an instrument developed by a theoretical framework produce equivalent hypothesized dimensions across groups? Byrne (2008) argued that structural level invariance should not be tested if there is no evidence that the measurement level invariance parameters are operating in the same way across groups. "Testing for equivalence entails a hierarchical set of steps that typically begin with the determination of a well-fitting multigroup baseline model for which sets of parameters are put to the test of equality in a logically ordered and increasingly restrictive fashion" (Byrne, 2008, p. 872). Measurement level invariance consists of five hierarchical invariance tests. The hierarchical nature of the tests implies that there is no utility in testing for higher level invariance unless the lower level invariance tests have been acceptable. Once all required measurement level invariance tests are examined and found to support the assumption of invariance, then structural level invariance tests can themselves be examined. The SEM literature is not consistent in confirming if

researchers should or should not test for structural level invariance (Byrne, 2010). However, Byrne (2010) argues that it is important for construct validity researchers to test the structural level invariance if they are interested in testing whether a dimensional construct holds across groups. Another method used to investigate measurement invariance (i.e. item or test bias) is called multiple causes multiple indicators (MIMIC) model (Jöreskog, & Goldberger, 1975).

MIMIC models are used to estimate group differences on latent variables. MIMIC models also are used for testing items on psychological assessments and/or questionnaires to determine if they measure the same underlying construct and possess the same measurement properties for all groups (Woods, 2009). With MIMIC models, latent variables with effect indicators are regressed on one or more dichotomous cause indicator that represents group membership (i.e. gender, ethnicity, socioeconomic status, etc.). The cause indicator, the grouping variable, has direct effects on latent variables. In MIMIC models, group differences are interpreted by examining the significance and magnitude of the structural regression variables. Group differences can still be detected in MIMIC models even if measuring instruments are not invariant (though some model adjustments may be needed). In other words, MIMIC model analysis does not require that measurement invariance be established before testing group differences. In fact, MIMIC models may help identify breakdowns in invariance.

Purpose of Study

This current research intends to make two distinct contributions. First, the analysis provides evidence for the appropriateness of using the SWS with different age

groups (i.e., conducts an invariance study). Second, the research serves as didactic explanation of the use of invariance testing (both multigroup CFA and MIMIC) in an applied setting. To accomplish these goals, several steps are taken. These steps are organized into explicit research hypotheses.

Conducting SEM analyses provide partial validation of the SWS instrument for researchers and practitioners to use. A validated SWS instrument adds supplemental information for helping better understand the needs of women and ways to provide effective interventions and treatments of health related issues, particularly for African American women (Woods-Giscombé, 2010). The multiple group study of the SWS will add to the literature and provide a different perspective of the superwoman phenomenon across age in women. Also, there are few applied studies that have explicitly described and compared multigroup CFA and MIMIC models. This current study addresses that shortcoming in the literature.

Research Questions

The problem described above leads to 3 multifaceted research questions which includes:

1. Does the five-factor structure proposed by the Superwoman Schema instrument adequately describe survey responses from women in the intended populations?
2. Is the superwoman schema instrument invariant across both young and middle-aged women in the intended population?
    a. Configural invariance: Do the groups have the same factor structure?
    b. Metric invariance: Do the groups have the same factor loadings?

c. Intercept invariance: Do the groups have the same item intercepts?
d. Residual variance invariance: Do the groups have the same item residual variances?
e. Latent mean invariance: Are the latent means invariant across groups?

3. Does using a MIMIC model lead to conclusions that are similar to those conclusions reached using a multigroup CFA?

CHAPTER II

REVIEW OF LITERATURE

This chapter introduces the review of the Superwoman Schema (SWS) instrument, the concept of superwoman, age differences, validity, item response theory (IRT), structural equation modeling (SEM), confirmatory factor analysis (CFA), multigroup CFA, multiple indicators multiple causes (MIMIC) model, and model fit. The discussion of SEM is descriptive for a lay audience only somewhat familiar with the theory and practice of the methodology. Further discussion involves the specific process of conducting a multigroup CFA and MIMIC model analysis. The discussion of methods is discussed in the context of how analyses assist in aspects of the SWS instrument.

Superwoman Schema Instrument

The SWS instrument was developed from Woods-Giscombé's (2010) pilot research study, which investigated the superwoman phenomenon among African American women. From this focus group pilot research study, a conceptual SWS framework was developed based on the data collected from the women in the study. The SWS conceptual framework suggests that sociohistorical factors (i.e., racial and gender stereotyping or oppression) may result in emotional suppression, determination to achieve goals despite limited resources, and limited prioritization of self-care. In the focus group study, Woods-Giscombé (2010) collected data to develop this framework by investigating how African American women characterized the superwoman role, what

women believed to be the contributing contextual factors of the superwoman role, and what women described as the benefits and liabilities of the superwoman role in relation to their general well-being. This focus group study was designed to develop a conceptual framework of superwoman to operationalize the superwoman role, and to develop an instrument to measure the characteristics of a superwoman to facilitate an empirical examination of its impact on the health of African American women (Woods-Giscombé, 2010).

Woods-Giscombé's (2010) focus group study included 48 African American women from diverse age and educational backgrounds. A total of eight focus group sessions were conducted between December 2006 and June 2007. The eight focus groups were held during eight different sessions based on age (ranging from age 19 to 72) and educational background (ranging from individuals without high school diplomas to those with terminal degrees such as J.D., Ph.D., etc.). The focus group participants represented a community-based sample located in a large metropolitan area in the southeastern region of the United States. (Woods-Giscombé, 2010). The focus groups sessions were held in private rooms located in community locations such as public libraries and colleges. The sessions were approximately two hours for each group. Each participant was compensated $30 for their time and they were also provided a meal.

In each focus group session, the moderator provided a brief summary of the study, administered consent forms to each participant, and conducted an icebreaker activity. After the icebreaker activity, the moderator began the study by asking key questions

related to stress, coping, and the superwoman role. Focus group discussions included the following questions (Woods-Giscombé, 2010):

1. When I say the word stress, what does it mean for you?
2. What causes stress in your life?
3. How do you cope with stress?
4. How did you see the women (mothers, grandmothers) in your life cope with stress?
5. Have you ever heard the term Strong Black Woman/Black Superwoman?
6. What is a Strong Black Woman/Black Superwoman?
7. What are her characteristics?
8. How did they develop?
9. Is being a Strong Black Woman/ Black Superwoman a good thing?
10. Is there anything bad about being a Strong Black Woman/Black Super- woman?

In addition to the key questions asked above, the participants completed a brief demographic questionnaire to obtain background information such as age, occupation, and household income. A research assistant co-facilitated the focus groups and recorded field notes. Each focus group session was audio-recorded and transcribed.

After each focus group session was transcribed, an analytic induction was used to analyze the data (Woods-Giscombe, 2010). An analytic induction is also referred to as deviant case analysis, which involves a prescribed process for systematic analysis of the data (Frankland and Bloor, 1999). Key words and thoughts were grouped together to

form index-coded categories during the analytic induction analysis. Systematic comparisons of the index-coded categories were conducted to identify the most relevant data to the topic or index code (Woods-Giscombé, 2010). This process was cyclical which means that as more data was collected and transcribed, new index-coded categories were identified and subcategories were created (Woods-Giscombé, 2010). Data that did not fit into the index-coded categories were not discarded but used to further contextualize the data (Woods-Giscombé, 2010). Concepts and items were identified for the preliminary development of the SWS instrument based on the identified index-coded categories and subcategories (Woods-Giscombé, 2010). Next, the focus group participants received a summary of the results from the focus group study via postal mail after all eight focus group sessions were transcribed and analyzed. Once the participants received a summary of the focus groups results, they were invited to communicate feedback to the research team through written or verbal correspondence.

The results of the focus group study demonstrated that the superwoman framework is a multidimensional phenomenon encompassing characteristics such as an obligation to present an image of strength, an obligation to emotional suppression, a resistance of being vulnerable, an intense motivation to succeed, and an obligation to help others (Woods-Giscombé, 2010). In addition, the superwoman framework was identified as having contributing contextual factors (e.g., historical events, spiritual values, etc.), and perceived benefits and liabilities (e.g., self-survival, stress, etc.). These findings contributed to the preliminary development of the SWS instrument (Woods-Giscombé, 2010). Contributing contextual factors identified included historical legacy of racial or

gender stereotyping or oppression; lessons from foremothers; past history of disappointment, mistreatment, or abuse; and spiritual values. The perceived benefits of the superwoman role identified were preservation of self and/or survival; preservation of the African American community; and preservation of the African American family. The perceived liabilities of the superwoman role identified included strain in interpersonal (e.g., romantic) relationships; stress-related health behaviors (e.g., postponement of self-care, emotional eating, poor sleep); and embodiment of stress (e.g., anxiety, depressive symptoms, adverse maternal health.

Once the focus group research study was completed, phase two of the focus group study was conducted to begin operationalizing the SWS framework. Phase 2 involved two additional focus groups which included 21 African American women from phase 1 study. These women examined the item clarity, readability, and content validity of the preliminary SWS scale. There were a total of 144 preliminary items developed from the original focus group study (phase one study). From the results of the focus group analysis (phase two study), the number of preliminary items reduced from 144 items to 60 items. Additional item analyses and content validity analysis were conducted during phase two, which reduced the item count to 35. The current study (phase three) will evaluate the psychometric properties of the 35 items of the preliminary SWS instrument with five subscales (see Appendix A).

SWS Subscale One. An obligation to present an image of strength is described as a woman having to present to others that she is "strong" (i.e. I feel obligated to present an

image of strength for my family) among family, friends, coworkers, etc. Four items are included in this subscale.

SWS Subscale Two. An obligation to suppress emotions is described as women hiding and not sharing their emotions with others (i.e. my tears are a sign of weakness). The obligation to suppress emotions subscale consists of six items.

SWS Subscale Three. The third subscale is called resistance to being vulnerable. This dimension is described as women denying or seeking help when needed, because they do not want others to perceive them as being vulnerable (e.g., I try to do everything by myself). This subscale includes six items.

SWS Subscale Four. An intense motivation to succeed is described as a women seeking success regardless of limited resources available to them and working relentlessly to achieve goals (i.e. routinely working late, skipping meals, and sacrificing sleep). This subscale includes four items.

SWS Subscale Five. An obligation to help others is described as a women fulfilling various duties to help others and/or being involved in numerous activities outside of work and/or school (i.e. participation in organizations and groups, helping family and friends in need). Women with these characteristics tend to prioritize the needs of others over their own personal needs. There are eight items included in this subscale.

Concept of Superwoman

The concepts of superwoman which derives from the characteristics of a strong Black woman are not new (Mullings, 2006; Romero, 2000; Wallace, 1990); however, there is limited empirical literature that exists on these topics. As recently as ten years

ago, most of the literature published on superwoman and strong Black woman came from non-fiction literature (Woods-Giscombé, 2010). Although more current empirical work has begun to focus on this concept, the research tends to be descriptive. According to Beauboeuf-Lafontant (2007), Black feminists identified that being strong was a specific culturally expectation placed on African American women. These women had to uphold the standards of being strong in their culture. Because African American women were expected to be strong, many of them became silent about their expectations of being strong, Black women. Beauboeuf-Lafontant (2007) identified this silence among African American women as the silencing paradigm. The silencing paradigm is described as "…normative expectations for women insist that they be overly attuned to others' needs, often at great cost to their own goals, desires, and feelings" (Beauboeuf-Lafontant (2007). Women silenced their true feelings and concerns because they felt that family and friends would not accept their discourse-discrepant feelings and thoughts about being a strong woman (Beauboeuf-Lafontant, 2007). Because women silence their thoughts and feelings, they internalize how they truly feel about being a good woman defined by others in the community. The silencing paradigm may lead to depression and other-related mental disorders in women.

This concept of being strong is sometimes considered honorable among African American women despite the negative stereotypes placed among African American women in society. Strength is considered as a moral characteristic, independence, and/or the capacity to complete a goal (Beauboeuf-Lafontant, 2007). According to Beauboeuf-Lafontant (2007), the discussion of strength developed its authority from contrasting

African American women to middle-class Caucasian women. In addition, strength is rooted in several conflicting assumptions: African American women may feel they must be strong as opposed to having freedom to display vulnerabilities due to racial and stereotypical oppression; strength is a natural quality apart of the African American womanhood; and that being strong characterizes all African American women's behaviors and attitudes. Furthermore, the term strong Black woman is typically used in the African American community to describe a woman who appears to be strong, resilient, tough, and self-sufficient which are the same characteristics of being a superwoman (Thomas et al., 2004).

The concept of superwoman developed during the feminine mystique era of the 1950s (Jacques, 2008). The meaning of a superwoman during the feminine mystique era helped define the role of a superwoman today. In the era, a superwoman was defined and known as having her life all under control – her life is prefect, she can cope with anything, nothing stresses her out, she is great in all roles, she fulfills multiple roles (housewife, mother, daughter, etc.) and she is a strong, independent career woman (Crago, Yates, Fleischer, Segerstorm, & Gray, 1996; Jacques, 2008; Herrera & DelCampo, 1995). The superwoman term has become an increasingly present topic in the media particularly among African American women because they tend to identify with the characteristics of being a superwoman (Black & Peacock, 2011). Many African American women describe the superwoman role as a survival mechanism in their world which is like a weapon to withstand discrimination of race and gender. (Woods-Giscombé, 2010). The multiple roles and responsibilities are standards (gender

stereotypes) that are set upon women by society and their culture which makes it extremely challenging for women (Beauboeuf-Lafontant, 2007). African American women simultaneously deal with race and gender issues which makes it even challenging for this group of women display characteristics of a superwoman (Settles, Pratt-Hyatt & Buchanan, 2008; Thomas, Witherspoon, & Speight, 2004). However, these women try to live up to society standards by being a superwoman or strong Black woman.

The obligation to display strength has been associated with distress in women which includes anxiety, anger, or depression particularly among African American women (Beauboeuf-Lafontant, 2007). The characteristics of the superwoman role cause women to internalize emotions which prevents them from expressing their needs and wants. This internalization may lead to psychological distress and health issues (Jacques, 2008; Hart & Kenny, 1997). Women who are expressing superwoman attributes may have psychological distress and health-related problems because they feel obligated to portray an ability to do everything on their own without support from others. Therefore, women with superwoman characteristics tend not to seek for any type of help from anyone because they feel obligated to present the image of being a superwoman or strong woman (Thomas, Witherspoon, & Speight, 2004).

The superwoman concept is considered a threat to women's emotional, mental, and physical health, predominantly in African American women (Romero, 2000; & Thompson, 2000; Thomas, Witherspoon, & Speight, 2004). Superwoman characteristics may play a role in the disparate health conditions experienced by African American women, including cardiovascular disease, obesity, and diabetes. Superwoman

characteristics may also influence misdiagnosed and undertreated psychological conditions. Previous research (Hamin, 2008) has argued that many African American women have been misdiagnosed due to health care professionals not fully understanding the issues and needs of African American women. Most African Americans do not seek professional help like other racial groups because of the stigma related to seeing a psychiatrist or counselor to help with personal and/or family issues (Utsey, Giesbrecht, Hook & Stanard, 2008). Seeking professional help (e.g., counseling, psychiatric care) is often frowned upon in the African American community because coping mechanisms such as religion and spirituality, and strong bonds among family and friends are most often preferred to help with stress-related issues (Utsey et al., 08). Hamilton-Mason, Hall, & Everett (2009) stressed the importance for professionals (i.e., educators, practitioners, and researchers) to understand and incorporate the conceptualizations of multiple theories (i.e., superwoman role, strong Black woman, racial and stereotypical issues, etc.) into their professional work when working with African American women without marginalizing them. In addition, Hamilton et al. (2009) reemphasized that understanding the impact of race, gender, social class, stress, and coping across the life span offers another perspective about the psychological well-being and mental health needs of African American women.

Age Differences

Examining the differences in age groups using the SWS instrument adds to the existing literature related to psychosocial development stages of life. The relationship between age and stress varies across different life spans. According to previous literature

(Carstensen, Isaacowitz, & Charles, 1999; Charles, Reynolds, & Gatz, 2001; Kanfer & Ackerman, 2004; Phillips, Henry, Hosie, & Milne, 2006; Rauschenbach & Hertel, 2011), older adults have decreased life stressors (e.g., relationship issues, unemployment, health issues, etc.) when compared to younger adults. It has been reported that older adults handle stressors in life differently from younger and middle-aged adults. In Charles et al. (2001) study, they reported that it is possible that older adults may become more skillful, insightful, and/or flexible in coping with life stressors. The experiences and coping techniques help older adults better manage stress which reduces the stress levels of daily stressors.

Another possibility of older adults demonstrating less stress levels may be caused by reduced exposure to daily stressors, which is explained by the life-span theory of motivation called, socioemotional selectivity theory (Carstensen, 1995; Charles et al., 2001; Phillips et al., 2006). The socioemotional selectivity theory argues that as people get older, they are more selective in how they spend their remaining time living particularly spending more quality time with close family members (Carstensen,1995; Charles & Carstensen, 2007; Freund & Baltes, 2002). In particular, older adults value their limited time by investing in more emotionally meaningful goals and activities as compared to younger generations. There are two goal-related stages in the socioemotional selectivity theory called emotion-related goals and knowledge-related goals. The emotion-related goals stage focuses on emotional regulation, emotional gratifying interactions with social partners, and other activities that can benefit the individual in the present moment. Older adults are considered to be in the emotion-related goal stage

because they typically tend to spend more quality time with social partners and family members which may reduce their exposure to daily stressors (e.g., relationship issues). In the knowledge-related goals stage, individuals are aimed at gaining knowledge, career planning, developing new social relationships and seeking other endeavors that may benefit them in the future. Younger adults are considered to be in the knowledge-related goals stage of the socioemotional selectivity theory. Adults in the knowledge-related goals stage typically are not concerned about their time left to live because they think they have more time to live compared to older adults. Based on these findings, one can conclude that younger adults may have higher levels of stress considering that they are more concerned with developing a career and building social relationships for their future. Conversely, older adults are in the process of retiring, enjoying and maximizing their time with love ones which is not as much stress of building a career and social networks like younger adults.

Based on the life span developmental theories, it is hypothesized that younger and older women in this study will differ on the superwoman characteristics. It is expected that younger women will have more stress levels compared to the older woman. Almedia and Horn (2004) revealed that younger adults tend to report greater numbers of daily life stressors than older adults. To conclude, age-related decreases in daily stressors are apparently observed when comparing younger, middle-aged, older adults (ages ranging from 25 to 74 years old) (Almeida & Horn, 2004). In spite of Almedia and Horn's findings among young, middle, and older adults, it is expected that young adult women will exhibit more qualities of the superwoman role in the SWS instrument because of the

difference in psychosocial stages and age differences between the young and middle aged women in this study. In the current study, the groups are based on the life-span developmental theory which includes young women include ages 18 to 39, and middle-aged women include ages 40 to 65 (Santrock, 1995).

The examination of the SWS instrument with the young and middle-aged women groups is appropriate in helping validate the use of the instrument across the age span of the study's population (18-65 years). Multiple group confirmatory factor analysis (multigroup CFA) is a robust method for examining the appropriateness of the SWS instrument with the two age groups in this study. Multigroup CFA is discussed later in detail along with a brief description of CFA.

Validity

There is a continuous debate about the use and understanding of validity since Messick's definition of validity as "an integrated evaluative judgment of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of inferences and actions based on test scores or other modes of assessment" (Messick, 1989). Since Messick (1989) and others (e.g., Kane, 1992, 2006, 2009; Lissitz & Samuelsen, 2007), modern validity theorists have argued that an instrument is not tested for validity. Instead, the use and interpretations of scores of an instrument are tested for validity. The debate about validity is particularly focused on the misunderstanding of the term *validity*. Frisbie (2005) discussed that the continued misuse and misunderstanding of validity could lead to negative consequences (i.e., weak

validation studies, miscommunication inside and outside the measurement field, misinterpretation of scores, etc.).

The concept of the modern validity theory developed its roots from construct validity. There are three aspects of construct validity that emerged as the basic principles of modern validity theory (Kane, 2012): 1) validation of a proposed interpretation or use of scores, 2) empirical evaluation of various implications of defining theory, and 3) challenge of proposed interpretations and consider alternate interpretations. These three aspects of validity focuses on the importance of shifting from discussing the validity of a test to discussing the validation of proposed use and interpretation of test scores. Also, Kane (2012) discussed how modern validity theory is an on-going process as compared to the earlier frameworks of validation studies where single empirical validation studies were conducted. By the mid-1980s, the three aspects of validity led to the development of the concept of modern validity theory. The focus on use and interpretations of test scores of modern validity theory were developed and refined by Cronbach (1971), Kane (1992, 2006), Messick (1989), among others. Among the redefined meaning and use of validity, an argument-based approach to validation was developed. There are two steps of the argument-based approach (Kane, 2006, 2012): (1) interpretive argument which specifies the proposed interpretations and uses of the test scores, and (2) validation argument which evaluates the overall plausibility of the proposed interpretations and uses of test scores. Any interpretation or use of test scores can be proposed; however, evidence must be provided to support the proposed interpretation and use of test scores.

Messick's definition of validity is consistent with the Standards of Educational and Psychological Testing (AERA, APA, & NCME, 1999) in which both meanings of validity focus on the evidence and theory that supports the interpretations and uses of tests scores. Sireci and Parker (2006) stated that evidence is collected to develop a scientifically sound validity argument that supports the intended interpretation of test scores and their relevance to the proposed use. The evidence includes gathering and analyzing data that is important to the degree to which test scores fulfill their intended purpose (Sireci & Parker, 2006). Furthermore, theory helps guide the development of a construct's meaning and the construct's interpretation and use of test scores. Sireci and Parker stated that theory involves answering the following questions: (1) what is the underlying meaning of the test? and 2) what is the construct being measured? In the current research study, the theory of the SWS framework operationally defines and measures the characteristics of a superwoman represented by five subscales: an obligation to present an image of strength, a resistance to being vulnerable, an intense motivation to succeed, an obligation to help others, and an obligation to suppress emotions.

According to the Standards, there are five sources of validity evidence that are used to evaluate the proposed interpretation of test scores for a particular use. The five sources of validity include (1) test content; (2) response processes; (3) internal structure; (4) relations to other variables; and (5) consequences of testing. In earlier concepts of validity, the first source of validity evidence (test content) was called content validity. This first source of validity includes all of the aspects of content validity such as item

writing, statistical reviews, etc. The second source of validity evidence (response processes) is based on examining the fit between the construct measure and the participants' response to test items (Kane, 2006). This source of validity evidence focuses on an individual's thinking process. For example, a way to identify response processes is to ask test takers while taking an assessment what are they doing or how are they answering a particular item on a test. The third source of validity evidence (internal structure) is most concerned with what is going on underneath the surface of observed responses. Sub-score data is a common way to understand internal (test) structure. Factor analysis and multidimensional scaling are used to investigate the dimensions measured by an assessment, which provides evidence for internal structure (Kane, 2006). The fourth source of validity evidence is based on relations to other variables, which are commonly investigated using correlations. There is also convergent and discriminant validity evidence that could be used to provide evidence for relations to other variables. Convergent validity is when constructs are related to each other and evidence shows that the constructs are strongly related. Discriminant validity is when constructs claim to measure different things and evidence shows that these constructs are not highly related (low correlations). In this current study, convergent validity evidence can be provided from factor analysis where items load on their intended scale of the SWS, and discriminant validity evidence can be provided from factor correlations where correlations may be low. Lastly, the fifth source of validity evidence is based on consequences of testing also known as consequential validity. This source of evidence evaluates the intended and unintended consequences associated with an assessment

(Kane, 2006).). For this study, the internal structure is used as a source of validity evidence to evaluate the SWS interpretation of test scores using the intended population. Also, another source of validity evidence based on relations to other variables is examined in this study using the multigroup CFA approach. In the multigroup CFA, the relationships between the factors are examined.

In validation studies, the proposed interpretation and use of tests is important which is the one of the purposes of examining the SWS instrument. The use of structural equation modeling techniques is used to initially validate the use and interpretation of scores of the SWS. Evidence is collected to support the intended purpose of the SWS instrument. This evidence includes examining the response processes (e.g., evidence concerning the fit between the construct and the examinees' responses or performances) and internal structure (e.g., evidence includes statistical analysis of item and sub-score data with multigroup confirmatory factor analysis) (Sireci & Parker, 2006). As such, this study plays a critical role in the process of validating the SWS for use as research tool. To assist in collecting evidence to validate the use and interpretation of the SWS instrument, SEM techniques are used to examine invariance across groups in this study. Before discussing the SEM techniques, another approach is briefly discussed next that also measures invariance.

Item Response Theory

Measurement invariance can be examined using an item response theory (IRT) method or a CFA method. While a discussion of IRT is beyond the scope of the current work, the reader should be aware of that IRT based approaches are available. IRT and

CFA have been compared regarding how well they establish measurement invariance across different populations (Raju, Laffitte, & Byrne, 2002; Meade & Lautenschlager, 2004). Meade and Lautenschlager's (2004) comparisons of CFA & IRT were made using simulated data. Meade & Lautenschlager (2004) found that CFA and IRT each have advantages when assessing measurement invariance across different populations. IRT analysis is most preferable if the invariance of a single scale or specific scale item(s) is of interest for a particular research study (Meade & Lautenschlager, 2004). In addition, IRT analyses require larger sample sizes in order to adequately estimate parameters (e.g., item parameters and latent trait scores). The CFA analysis is preferable when the invariance of a multidimensional framework is being assessed for measurement invariance (Meade & Lautenschlager, 2004). Based on the results of Meade and Lautenschalger's study, the various measurement invariance tests between the IRT and CFA analyses provided different information based on sample sizes and/or the number of scale items; however, the results from an IRT method was similar to CFA results. Like other studies (Reise, Widaman, & Pugh, 1993), Raju et al. (2002) noted similar comparisons of the IRT and CFA approaches when establishing measurement invariance across populations. One difference between these two methods is that the IRT approach postulates a nonlinear relationship between the latent variable and the observed variable. Instead, the CFA approach often assumes a linear relationship between the latent variable and observed variable. The CFA approach can assess multiple dimensions and multiple populations simultaneously when assessing measurement invariance. Conversely, many of the IRT methods used to examine measurement invariance are typically confined to

unidimensional scales. Despite the differences between the IRT and CFA methods, both examine the relationship between latent variables and observed variables and more importantly, these two methods both provide a statistical framework within which between-group equality can be evaluate for the item parameters.

Previous studies (Cohen, Kim, & Wollack, 1996; Finch, 2005; Stark, Chernyshenko, & Drasgow, 2006; Wang & Shih, 2010; Woods, 2009) have used IRT methods to detect measurement invariance across heterogeneous groups (e.g., age, gender, race, ethnicity, etc.). From an IRT perspective, measurement invariance of items and subscales, or tests across subpopulations can be assessed using IRT-based techniques developed for studying differential item functioning (DIF) or item bias (Raju, Laffitte, & Byrne, 2002). There are several IRT-based techniques for investigating differential functioning of items and tests: Lord's (1980) chi-square; Raju's (1988, 1990) area measures; Thissen, Steinberg, and Wainer's (1988) likelihood ratio test; and Raju, van der Linden, and Fleer's (1995). These IRT-based DIF techniques examine the invariance of item parameters across two populations which are commonly referred to as the focal group and the reference group in DIF literature (Raju et al., 2002). If item parameters are invariant across the two populations, items are said to have measurement invariance or non-DIF according to the IRT-based DIF literature (Raju et al., 2002).

In IRT analyses, the item discrimination and item difficulty parameters are estimated when assessing measurement invariance which is analogous to the factor loadings and intercepts in multigroup CFA, respectively. In the MIMIC model, the intercept is estimated when measurement invariance is tested. Despite the similarities and

differences in both methods, SEM techniques (MIMIC models and multigroup CFA) share advantages over IRT methods. The advantages include: (1) both the MIMIC model and multigroup CFA methods can model multidimensional data while in IRT unidimensionality is typically assumed and (2) there are a greater number of well-established model fit indices in SEM than with IRT models (Finch, 2005; Kaplan, 2009). Thus, SEM methodological techniques (MIMIC model & multigroup CFA) are more favorable for this study because of the multidimensional framework of the SWS instrument is used to examine measurement invariance between the populations in this study.

Structural Equation Modeling

Byrne (1998) described SEM as a statistical methodology with a confirmatory approach (i.e. hypothesis-testing) to the multivariate analysis of a structural theory based on some phenomenon. SEM is also referred to as causal modeling, causal analysis, simultaneous equation modeling, and analysis of covariance structure (Kline, 2005). There are two important features of SEM (Byrne, 1998):

1. The causal processes under study are represented by a series of structural equations such as regression.
2. The structural relations are modeled in a diagram to provide visual conceptualization of the theory under study which represents the hypothesized model.

Using SEM techniques, a hypothesized model can be tested to examine the relationships between the latent and observed variables. Latent variables are variables that cannot be directly measured (Kline, 2005). A latent variable is sometimes called a factor or an underlying construct. Observed variables are variables that are directly measured (e.g., the measurement of temperature or weight) (Kline, 2005). Observed variables are also called measured variables, indicators or manifest variables. For example, the SWS instrument measures superwoman characteristics among women which include five subscales such as obligation to present an image of strength. Obligation to present an image of strength (from the SWS instrument) is a latent variable because it cannot be measured directly; however, it can be assessed indirectly using several observed variables. This latent variable include observed variables such as "I have to be strong," "I try to present an image of strength," and "I am expected to be the strong on in my family."

Figure 1 is a path diagram which depicts the hypothesized set of relationships described above in the obligation to present an image of strength subscale (latent variable). The path diagram provides a visual representation of the hypothesized model to be tested. In path diagrams, the latent variables (i.e. obligation to present an image of strength) are represented by circles or ovals and the observed variables (i.e. "I have to be strong," "I try to present an image of strength," "I am expected to be the strong one in my family") are represented by rectangles or squares. The arrows represent the relationships between the observed and latent variables. One way arrow represents a hypothesized

direct relationship between variables, and two way arrows represent a correlation between variables with no implied direction of effect.

In SEM, the hypothesized model is specified first and then estimated to examine the linear relationships among the latent and observed variables (Byrne, 1998; MacCallum & Austin, 2000). Once the hypothesized model is estimated, the model is evaluated to determine how well the model fits data. If the model fits well, the model suggests that hypothesized relations among variables are plausible (Byrne, 1998; Kline, 2005). If the model does not fit, the tenability of the hypothesized relations among the variables is rejected (Byrne, 1998; Kline, 2005). A misfit model can be tested again after re-specifying the hypothesized model according to the evaluation of the previously estimated model. More details will be discussed later about the primary steps of SEM.

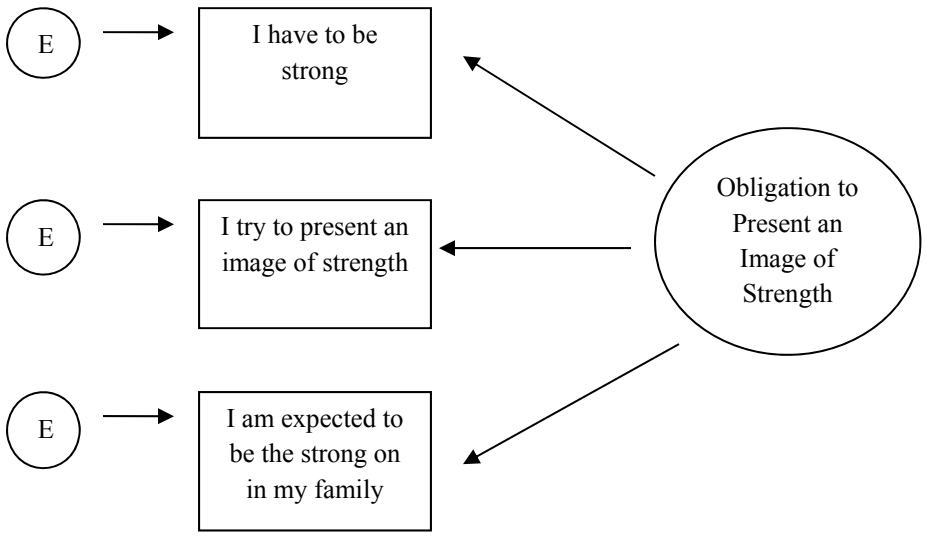

Figure 1. Path Diagram

LISREL is used in this study to examine the hypothesized model of the SWS. LISREL is the acronym for LInear Structural RELations which is based on the Jöreskog-Keesling-Wiley approach that represents systems of structural equations (Byrne, 1998). The LISREL full latent variable model in SEM includes two components: measurement model and structural model. The measurement model comprises the relationships between the latent variables and the observed variables. Measurement properties of the observed variables such as reliability and some aspects of validity are also described in the measurement model (Byrne, 1998). The structural model includes only the relationships among the latent variables. Specifically, the structural model specifies which latent variable(s) directly or indirectly causes changes in the values of other latent variable(s) (Byrne, 1998). This aspect assists in the evaluation of factorial validity of an instrument and/or construct. Factorial (structural) validity is defined as the degree to which the measure of a construct fits to the theoretical definition of the construct (Messick, 1995). Factorial validity is established by testing the fit of a theoretical based measurement model for describing the variances and covariances underlying items on a scale using CFA (Bollen, 1989). In addition, factorial validity is an aspect of construct validity that is established through factor analysis. Previous studies (Byrne, 1994; Barton, Andrew, & Schwab, 1994; Hull, Beaujean, Worell, & Verdisco, 2010; Bradely, Bagnell, & Brannen, 2010) have made predictions about how test scores of an instrument should behave based on a theory regarding the trait being measured using factorial analysis procedures to examine the psychometric properties of instruments.

SEM is a large sample technique and the type of estimation used in the analysis can affect sample size requirements (Byrne, 1998). Small sample sizes can cause some issues in analyses and limit the power of the analysis. There are several methods of estimations used for LISREL: instrumental variable method (IV), two-stage least squares (TSLS), generalized least squares (GLS), unweighted least squares (ULS), maximum likelihood (ML), weighted least squares (WLS), and diagonally weighted least squares (DWLS). Because there are various purposes and underlying assumptions of the parameter estimations, ML and DWLS are discussed in the context of this study. Several resources are available for more complete discussion of each method (Bollen, 1989; Hayduk, 1987; Jöreskog & Sörbom, 1993, 2006b).

In LISREL, ML estimation is the default estimator and is the most common estimator used to estimate parameters. ML estimation is an information estimator that simultaneously estimates all parameters and accounts for the full system equations including constraints and restrictions when developing estimates (Kline, 2005; Hambelton & Swaminathan, 1985). In other words, ML estimation maximizes the likelihood of a sample that is observed. ML estimation is iterative which means that it derives an initial solution and then attempts to improve estimates through subsequent cycles of calculations. The ML estimator is known to be consistent, asymptotically unbiased, asymptotically efficient, and asymptotically normal (Browne, 1984; Jöreskog, 1994; Kirby & Bollen, 2009). ML estimation assumptions in SEM include independence of observations, multivariate normality of endogenous variables, independence of the

exogenous variables and disturbances, and correct specification of the model (Kline, 2005).

An alternative estimator, DWLS, is considered if major violations of ML assumptions occur. DWLS is a weighted least squares estimator for ordered-categorical data and is defined as

$$F_{DWLS} = (r - p)' W^{-1}_{D} (r - p),$$

where p is a vector containing the unique elements of the $p \times p$ model implied correlation matrix (Jöreskog, 1994), and r is a vector containing unique elements of a $p \times p$ sample polychoric correlation matrix; W^{-1}_{D} contains only the diagonal elements of the full weight matrix which reduces the number of nonzero elements and reduces the computational burdens (Wirth & Edwards, 2007). DWLS requires chi-square and standard error adjustments (Wirth & Edwards, 2007). To adjust for the biased standard errors and test statistic (chi-square), the Satorra-Bentler scaled chi-square and robust standard errors (Satorra & Bentler, 1994; Yuan & Bentler, 1998) are used in this study. These methods adjust for the chi-square test statistic and standard errors of the parameters while the degrees of freedom are left unadjusted (Wirth & Edwards, 2007). According to Flora & Curran (2004), the Satorra-Bentler scaled chi-square and robust standard errors method work well in practice.

Confirmatory Factor Analysis

SEM uses a confirmatory technique to explain how the observed and latent variables are related to one another. Confirmatory factor analysis (CFA) is one of several

techniques used in SEM which entails testing *a prior* set of relationships between particular indicators (observed variables) and factors (latent variables). CFA is often referred to as the measurement model, because it focuses on how observed variables are linked to latent variables and not with causal relations among latent variables.

In order to test a CFA model, SEM requires that several steps are taken to build the hypothesized model (Byrne, 1998; Kline, 2005). First, specify the model using structural equations and/or diagrams to describe the hypothesized model to be tested. The equations used to describe the hypothesized model correspond to the presumed relations among observed and latent variables which are estimated by SEM software program using sample data (Kline, 2005). Second, decide if the model is identified. A model is identified when unique solutions for the values of parameters are found, the parameters are therefore estimable which makes the model testable (Byrne, 1998). Conversely, if the model cannot be identified, many sets of various different parameter estimates could fit the data equally (Byrne, 1998). Third, select measures of the variables represented in the model. Once this is completed, collect, prepare and screen the data for multivariate normality, homoscedasticity, missing data and multicollinearity. Fourth, select a computer program to estimate the model. In estimating the model within a computer program framework, the following steps are executed (Kline, 2005):

a) Evaluate the model fit. The model fit determines how well the model fits the data. If the model doesn't explain the data well, skip the rest of the steps and proceed to the fifth step.

b) Interpret the parameter estimates. Examine the linear association among the observed and latent variables.
c) Consider equivalent models. Equivalent models offer a competing account of the data compared to the preferred hypothesized model. Kline argues that a researcher should explain why the preferred model should not be rejected in favor of statistically equivalent models.

Fifth, re-specify the model and evaluate the fit of the revised model using the same data (only if needed). Sixth, accurately and completely describe the analysis results based on a satisfactory model obtained. Seventh, replicate the results of the study if possible. Kline (2005) mentions that many studies aren't replicated due to SEM general need for large samples which makes it hard for researchers to replicate SEM models. Eighth, apply the results which can be used to contribute to existing research and policy. For the purpose of this study, hypothesized models are evaluated to determine how well the collected data fits the models, parameter estimates are interpreted and if needed, models are re-specified and evaluated for fit including describing the final results of the study. Steps seven and eight are beyond the scope of the current work.

Multigroup Confirmatory Factor Analysis

Multigroup confirmatory factor analyses (multigroup CFA) is tested using SEM within the framework of a confirmatory factor analysis (CFA). Multigroup CFA is used to assess measurement invariance between latent and observed variables across groups. In order to examine group differences in a SEM using the framework of a multigroup CFA model, several hierarchal steps are taken to properly examine invariance across groups

using the two major components of invariance: measurement level invariance and structural level invariance. The hierarchical steps include: (1) configural invariance; (2) metric invariance; (3) intercept invariance; (4) residual variance invariance; (5) construct variance invariance; (6) construct covariance invariance; and (7) latent mean invariance. Hierarchical steps 1-4 are the measurement level invariance tests and hierarchical steps 5-7 are the structural level invariance tests. Not all invariance steps are required when measuring invariance depending on the purpose of the research study (Schmitt & Kuljanin, 2008; Vandenberg & Lance, 2000). The intention of this current study is to test: configural invariance, metric invariance, intercept invariance, residual variance invariance (i.e. measurement level invariance) and latent mean invariance (i.e. structural level invariance). These specific invariance tests address the research questions about the SWS instrument. The evaluation of measurement invariance models involves comparisons of nested models, constrained (cross-group equality) models and less constrained models. The constrained and less constrained models are tested using a goodness-of-fit index (i.e. chi-square difference test) to examine the comparisons of relative fit. If the fit of the more constrained model is worse than the less constrained model, the measures are not invariant and no higher level tests are needed to be examined. However, if the fit of the constrained model is not considerably worse than the less constrained model, then the observed variables (i.e. items) can be assumed to measure the factors of interest in comparable ways across groups.

The first hierarchical step addresses configural invariance. Configural invariance indicates that the same factors and pattern of factor loadings explains the variance-

covariance matrices associated with the groups' responses. In other words, configural invariance means that the factor structure for all groups is the same; however, the values of the parameters in the model may vary (Schmitt & Kuljanin, 2008). The configural invariance model is the baseline against which other more restrictive models of the data are compared to (Byrne, 1998; Schmitt & Kuljanin, 2008). So, in the current study this level of testing will check whether the five SWS subscales have the same number of factors and pattern of factor loadings associated with the responses with both young and middle aged women.

The second hierarchical step examined is the metric invariance. Metric invariance model examines the equivalent factor loadings across groups. For example, the values of the factor loadings in a model are constrained equal for each group being compared. Metric invariance can be considered strong invariance whereas configural invariance is a weaker form of invariance (Schmitt & Kuljanin, 2008).

The third hierarchical step is the examination of intercept invariance. Intercept invariance is also called scalar invariance or strong factorial invariance. Intercept invariance tests the invariance of item intercepts across groups. Intercept invariance is a prerequisite for the comparison of latent means across groups which indicate that the measurement scale has the same operational meaning across groups (Cheung & Rensvold, 2002). Without intercept invariance, the comparison of latent means across groups can be ambiguous.

The fourth hierarchical step tests residual variance invariance, which is also known as the equality of uniqueness. Residual variance invariance tests whether items on

a scale have the same internal consistency for groups or individuals (Cheung & Rensvold, 2002). Researchers testing for residual variance invariance are interested in knowing if groups have the same item residual variances. In testing the equality of uniqueness, the residuals of the regression equations for each indicator are equivalent across groups (Schmitt & Kuljanin, 2008). This step is essentially testing for the equality of the reliability across groups. Some researchers have argued that testing the equality of uniqueness is only legitimate if the latent factor variances are equal, and others consider this test of invariance difficult to obtain and is not needed even if you want to test the differences in latent means (Meredith, 1993; Schmitt & Kuljanin, 2008; Vandenberg & Lance, 2000). However, this study tests the residual variance invariance to see if both groups have equivalent residual variance invariance in the SWS instrument. If residual variance invariance is not established, it does not affect continued testing of measurement invariance.

The fifth hierarchical step addresses construct variance invariance. Construct variance invariance is also called equivalence of construct variance. Construct variance invariance tests whether the variances of constructs are the same across groups. Construct variance invariance must hold to compare correlations of constructs across groups (Cheung & Rensvold, 2002).

The sixth hierarchical step is the test of construct covariance invariance. Construct covariance invariance is also called equivalence of construct covariance. Construct covariance invariance tests whether the covariances are invariant across groups.

The seventh hierarchical step is called latent mean invariance or equivalence of latent means. Latent mean invariance tests the invariance of latent means across groups. The intercept invariance must be established to compare means across groups. So, the intercept invariance must hold before testing the latent mean invariance using the SWS instrument.

If invariance cannot be established in the steps mentioned above for this current study, attempts will be made to proceed with a test of a model which includes separate estimates of a subset of the subgroup parameters (e.g., some factor loadings, some intercepts). In other words, some but not all of the measurement parameters specified in the hypothesized model are constrained equal across groups in testing for measurement invariance which is called partial measurement invariance (Byrne, Shavelson, Muthén, 1989). Partial measurement invariance is where some parameters (equivalent) but not all of the parameters (non-equivalent) in the model are constrained equal across groups. Byrne et al. (1989) argues that if full measurement invariance is not established, then partial measurement invariance can be used to continue testing measurement invariance among other invariance models. Partial measurement invariance is only explored if full measurement invariance is not evident. In the current study, this level of invariance is examined if items on the SWS are equally discriminating (similar loadings) for the two age groups of women. If the items are not equally discriminating, they have different levels of importance of defining the constructs in the groups.

In review, the measurement level invariance includes hierarchical steps: configural invariance, construct-level metric invariance, item-level metric invariance,

residual variance invariance, and intercept invariance. The structural level invariance includes hierarchical steps: construct variance invariance, construct covariance invariance, and latent mean invariance. Overall, the following hierarchical steps are tested in this study to examine the invariance of the SWS instrument between young and middle aged African American women: configural invariance, metric invariance (construct- and item-level metric invariance), intercept invariance, residual variance invariance, and latent mean invariance.

Multiple Indicators Multiple Causes Model

Multiple indicators multiple causes (MIMIC) model is an alternative model used for estimating group differences on latent variables which was proposed by Jöreskog and Goldberger (1975). Latent variables with effect indicators are regressed on one or more dichotomous, cause indicators that represent group memberships (Jöreskog, & Goldberger, 1975). Instead of fitting the model to different groups separately in multigroup CFA, MIMIC models combine the groups in one variable which incorporates the membership variables as the cause indicators into the model. For example, suppose there is a simple factor model with one latent variable (motivation to succeed) and three observed variables ("I accomplish my goals with limited resources," "No matter how hard I work, I feel like I should do more," and "I put pressure on myself to achieve a certain level of accomplishment"), and the goal is to compare the factor mean of this model between two groups. The three observed variables are called effect indicators in MIMIC models. The latent variable, motivation to succeed, with its three effect indicators are regressed on the dichotomous (0 = young women and 1 = middle aged women) cause

indicator, age. The path coefficients for the direct effect of the age variable provides information about the degree to which the difference between young and middle aged women predicts the latent variable, motivation to succeed. The overall model can be presented as:

$$\eta = \nu + \gamma X + \zeta$$

where η is the factor, ν is the latent intercept (or the mean of the group with X coded as 0), ζ is the residual of the latent factor, and γ is actually the mean difference of the latent factor between the two groups.

In MIMIC models, the group differences are determined by examining the significance and magnitude of the factor loadings (Muthén, 1989). A limitation of MIMIC models is the assumption of invariance. However, if an instrument is not invariant across groups, group differences can still be examined using the MIMIC model because it does not require for constructs to be equivalent across groups. Conversely, the MIMIC model is an efficient method for handling population heterogeneity for validation research and is used to investigate potential differential item functioning in observed indicators of latent variables (Muthén, 1989).

There are several other advantages of using the MIMIC model: smaller sample sizes are permissible; grouping variable with two or more levels; and less parameter estimation. MIMIC models are better with small sample sizes compared to multigroup CFA because there is no need to divide the sample into different groups (Muthén, 1989). Because groups are analyzed separately in multigroup CFA, the sample size needs to be

relatively large for each group in order to have stable and accurate parameter estimates. Muthén (1989) argues that a grouping variable (i.e. cause indicator) increases the power to detect true heterogeneity in the sample. Furthermore, investigating group differences when more than two groups are present using multigroup CFA are difficult when compared to the MIMIC model where factors are regressed on one or more dichotomous indicators representing group membership (e.g. dummy variable, 0 = young women; 1 = middle aged women). The importance of using MIMIC models and other related group differences methodologies (e.g., multigroup CFA) is to ensure that researchers infer accurate explanation and comparisons of latent factor means. If invariance is not established for in a construct, then precise comparisons cannot be made among heterogeneous populations. In addition, the MIMIC model can be used to identify the source of invariance, if a test lacks invariance. Rubio, Berg-Weger, Tebb, & Rauch (2005) used the MIMIC model in conjunction with a multiple group model in measuring family care-giving experience among two populations, caregivers and non-caregivers. Using the MIMIC modeling combined with multi-group modeling, Rubio et al. (2005) were able to evaluate measurement invariance among the caregivers and non-givers using a caregiver well-being scale in their study. In addition, Rubio et al. (2005) were able to use MIMIC modeling to find the specific source of invariance via demographic variables used in the study when the multiple group modeling indicated factorial invariance in the model.

The MIMIC model is used as a supplemental analyses in this study to add to the examine if the MIMIC model findings lead to similar findings of the MGCFA model

across the young and middle-aged women groups using the SWS instrument. The same hypothesized model for the MGCFA is used for the MIMIC model. The five subscales of the SWS instrument (obligation to present an image of strength, obligation to suppress emotions, resistance to being vulnerable, intense motivation to succeed, and obligation to help others) are regressed on the dichotomous cause indicator, age variable. The age variable is a direct effect of the SWS five subscales which has 35 effect indicators. The path coefficients for the direct effects of the age variable will provide information about the degree to which the difference between young and middle aged women predicts each of the five subscales.

Model Fit

Goodness-of-fit indexes (GFI) are used to evaluate overall model fit for multigroup CFA and MIMIC model analyses in this study. A model has adequate fit (acceptable value of GFI) if the covariance structure implied by the model being tested is similar to the covariance structure of the sample data. In selecting the best-fitting models for multigroup CFA and MIMIC models, the following GFI are used in this study: chi-square (χ^2) statistics, comparative fit index (CFI), root mean square error of approximation (RMSEA), Akaike information criterion (AIC) and consistent AIC.

Chi-square (χ^2) is a most commonly fit statistics used in SEM which is the product (N-1) F_{ML}, where N is the sample size and F_{ML} is the value of the statistical criterion minimized in maximum likelihood (ML) estimation (Kline, 2005). ML estimation maximizes the likelihood of a sample observed and it assumes that population distribution is multivariate normal (Kline, 2005; Cheung & Rensvold, 2002). A

nonsignificant value of χ^2 statistics indicates failure to reject the null hypothesis that the hypothesized covariance matrix is identical to the observed covariance matrix (Cheung & Rensvold, 2002). This result means that there is adequate fit in the model. Despite χ^2 statistics popular use, it is assumed that samples are large and multivariate normal which can be an issue when small sample sizes are used. Cheung and Rensvold (2002) noted that other GFIs are proposed as alternatives to χ^2 statistics due to issues of sample size: CFI, Tucker-Lewis index, normed fit index, and RMSEA. Some of these GFIs do not have known sampling distributions instead criterion values are used to assess model fit; however, RMSEA has sampling distribution which is discussed in more details below.

The CFI is also a common model fit index used in the SEM literature. CFI tests the absolute fit of the model to data and it is considered to be very sensitive to sample size (Hu & Benlter, 1995). The CFI ranges in values from 0.00 to 1.00. A value of 0.90 indicates a good fit of the model to the data and a value of 0.95 or greater suggests that there is an excellent fit of the model to the data (Hu & Benlter, 1995; Wu, 2010).

RMSEA measures how well the model with optimally parameter estimates fit the population covariance matrix (Byrne, 1998). RMSEA is based on the noncentrality parameter which measures the degree of falseness in the null hypotheses. RMSEA values less than or equal to 0.05 is considered adequate fit, values between 0.05 and 0.08 is considered reasonable error of approximation, and values greater than .08 are considered poor fit of the model to the data (Browne & Cudeck, 1993).

The AIC and CAIC are parsimony fit indices also known as information criteria indices which adjust for sample sizes (Akaike, 1974). For AIC and CAIC, smaller values

suggest good model fit or indicate a parsimonious model. The AIC and CAIC are not normed to a 0-1 scale which makes it difficult to require a cut-off value other than suggesting that the model that produces the lowest value is the most superior (Akaike, 1974). Both the AIC and CAIC are used as supplemental indexes to compare nested models in this current study.

Once models have been assessed using model fit indexes, nested models are compared to determine which model best fits the data. The most common test used to evaluate nested models is called the likelihood ratio (LR) test. The LR test is also known as the chi-square difference ($\Delta\chi^2$) and is used to evaluate model fit differences between nested models (Bollen, 1989). The $\Delta\chi^2$ is calculated as

$$\Delta\chi^2 = \chi^2_c - \chi^2_{uc}$$

where χ^2_c is the value of the constrained model and χ^2_{uc} is the value of the unconstrained model. Significance is determined by the degrees of freedom (Δdf)

$$\Delta df = df_c - df_{uc}$$

The Satorra-Bentler scaled statistic (S-B χ^2) is used to evaluate the goodness of fit of models estimated using the DWLS estimator (Satorra & Bentler, 2001; Satorra & Bentler, 1988). S-B χ^2 is a correction for the χ^2 statistic when distributional assumptions are violated. Because of the reduction of information when using the DWLS estimator, the test statistic and standard errors are left biased. The S-B χ^2 adjusts the χ^2 statistic and standard errors of the parameters leaving the degrees of freedom unadjusted (Wirth &

Edwards). According to Hu, Bentler, & Kano (1992), S-B χ^2 is considered the most reliable test statistic for evaluating covariance structure models under various distributions and sample size.

A S-B χ^2 difference test is used to compute the difference between two S-B scaled goodness-of-fit test statistics for parameters estimated under the DWLS method. S-B scaled difference test statistic is defined as $T_d := Td/\hat{c}_d$ where $\hat{c}_d$ is a consistent estimate of

$$\hat{c}_d := 1/m(\text{tr } U_d \Gamma)$$

with

$$U_d = V \Pi P^{-1}A'(AP^{-1}A)^{-1} A P^{-1}\Pi'V$$

Satorra and Bentler (2001) noted that the nonnull eigenvalues of $U_d \Gamma$ are equal, then the scaled statistic T_d is asymptotically a chi-square statistic. For this study, a computer program called SBDIFF.EXE, which computes a significance test on the difference between Satorra-Bentler, scaled chi-square statistics (Crawford & Henry, 2003; Satorra & Benter, 2001).

CHAPTER III

METHODOLOGY

In this chapter, the sample and the procedures used to collect data are described including a discussion of the instrument, hypothesized model, and analyses used: multigroup confirmatory factor analysis and multiple indicators multiple causes model.

Sample

The sample data used in this study includes 561 women between the ages of 18 and 65. For the purposes of examining group differences, the women are divided into two age groups: young women (age < 39, n = 386) and middle-aged women (age > 40, n = 175). This sample (N = 561) originated from a sample of 674 women, which reflects the degree of missing data mainly due to the age variable. In addressing this issue of incomplete data, listwise deletion was used to delete missing cases. Listwise deletion also known as complete-cases analysis, removes any cases with any missing values on variables to be included in the model (Hoyle, 2012). Listwise deletion was chosen because much of the missing data (17%) occurred on the age variable. The final sample size after listwise deletion included 561. Age is critical variable in this study and did not seem appropriate for imputation.

Procedure

The research team recruited participants by contacting various organizations and groups at colleges and universities, organizations, and groups across the United States. Email correspondents and flyers were distributed to recruit study participants. Interested

groups and individuals were contacted by the research team to conduct a preliminary screening for study eligibility. Eligible participants were provided a date and location for data collection or administered the survey using an online survey software called Qualtrics. Study participants were entered into a drawing for a chance to win one of three $30 gift certificates by completing the survey. Secondary data is included in the study sample. The secondary sample included a sample of 163African American women from the San Francisco Bay area who were administered the SWS via an interview, and 57 African American women who were administered the SWS using a paper-pencil version. The identity of study participants from the secondary data samples were de-identified for the purposes of this study and to ensure confidentiality/anonymity requirements were met by according to the Institutional Review Board (IRB). The data collected is stored in a secured file, which is accessible by the research team. The data collection administered online included 453 participants which was collected from four to six weeks during the summer of 2012 (May - June).

Instrumentation

The SWS instrument is a 35-item measure including five subscales: obligation to present an image of strength (six items), obligation to suppress emotions (seven items), resistance to being vulnerable (seven items), intense motivation to succeed (six items), and obligation to help others (nine items). All items on the SWS instrument are statements which the participants rate using the following response scale: *this is not true for me* and *this is true for me rarely, this is true for me sometimes, and this is true for me all of the time.* Also, the participants are instructed that if they rate an item as *true for me*,

they have the option to indicate if an item bothers them using the following rating scale: *very much, somewhat, or not at all*. However, this information was not be used in this current study. A copy of the SWS instrument is presented Appendix A.

Table 1 shows the subscales and total items included in the SWS instrument. The current SWS instrument was reconstructed based on content expert recommendations and statistical analyses findings. Reliability analyses were constructed from a pilot study of the SWS framework using the very first version of the SWS instrument. This pilot study included a small sample of 28 African American women who completed the first SWS instrument. The first SWS instrument included eleven subscales: obligation to present an image of strength (10 items; $\alpha = 0.72$); obligation to suppress emotions (32 items; $\alpha = 0.98$); resistance to being vulnerable (20 items; $\alpha = 0.96$); resistance to dependence (16 items; $\alpha = 0.94$); taking on multiple roles and responsibilities (22 items; $\alpha = 0.89$); defy the odds (29 items; $\alpha = 0.93$); intense motivation to succeed (9 items; $\alpha = 0.89$); lack of self-care (14 items; $\alpha = 0.91$); obligation to help others (10 items; $\alpha = 0.98$); expectations to achieve (14 items; $\alpha = 0.98$); and difficulty relinquishing control (7 items; $\alpha = 0.92$).

Table 1. SWS Instrument Subscales

Superwoman Schema Subscale	New Item Total
Subscale 1: Obligation to present an image of strength	6
Subscale 2: Obligation to suppress emotions	7
Subscale 3: Obligation to help others	9
Subscale 4: Resistance to being vulnerable	7
Subscale 5: Intense motivation to succeed	6

Based on the expert content recommendations and statistical analyses findings, the current SWS instrument in this study were reconstructed which now includes five subscales and 35 items. This pilot study of the SWS is presented to provide some background information about the psychometric properties of the instrument since no previous psychometric analysis has been conducted before this study

Hypothesized Model

The confirmatory factor analysis (CFA) model in the present study hypothesized *a priori* that: (a) responses to the SWS could be explained by five subscales, (b) each item would have a non-zero loading on the SWS subscale it was designed to measure, and zero loadings on all other subscales, (c) the five subscales would be correlated and, (d) measurement error terms would be uncorrelated (Byrne, 1994) (see Figure 2).

Preliminary Analysis

Preliminary analyses were conducted to examine the quality of the data and multivariate assumptions before conducting a SEM analyses on the hypothesized model of the SWS instrument. Byrne and Campbell (1999) argued that reporting of preliminary analyses is important, particularly in cross-group studies, which can have an impact on the findings of a multigroup comparison. These analyses may impact the choice of estimation technique (ML versus DWLS) if multivariate assumptions are violated.

Analysis Overview

First, this study examines the factorial structure of SWS instrument by examining the baseline model of the multigroup CFA on the sample of African American women.

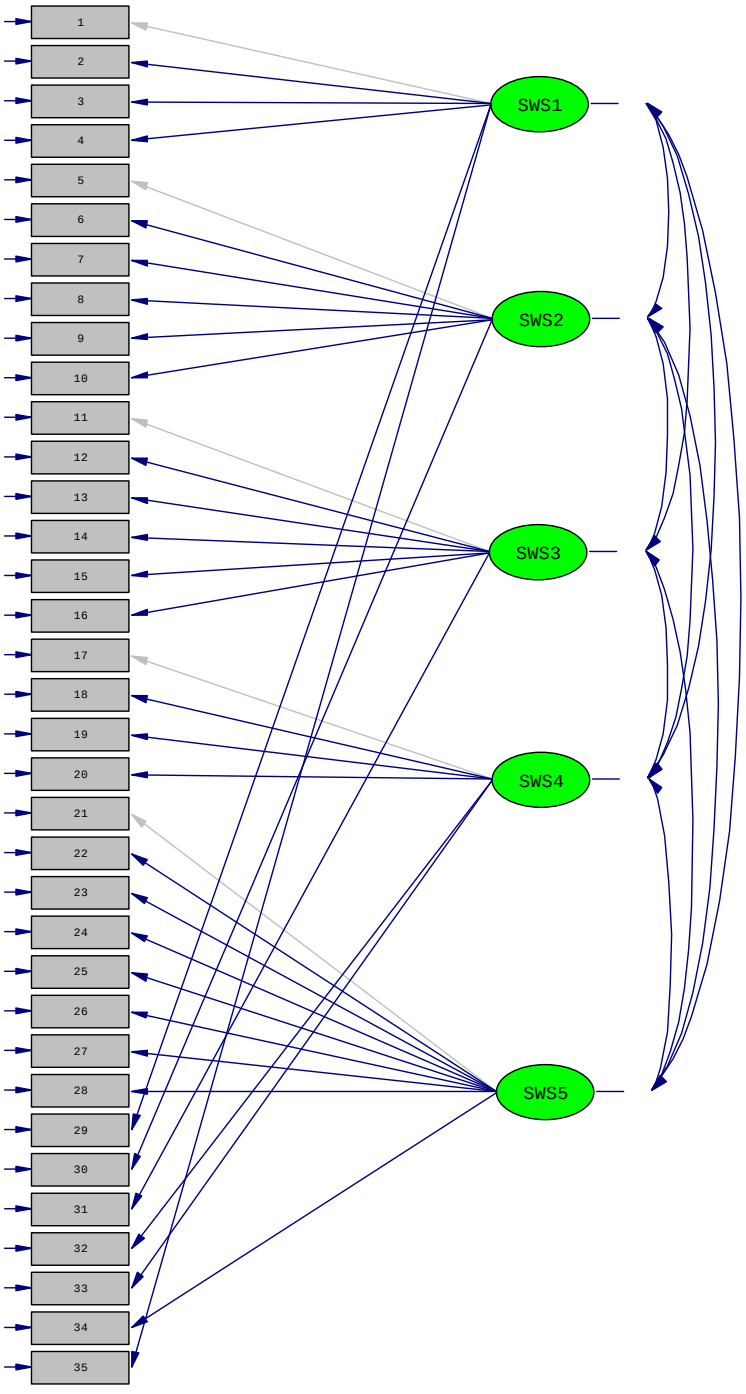

Figure 2. Superwoman Schema Hypothesized Model

Next, measurement invariance is tested using a multigroup CFA. Lastly, a MIMIC model is examined using the same hypothesized model used for the multigroup CFA model. Comparisons are made between the two SEM approaches, MIMIC model and multigroup CFA.

Estimation

Maximum Likelihood (ML) estimation is a common estimator used to analyze parameters in LISREL. ML is asymptotically consistent, unbiased, and normally disturbed under the assumption of multivariate normality of observed variables. Even though ML produces relatively unbiased parameter estimates, the model chi-square tends to be inflated and the standard error estimates are deflated under non-normality (Bollen, 1989). However, Satorra and Bentler (1994) provided an alternative to estimate parameters which provide a better approximate of the chi-square distribution known as a robust asymptotic covariance matrix, Sattora-Bentler scaled statistic.

Another estimator used to estimate parameters is called the diagonally weighted least squares (DWLS). DWLS is often used when parameters are non-normal and categorical. It is also commonly used with alternative correlations that estimate association between latent variables for ordinal data (polychoric correlation between categorical variables). When ML is used to estimated ordinal variables, parameter estimates may be underestimated and standard errors estimates may be negatively biased. DWLS helps correct and provide better parameter estimates and standard errors for categorical variables. DWLS is used in this study because the responses to the items are ordinal and categorical.

The estimated parameters in the CFA model are evaluated using the following goodness-of-fit indices as previously mentioned in the literature review: Chi-square statistic (χ^2), Satorra-Bentler scaled chi-square test statistic (S-B χ^2), the comparative fit index (CFI), root mean square error of approximation (RMSEA), akaike information criterion (AIC) and consistent akaike information criterion (CAIC).

Factorial Structure Test

The first research question addresses the issue of determining the factorial structure of the SWS instrument: *Does the five factor structure proposed by the superwoman schema exist within a population of African American women?* The hypothesized structure described in chapter 2 and mentioned above is tested using DWLS estimation in the baseline model of the multigroup CFA model. Model fit is evaluated using the criteria described in the preceding section.

Measurement Invariance Tests

The second research question pertains to conducting a cross-group study using the SWS instrument: *Is the superwoman schema instrument invariant across both young and middle-aged women in the intended population?* Multigroup CFA using SEM techniques was used to examine this research question and its sub-questions with DWLS estimation. Model fit was evaluated using the criteria described above.

The following hierarchical analyses were conducted to test measurement invariance: configural invariance, metric invariance, intercept invariance, residual variance invariance, and latent mean invariance. Table 2 (adapted from Cheung & Rensvold, 2002) provides an overview of the proposed models to be tested for invariance

of the SWS instrument. These models are hierarchical in that they cannot be examined unless the first invariance model holds. Each analysis is described under the corresponding hypothesis subheading below.

Configural Invariance. *Do the groups have the same factor structure?* Configural invariance examined whether the five-factor model of the SWS holds for both the young and middle-aged women. For young and middle-aged women, the fixed and free factor loadings have the same pattern in the configural invariance model, but no equality constraints. Model fit of the configural invariance is evaluated using the index criteria described above. This configural invariance model serves as a baseline model to which other restrictive models were compared.

Metric Invariance. *Do the groups have the same factor loadings?* Metric invariance examined if the strength of the relationship between factors and items were the same for both young and middle-aged women. The factor loadings are constrained to be equal across groups, but no other equality constraints are imposed. Model fit of the metric invariance is evaluated using the index criteria described above. The model is evaluated in relation to the previous model (configural model) using the criteria described in the preceding section.

Intercept Invariance. Do the groups have the same item intercepts? Intercept invariance tested whether the latent means indicated by the SWS items are the same for both young and middle-aged women. The factor loadings and intercepts are constrained to be equal across groups. Model fit of the configural invariance is evaluated using the same criteria as with previous models. The model ias evaluated in comparison to the

previous model (metric invariance) using the criteria described in the preceding section if the previous model holds.

Residual Variance Invariance. Do the groups have the same item residual variances? Residual variance invariance examined if the SWS items have the same internal consistency for young and middle-aged women. The factor loadings, intercepts, and theta deltas are constrained to be equal across groups. The same fit index criteria used with the earlier model are used to evaluate the model fit. This model is evaluated in comparison to the previous model (metric invariance) using the same criteria mentioned above.

Latent Mean Invariance. Are the latent means invariant across groups? The latent mean invariance model examined if the latent means were invariant across both young and middle-aged women. The factor loadings, intercepts, theta deltas, and kappas are constrained equal across groups for this model. Model fit of the configural invariance is evaluated using the index criteria described above. This model is evaluated in comparison to the intercept invariance model using the criteria described above if the previous model holds.

Multiple Indicators Multiple Causes Model

Multiple indicators and multiple causes (MIMIC) model is used to investigate the remaining research question: *Does using a MIMIC model lead to conclusions that are similar to those conclusions reached using a multigroup CFA?* The hypothesized model (see Figure 2) used in examining invariance in this study is used for the MIMIC model. In assessing group differences, factors with effect indicators are regressed on one

dichotomous (age) cause indicator which represents group membership. The path coefficients for the direct effects of the grouping variables provide information about the degree to which the differences of age predicts the SWS subscales: obligation to present an image of strength, obligation to suppress emotions, obligation to help others, resistance to being vulnerable, and intense motivation to succeed. The MIMIC model analyses are estimated in LISREL 8.8 (Jöreskog & Sörbom, 2006a) software using ML and DWLS estimations and model fit indexes (χ^2 statistics, CFI, RMSEA, AIC, CAIC) are used to examine the overall model fit of the MIMIC models in the same fashion as the multiple-group models.

Table 2. Hypotheses of Measurement Invariance

Model	Hypothesis	Hypothesis Test	Hypothesis Name	Description
1	H_{form}	Overall fit	Configural invariance	The test of whether the 5-factor model of the SWS holds for both the young and middle-aged women
2	H_{Λ}	2-1	Metric Invariance	The strength of the relationship between factors and items are the same for both young and middle-aged women
3	$H_{\Lambda,\nu}$	3-2	Intercept Invariance	The intercepts are invariant across the young and middle-aged women
4	$H_{\Lambda, \nu, \Theta(\delta)}$	4-2	Residual variance invariance	The items have the same internal consistency for young and middle-aged women
5	$H_{\Lambda,\nu\,\Theta(\delta),\kappa}$	5-2	Latent mean invariance	The latent means are invariant across young and middle-aged women groups

CHAPTER IV

RESULTS

In this chapter, the results of the study are provided using the research questions provided in Chapter I. First, preliminary analyses are presented which provide an overview of the data collected. Next, the factorial structure of the Superwoman Schema (SWS) is presented based on the hypothesized model described in Chapter III. Then, the results of the measurement invariance tests are described for the SWS instrument. Lastly, the MIMIC model results are presented which summarizes the group differences based on the SWS hypothesized model.

Preliminary Analysis

Missing Values. Missing value analysis was conducted on the SWS data to examine the pattern of missing data. In particular, the missing value analysis evaluated the location of missing values, the extent of missing values, and the likelihood that values were missing at random (IBM SPSS, 2011). In addition, the missing values analysis performed the following functions: estimated means, standard deviations, and expectation maximization (EM) method. The results are discussed below.

An overview of the extent of the missing data for SWS subscale one *(obligation to present an image of strength)* is displayed in Table 3. The full sample from the data collection (n = 674) was used for the missing data analysis conducted in SPSS. The final

sample size was reduced from 674 to 561 after analyzing the missing data. The number of missing values for each of the items includes item 1= 5; item 2 = 11; item 3 = 15; item 4 = 23; item 29 = 109; and item 35 = 114. Item 35 ("I have to be strong because I am a woman") has the highest number missing values (114). Note that the survey items were administered in order despite the ordering presented here. Item 1 ("I try to present an image of strength") has the least number (0.7%) of missing values. Approximately 17% of age values are missing for the overall SWS scale.

Table 3. Missing Values for SWS Subscale 1

				Missing	
	N	**M**	**SD**	**Count**	**Percent**
Item 1	668	2.50	0.67	5	0.7
Item 2	662	2.52	0.64	11	1.6
Item 3	658	2.36	0.91	15	2.2
Item 4	650	2.38	0.81	23	3.4
Item 29	564	1.76	1.12	109	16.2
Item 35[a]	559	2.01	1.07	114	16.9
Age	561	35.16	8.79	112	16.6

[a] This item is located towards the end of the scale in chronological order.

An *obligation to suppress emotions* subscale missing values are displayed in Table 4. The number of missing values for item 5 = 26; item 6 = 32; item 7 = 35; item 8 = 35; item 9 = 37; item 10 = 42; and item 30 = 109. Item 30 ("I keep my problems to myself to prevent from burdening others") has the most missing values; however, this item was located at the end of the SWS scale. Item 5 ("I display my emotions in privacy") has the least (3.9%) amount of responses missing.

Resistance to being vulnerable subscale missing values are presented in Table 5. The number of missing values for item 11 = 45; item 12 = 46; item 13 = 46; item 14 = 51; item 15 = 55; item 16 = 58 and item 31 = 110. The item with the most missing values is item 31 ("I do things by myself without asking for help"), which is not surprising because this item was located towards the end of the scale. Item 11 ("It's hard for me to accept help from others") has the least number (6.7%) of missing cases.

Table 4. Missing Values for SWS Subscale 2

				Missing	
	N	**M**	**SD**	**Count**	**Percent**
Item 5	647	2.10	0.80	26	3.9
Item 6	641	1.80	0.89	32	4.8
Item 7	638	0.88	1.04	35	5.2
Item 8	638	1.56	1.00	35	5.2
Item 9	636	1.69	0.95	37	5.5
Item 10	631	1.17	1.03	42	6.2
Item 30[a]	564	1.91	1.00	109	16.2

[a] This item is located towards the end of the scale in chronological order.

Intense motivation to succeed subscale missing values are provided in Table 6. The number of missing values for item 17 = 59; item 18 = 64; item 19 = 65; item 20 = 69; item 32 = 112; and item 33 = 114. As expected, item 33 ("I am a perfectionist") and item 32 ("The only way for me to be successful is to work hard"), had the most missing values. Item 17 ("I accomplish my goals with limited resources") had the least number (8.8%) of missing values.

Table 5. Missing Values for SWS Subscale 3

				Missing	
	N	**M**	**SD**	**Count**	**Percent**
Item 11	628	1.59	1.02	45	6.7
Item 12	627	1.75	0.95	46	6.8
Item 13	627	1.67	1.01	46	6.8
Item 14	622	1.64	1.08	51	7.6
Item 15	618	1.36	1.08	55	8.2
Item 16	615	2.23	0.81	58	8.6
Item 31[a]	563	2.12	0.79	110	16.3

[a] This item is located towards the end of the scale in chronological order.

Table 6. Missing Values for SWS Subscale 4

				Missing	
	N	**M**	**SD**	**Count**	**Percent**
Item 17	614	1.96	0.88	59	8.8
Item 18	609	2.46	0.77	64	9.5
Item 19	608	1.86	1.01	65	9.7
Item 20	604	2.17	0.88	69	10.3
Item 32[a]	561	2.35	0.87	112	16.6
Item 33[a]	559	1.77	1.01	114	16.9

[a] This item is located towards the end of the scale in chronological order.

An *obligation to help others* subscale missing values are displayed in Table 7. The number of missing values for each item includes: item 21 = 73; item 22 = 74; item 23 = 77; item 24 = 81; item 25 = 79; item 26 = 82; item 27 = 86; item 28 = 84; and item 34 = 113. Overall, there seemed to be a consistent pattern with missing data in this study. The items towards the end of the scale had the most missing values which are demonstrated in the tables above. The number of missing values increased as the items were listed chronological on the SWS scale. Respondents may have begun with

intentions of completing the survey but didn't complete the survey due to unforeseen circumstances or felt less motivated to complete the survey, particularly with participants taking the survey online.

Table 7. Missing Values for SWS Subscale 5

				Missing	
	N	M	SD	Count	Percent
Item 21	600	1.74	0.99	73	10.8
Item 22	599	1.66	1.10	74	11.0
Item 23	596	1.99	0.93	77	11.4
Item 24	592	2.05	0.94	81	12.0
Item 25	594	1.88	0.91	79	11.7
Item 26	591	1.66	1.06	82	12.2
Item 27	587	1.32	0.99	86	12.8
Item 28	589	1.11	1.11	84	12.5
Item 34[a]	560	1.45	1.04	113	16.8

[a] This item is located towards the end of the scale in chronological order.

Missing Completely At Random. Table 8 shows the Missing Completely At Random (MCAR) analysis for the 35 items of the SWS instrument including the age variable. MCAR is established when missingness does not depend on the values of variables in the data set (Little, 1988). Little's chi-square statistic for testing whether values are missing completely at random is used in this study. If chi-square value is significant, the data are not missing completely at random. However, if the chi-square is not significant, then the data are missing completely at random and multiple imputation methods can be conducted to generate possible values for missing cases. Based on the results of the MCAR analysis, the data are not missing completely at random, $\chi^2 = 2297.18$, df = 1937, $p \leq .01$. In this current study, listwise deletion is used since the data

are not missing completely at random. With listwise deletion, approximately 100 cases were excluded out of a total of 674 cases, which still provided an appropriate sample size to analyze the data.

Table 8. MCAR Estimates for SWS 35 Items

Variables	EM Means[a]	Variables	EM Means[a]	Variables	EM Means[a]
Item 1	2.50	Item 15	1.36	Item 29	1.75
Item 2	2.52	Item 16	2.23	Item 30	1.92
Item 3	2.36	Item 17	1.96	Item 31	2.12
Item 4	2.38	Item 18	2.46	Item 32	2.35
Item 5	2.10	Item 19	1.86	Item 33	1.77
Item 6	1.80	Item 20	2.17	Item 34	1.45
Item 7	0.88	Item 21	1.74	Item 35	1.99
Item 8	1.56	Item 22	1.66	Age	35.14
Item 9	1.69	Item 23	1.99		
Item 10	1.17	Item 24	2.05		
Item 11	1.59	Item 25	1.88		
Item 12	1.75	Item 26	1.66		
Item 13	1.67	Item 27	1.32		
Item 14	1.64	Item 28	1.11		

[a] MCAR test: $\chi^2 = 2297.18$, $df = 1937$, $p \leq .01$

Scale Reliability. The scale reliability for SWS was assessed by calculating the Cronbach's alpha for each subscale. The Cronbach's alpha is a common measure of reliability used to assess the internal consistency reliability of items and/scores (Cohen, 1988). A commonly acceptable rule of thumb for describing internal consistency using Cronbach's alpha (Cronbach, 1951) includes: $\alpha \geq 0.9$ is excellent; $0.7 \leq \alpha < 0.9$ is good; $0.6 \leq \alpha < 0.7$ is acceptable; $0.5 \leq \alpha < 0.6$ is poor; and $\alpha < 0.5$ is unacceptable. According to Cronbach's (1951) acceptable values, the following SWS subscales provide good internal consistency (Table 9): obligation to suppress emotions ($\alpha = .85$), resistance to being vulnerable ($\alpha = .86$), and obligation to help others ($\alpha = .87$). The obligation to

present an image of strength (α = .70) and an intense motivation to succeed (α = .71) have low subscale reliability and should be improved in future versions of the scale.

Subscores. The raw score ranges for each of the SWS subscales included the following: an obligation to present an image of strength with a range of 0 to 18, an obligation to suppress emotions with a range of 0 to 21, a resistance to being vulnerable with a range of 0 to 21, an intense motivation to succeed with a range of 0 to18, and an obligation to help others with a range of 0 to 27 (see Table 10). The average subscores for the SWS sample are also presented in Table 10. The following are the subscale items' averages: obligation to present an image of strength (13.61); obligation to suppress emotions (11.06); resistance to being vulnerable (12.34); intense motivation to succeed (12.63); and obligation to help others (14.94).

Table 9. SWS Subscale Reliability (N = 561)

Subscales	Cronbach's Alpha	N of Items
1. Obligation to Present an Image of Strength	0.70	6
2. Obligation to Suppress Emotions	0.85	7
3. Resistance to Being Vulnerable	0.86	7
4. Intense Motivation to Succeed	0.72	6
5. Obligation to Help Others	0.87	9

Table 10. SWS Raw Subscores (N = 561)

Subscales	N Items	Range	M Subscore	SD Subscore
1. Obligation to Present an Image of Strength	6	0 - 18	13.61	3.37
2. Obligation to Suppress Emotions	7	0 - 21	11.06	4.93
3. Resistance to Being Vulnerable	7	0 - 21	12.34	5.05
4. Intense Motivation to Succeed	6	0 - 18	12.63	3.52
5. Obligation to Help Others	9	0 - 27	14.94	6.45

Subscores' Correlations. To investigate if there was a statistically significant association among the SWS subscales, a correlation was computed for each subscale. Table 11 shows that the five subscales were significantly correlated. The strongest correlation was between SWS subscale 2 (obligation to suppress emotions) and SWS subscale 3 (resistance to being vulnerable), r (561) = .65, $p < .01$. The significant correlations show that women who feel an obligation to suppress their emotions are likely to feel more resistance to being vulnerable. The weakest correlation is between subscale 1 (obligation to present an image of strength) and subscale 3 (resistance to being vulnerable), r (561) = .28, $p \leq .01$. Although these subscales are significantly correlated, the level of correlation is low enough to suggest discriminant validity.

Table 11. SWS Subscale Correlation (N = 561)

Subscales	SWS1	SWS2	SWS3	SWS4	SWS5
1. Obligation to Present an Image of Strength	1				
2. Obligation to Suppress Emotions	.298*	1			
3. Resistance to Being Vulnerable	.278*	.651*	1		
4. Intense Motivation to Succeed	.424*	.462*	.577*	1	
5. Obligation to Help Others	.336*	.501*	.571*	.526*	1

*Correlation is significant at the 0.01 level

Descriptive Statistics. Descriptive statistics for the *obligation to present an image of strength* subscale are presented in Table 12. The highest mean score for this subscale includes item 2 ("I have to be strong"; M = 2.51 and SD = .65) and item 1 ("I try to present an image of strength"; M = 2.49 and SD = .69) compared to the other items. If item 29 ("The struggles of my ancestors require me to be strong") were deleted from this subscale, the Cronbach's alpha would increase to .72 from .70. Although items could be

removed to improve reliability, the violations were not great enough to make that decision at this point. Instead, items were retained in the interest of maintaining construct validity. Also, the item-total correlations for an obligation to present an image of strength subscale items imply good discrimination. (see Table 12).

The *obligation to suppress emotions* subscale descriptive statistics are presented in Table 13. Item 5 ("I display my emotions in privacy") has the highest mean score for this subscale (M = 2.09; SD = .82). The lowest mean for this subscale is item 7 ("My tears are a sign of weakness") with a mean of 0.87 and standard deviation of 1.05. Furthermore, if item 7 were deleted from this subscale, the Cronbach's alpha will increase to .85 from .84. The subscale 2 items are discriminating (item-total correlations) well among the participants.

Table 12. Descriptive Statistics for SWS Subscale 1 (N = 561)

Obligation to Present an Image of Strength	M	SD	Skewness	Kurtosis	Item-Total Correlation	Cronbach's Alpha if Deleted
Item 1	2.49	.693	-1.44	2.18	0.52	0.65
Item 2	2.51	.648	-1.35	2.12	0.54	0.64
Item 3	2.34	.930	-1.35	0.82	0.42	0.67
Item 4	2.37	.830	-1.34	1.28	0.50	0.64
Item 29	1.76	1.12	-0.39	-1.23	0.31	0.72
Item 35	2.01	1.07	-0.78	-0.69	0.46	0.66

The descriptive statistics for the *resistance to being vulnerable* subscale is depicted in Table 14. Item 16 ("If I want things done right, I do them myself") had the highest endorsement among the women in this study compared to the other items in this subscale, M = 2.22; SD = .84. Item 15 ("I resist help to prove that I can make it on my

own") has the lowest mean score of 1.38 and standard deviation of 1.09. Based on the Cronbach's alpha deleted, removing any of the items in this subscale would cause alpha to decrease.

Table 13. Descriptive Statistics for SWS Subscale 2 (N = 561)

Obligation to Suppress Emotions	M	SD	Skewness	Kurtosis	Item-Total Correlation	Cronbach's Alpha if Deleted
Item 5	2.09	0.82	-0.95	0.81	0.49	0.84
Item 6	1.79	0.91	-0.70	-0.22	0.68	0.82
Item 7	0.87	1.05	0.73	-0.91	0.44	0.85
Item 8	1.54	1.02	-0.25	-1.07	0.77	0.80
Item 9	1.67	0.96	-0.42	-0.75	0.60	0.83
Item 10	1.15	1.03	0.27	-1.23	0.65	0.82
Item 30	1.91	1.00	-0.60	-0.69	0.62	0.82

Table 14. Descriptive Statistics for SWS Subscale 3 (N = 561)

Resistance to Being Vulnerable	M	SD	Skewness	Kurtosis	Item-Total Correlation	Cronbach's Alpha if Deleted
Item 11	1.58	1.03	-0.29	-1.06	0.65	0.83
Item 12	1.76	0.97	-0.38	-0.80	0.54	0.85
Item 13	1.66	1.01	-0.32	-0.97	0.68	0.83
Item 14	1.63	1.09	-0.28	-1.23	0.76	0.82
Item 15	1.38	1.09	0.01	-1.35	0.68	0.83
Item 16	2.22	0.84	-1.01	0.54	0.51	0.85
Item 31	2.12	0.80	-0.88	0.69	0.53	0.85

Descriptive statistics for the *intense motivation to succeed* subscale are presented in Table 15. Item 18 ("It is very important to me to be the best at the things that I do") has the highest mean score (M = 2.44; SD = .79) for the sample in this study. Item 33 ("I am a perfectionist") has the lowest mean score for this subscale (M = 1.78; SD = 1.01). Item 17

has the lowest item total correlation among the other items (.31) which indicates good discrimination.

Table 15. Descriptive Statistics for SWS Subscale 4 (N = 561)

Intense Motivation to Succeed	M	SD	Skewness	Kurtosis	Item-Total Correlation	Cronbach's Alpha if Deleted
Item 17	1.94	0.89	-0.71	-0.10	0.31	0.72
Item 18	2.44	0.79	-1.37	1.31	0.49	0.67
Item 19	1.86	1.02	-0.52	-0.85	0.47	0.67
Item 20	2.15	0.90	-0.90	0.90	0.55	0.65
Item 32	2.36	0.87	-1.39	1.23	0.43	0.69
Item 33	1.78	1.01	-0.56	-0.74	0.48	0.67

Table 16. Descriptive Statistics for SWS Subscale 5 (N = 561)

Obligation to Help Others	M	SD	Skewness	Kurtosis	Item-Total Correlation	Cronbach's Alpha if Deleted
Item 21	1.73	0.99	-0.42	-0.84	0.58	0.86
Item 22	1.64	1.10	-0.22	-1.29	0.61	0.86
Item 23	1.99	0.94	-0.71	-0.35	0.70	0.85
Item 24	2.05	0.95	-0.78	-0.28	0.63	0.85
Item 25	1.88	0.92	-0.55	-0.44	0.61	0.86
Item 26	1.67	1.06	-0.32	-1.12	0.41	0.88
Item 27	1.33	1.00	-0.08	-1.22	0.62	0.86
Item 28	1.13	1.13	0.36	-1.36	0.66	0.85
Item 34	1.45	1.05	-0.13	-1.23	0.70	0.85

Descriptive statistics for the *obligation to help others* subscale are presented in Table 16. For this subscale, item 24 ("I feel obligated to take care of others") has the highest mean score (M = 2.05; SD = .955). Item 28 ("I feel guilty when I take time for myself") has the lowest mean score of 1.13 and standard deviation of 1.13. The item-total correlation for subscale 5 items has good discrimination.

Group Means. A set of independent t-tests was conducted to analyze the difference between the two groups in this study, young women (ages 18-39) and middle-aged women (ages 40-65), on the scale level among the five subscales. The results for the independent t-test analysis are displayed in Table 17. To control family-wise error, a Bonferroni correction was applied, making the appropriate alpha for significance .01 rather than .05. Three subscales showed significant differences (subscales 2, 3, and 4). For subscale two (obligation to suppress emotions), there was a significant ($p \leq .01$) small to medium effect ($d = .32$), with middle-aged women's scores ($M = 9.92$; $SD = 4.99$) significantly lower than scores for young women ($M = 11.53$; $SD = 4.83$).

Table 17. Comparison of Young and Middle-Aged Women on SWS Subscales (N = 386 young women and N = 175 middle-aged women)

Subscales	M	SD	*t*	*df*	*p*
Subscale 1			.734	528	.46
Young	13.68	3.23			
Middle-Aged	13.45	3.66			
Subscale 2			3.53	527	.00
Young	11.53	4.83			
Middle-Aged	9.92	4.99			
Subscale 3			3.28[a]	301[a]	.00
Young	12.86	4.81			
Middle-Aged	11.28	5.35			
Subscale 4			4.98[a]	290[a]	.00
Young	13.17	3.27			
Middle-Aged	11.50	3.75			
Subscale 5			.354	530	.72
Young	15.00	6.29			
Middle-Aged	14.79	6.79			

[a] The *t* and *df* were adjusted because variances were not equal.

Subscale three (resistance to being vulnerable) had a significant ($p \leq .01$) small to medium effect ($d = .31$), with middle-aged women ($M = 11.28$; $SD = 5.35$) scoring

significantly lower than young women (M = 12.86; SD = 4.81). For subscale four (intense motivation to succeed), middle-aged women scored (M = 11.50; SD = 3.75) significantly lower than young women (M = 13.17; SD = 3.27) which had a significant ($p \leq .01$) medium effect size (d = .47). Conversely, the two age groups did not differ significantly on subscale one (obligation to present an image of strength; p = .46) or subscale five (obligation to help others; p = .72).

Factorial Structure Results

Multi-group confirmatory factor analysis (CFA) was used to examine research question one: *Does the five factor structure proposed by the SWS instrument adequately describe survey responses from African American women?* Table 18 provides the factor loadings for the factorial structure of the SWS instrument. The goodness of fit statistics indicated a good model fit ($S\text{-}B\chi^2_{(1100)} = 2198.72$, $p < .01$; RMSEA = 0.06; CFI = 0.97; AIC = 2518.72; CAIC = 3371.47). The standardized factor loadings ranged between 0.93 (item 8) to 0.38 (item 29). Overall, the factor structure of the SWS looks good; however, item 29 (standardized factor loading = .38) may be problematic and should be monitored moving forward (see Table 18).

The factor correlations of the SWS instrument are presented in Table 19. The following subscales had higher factor correlations: subscale 3 and subscale 4 (.78), subscale 4 and subscale 5 (0.77), and subscale 2 and subscale 3 (0.72). Subscales 1 and 3 have the lowest correlation of 0.32. The measurement errors are presented in Table 23. Although a few of these correlations are high, they are below the .80 level usually used to indicate adequate discriminant validity. Note that these correlations are greater than those

reported in Table 11, because the CFA correlations have been disattenuated for the lack of reliability in the measures.

Measurement Invariance Results

Multigroup CFA was used to examine research question two: (*Is the SWS instrument invariant across both young and middle-aged women in the intended population?*) and its sub-questions on measurement invariance in the SWS instrument:

a) Configural invariance – Do the groups have the same factor structure?
b) Metric invariance – Do the groups have the same factor loadings?
c) Intercept invariance – Do the groups have the same item intercepts?
d) Residual variance invariance – Do the groups have the same item residual variances?
e) Latent mean invariance – Are the latent means invariant across groups?

Configural Invariance. Diagonally weighted least squares (DWLS) estimation was used to examine the configural invariance model as well as the other invariance models mentioned above. The goodness of fit statistics indicated a good fit ($S\text{-}B\chi^2_{(1100)} = 2151.65$, $p < .01$; RMSEA = 0.06; CFI = 0.97; AIC = 2471.65; CAIC = 3324.40) (Table 21).

Metric Invariance. Metric invariance is supported (Table 21). The goodness of fit statistics yielded a good fitting model ($S\text{-}B\chi^2_{(1130)} = 2203.75$, $p < .01$; RMSEA = 0.06; CFI = 0.97; AIC = 2463.75; CAIC = 3156.62) (Table 24). The Satorra-Bentler scaled difference test demonstrated that the metric invariance model fits the data as good as the

configural model, $\Delta SB\chi^2 = 43.76$, $df = 30$. The alternative difference tests results included: ΔRMSEA = 0.0, ΔCFI = 0.0, ΔAIC = 7.90, and ΔCAIC = 167.78 (Table 22).

Intercept Invariance. Intercept invariance test was overall supported, $S\text{-}B\chi^2_{(1160)} = 2296.08$, $p < .01$; RMSEA = 0.06; CFI = 0.97; AIC = 2636.08; CAIC = 3542.14 (Table 21). However, the difference test revealed that the intercept invariance model fit worse than the metric invariance model (Table 22). The Satorra-Bentler scaled chi-square difference test indicated the following: $\Delta SB\chi^2 = 195.25$ and $df = 30$. The alternative difference tests included: ΔRMSEA = 0.0, ΔCFI = 0.0, ΔAIC = 172.33, and ΔCAIC = 385.52 (Table 22). Modification indices were examined for the intercept invariance model to determine which intercept parameters contributed to the lack of fit. The modification indices suggested freely estimating the following items: 11, 14, 21, 22, 25, 27, and 32. Freely estimating the items resulted in a well-fitting partial intercept invariance model ($S\text{-}B\chi^2_{(1160)} = 2237.99$, $p < .01$; RMSEA = 0.06; CFI = 0.97; AIC = 2591.99; CAIC = 3535.35). Also, the difference tests revealed that the partial intercept invariance model fits the data at least good as the metric invariance: $\Delta SB\chi^2 = 24.99$, $df = 23$; ΔRMSEA = 0.0; ΔCFI = 0.0; ΔAIC = 128.24; and ΔCAIC = 378.73. The remaining invariance models are tested for partial invariance freely estimating the same seven items mentioned above.

Residual Variance Invariance. Partial residual variance invariance test is supported. The goodness of fit statistics indicated a good fit ($S\text{-}B\chi^2_{(1195)} = 2240.06$, $p < .01$; RMSEA = 0.06; CFI = 0.97; AIC = 2524.06; CAIC = 3280.88) (Table 21). The Satorra-Bentler scaled chi-square difference between metric invariance and partial

residual variance invariance indicated the partial residual variance invariance model fits good, $\Delta SB\chi^2 = 49.67$, $df = 58$; $\Delta RMSEA = 0.0$; $\Delta CFI = 0.0$; $\Delta AIC = 60.31$; and $\Delta CAIC = 124.26$ (Table 22).

Latent Mean Invariance. Partial latent mean invariance test is supported by the goodness of fit statistics which indicated a good fitting model ($S\text{-}B\chi^2_{(1200)} = 2287.50$; RMSEA = 0.06; CFI = 0.97; AIC = 2531.50; CAIC = 3181.73) (Table 21). However, the Satorra-Bentler scaled chi-square difference test indicated that the partial latent mean invariance model fits worse than the partial intercept model, $\Delta SB\chi 2 = 60.21$, $df = 55$ (Table 22). So, there is reason to believe that the latent means do differ across these age groups. This difference does not indicate a problem of measurement invariance. Instead, this difference suggests population difference on levels of the constructs themselves.

MIMIC Model. The SWS hypothesized model presented in Chapter III was examined using the MIMIC model analysis to answer the following research question: *Does using the MIMIC model lead to conclusions that are similar to those conclusions reach using multigroup CFA?* The goodness of fit statistics revealed a good fit ($S\text{-}B\chi^2_{(580)} = 1688.46$, $p < .01$; RMSEA = 0.06; CFI = 0.97; AIC = 1860.46; CAIC = 2318.81) (see Table 21). The factor loadings are presented in Table 23. These values are close in value compared to the multigroup CFA factor loadings in Table 23. The factor correlations along with age are included in Table 24 and the measurement errors for the MIMIC model are presented in Table 25. The gamma estimated values were all significant except SWS1 (gamma = 0.04, t-value = 1.62) and SWS5 (gamma = 0.01; *t-value* = 0.22) (see

Table 26). Notice, this pattern of significance is the same as was found in the multi-group analysis.

In summary, the measurement invariance models were acceptable for the configural invariance, metric invariance, intercept invariance, and residual invariance after modifying models for partial invariance where needed (see Table 27). The latent mean invariance was not supported indicating that there are some differences between the groups in this study. In Table 26, the standardized mean difference for SWS subscale 3 is fairly large compared to the other subscales. After partial invariance was examined, the mean differences in the multigroup CFA increased which may have caused the inflation of the standardized mean difference for subscale 3. The MIMIC model demonstrated an overall well-fitting model, and differences between the groups' factor means were consistent with the multigroup CFA models (Table 26). The differences revealed between the groups across analyses are discussed in more detail in the next chapter.

Table 18. SWS Parameter Estimates: Multigroup CFA

Parameter	Standardized	Unstandardized	Standard Error
Factor Loadings			
SWS1			
Item 1	0.65	1.00	---
Item 2	0.84	1.00	0.13
Item 3	0.58	0.85	0.12
Item 4	0.71	1.09	0.12
Item 29	0.38	0.82	0.13
Item 35	0.76	1.21	0.18
SWS2			
Item 5	0.55	1.00	---
Item 6	0.69	1.22	0.12
Item 7	0.66	0.91	0.16
Item 8	0.93	1.46	0.16
Item 9	0.75	1.31	0.14
Item 10	0.75	1.30	0.15
Item 30	0.92	1.43	0.16
SWS3			
Item 11	0.77	1.00	---
Item 12	0.67	0.79	0.08
Item 13	0.79	1.04	0.06
Item 14	0.86	1.09	0.06
Item 15	0.78	1.06	0.07
Item 16	0.64	0.95	0.08
Item 31	0.74	0.97	0.08
SWS4			
Item 17	0.49	1.00	---
Item 18	0.68	1.05	0.14
Item 19	0.86	1.42	0.23
Item 20	0.73	1.19	0.18
Item 32	0.56	1.19	0.17
Item 33	0.63	1.01	0.14
SWS5			
Item 21	0.77	1.00	---
Item 22	0.60	0.97	0.06
Item 23	0.75	1.14	0.07
Item 24	0.69	0.98	0.06
Item 25	0.70	1.20	0.08
Item 26	0.46	0.68	0.07
Item 27	0.73	1.05	0.08
Item 28	0.82	1.07	0.08
Item 34	0.78	1.09	0.08

Table 19. SWS Factor Correlation Matrix: Multigroup CFA

	SWS1	SWS2	SWS3	SWS4	SWS5
SWS1	1.00				
SWS2	0.29	1.00			
SWS3	0.34	0.72	1.00		
SWS4	0.59	0.53	0.78	1.00	
SWS5	0.49	0.49	0.62	0.77	1.00

Table 20. SWS Measurement Error: Multigroup CFA

Items	Theta-Delta	Items	Theta-Delta	Items	Theta-Delta
1	0.61	16	0.58	31	0.45
2	0.34	17	0.79	32	0.72
3	0.68	18	0.58	33	0.64
4	0.53	19	0.34	34	0.39
5	0.72	20	0.52	35	0.46
6	0.57	21	0.41		
7	0.60	22	0.63		
8	0.21	23	0.43		
9	0.49	24	0.51		
10	0.49	25	0.50		
11	0.42	26	0.79		
12	0.55	27	0.46		
13	0.37	28	0.32		
14	0.27	29	0.87		
15	0.39	30	0.22		

Table 21. Model Fit Indices

	S-Bχ^2 (*df*)	χ^2	*p*	RMSEA	CFI	AIC	CAIC
Model 1							
DWLS	2151.65 (1100)	4142.16	< .001	0.06	0.97	2471.65	3324.40
Model 2							
DWLS	2203.75 (1130)	4178.83	< .001	0.06	0.97	2463.75	3156.62
Model 3							
DWLS	2296.08 (1160)	4268.49	< .001	0.06	0.97	2636.08	3542.14
Model 3 ***Partial***							
DWLS	2237.99 (1153)	4204.06	< .001	0.06	0.97	2591.99	3535.35
Model 4							
DWLS	2240.06 (1188)	4249.11	< .001	0.06	0.97	2524.06	3280.88
Model 5							
DWLS	2287.50 (1208)	4344.47	< .001	0.06	0.97	2531.50	3181.73
MIMIC							
DWLS	1688.46 (580)	3648.28	< .001	0.06	0.97	1860.46	2318.81

Note. Model 1, 2, 3, & 4 are a part of measurement invariance level. Model 5 is a component of structural level invariance. Model 1, configural invariance; Model 2, metric invariance; Model 3, intercept invariance; Model 4, residual variance invariance; Model 5, latent mean invariance; MIMIC, multiple indicators multiple causes model; DWLS, diagonally weighted least squares estimation.

Table 22. Model Fit Indices Difference Tests

	$\Delta SB\chi^2$ (*df*)	$\Delta\chi^2$	ΔRMSEA	ΔCFI	ΔAIC	ΔCAIC
M2-M1						
DWLS	43.76 (30)	36.67	0.00	0.00	7.90	167.78
M3-M2						
DWLS	195.24 (30)	89.66	0.00	0.00	172.33	385.52
M3-M2Partial						
DWLS	24.99 (23)	25.23	0.00	0.00	128.24	378.73
M4-M2						
DWLS	49.67 (58)	70.28	0.00	0.00	60.31	124.26
M5-M3 DWLS	60.21 (55)	140.41	0.00	0.00	60.49	353.62

Note. M1, configural invariance; M2, metric invariance; M3, intercept invariance; M4, residual variance invariance; M5, latent mean invariance

Table 23. SWS Parameter Estimates: MIMIC Model

Parameter	Standardized	Unstandardized	Standard Error
Factor Loadings			
SWS1			
Item 1	0.65	1.00	---
Item 2	0.76	1.15	0.09
Item 3	0.58	0.89	0.07
Item 4	0.70	1.08	0.08
Item 29	0.43	0.66	0.07
Item 35	0.78	1.19	0.12
SWS2			
Item 5	0.57	1.00	---
Item 6	0.68	1.20	0.08
Item 7	0.62	1.10	0.08
Item 8	0.89	1.56	0.07
Item 9	0.73	1.28	0.05
Item 10	0.74	1.30	0.06
Item 30	0.87	1.53	0.11
SWS3			
Item 11	0.75	1.00	---
Item 12	0.65	0.86	0.05
Item 13	0.79	1.06	0.04
Item 14	0.84	1.12	0.05
Item 15	0.80	1.06	0.04
Item 16	0.68	0.91	0.05
Item 31	0.74	0.99	0.04
SWS4			
Item 17	0.48	1.00	---
Item 18	0.63	1.31	0.20
Item 19	0.80	1.66	0.21
Item 20	0.68	1.42	0.22
Item 32	0.61	1.26	0.27
Item 33	0.59	1.22	0.23
SWS5			
Item 21	0.77	1.00	---
Item 22	0.63	0.83	0.04
Item 23	0.79	1.03	0.04
Item 24	0.71	0.93	0.04
Item 25	0.78	1.01	0.04
Item 26	0.48	0.62	0.05
Item 27	0.73	0.95	0.05
Item 28	0.82	1.07	0.05
Item 34	0.79	1.03	0.04

Table 24. SWS Factor Correlation Matrix: MIMIC Model

	SWS1	SWS2	SWS3	SWS4	SWS5	AGE
SWS1	1.00					
SWS2	0.36	1.00				
SWS3	0.37	0.76	1.00			
SWS4	0.55	0.56	0.71	1.00		
SWS5	0.42	0.58	0.66	0.69	1.00	
AGE	-0.06	-0.20	-0.16	-0.29	-0.01	1.00

Table 25. SWS Measurement Error: MIMIC Model

Items	Theta-Epsilon	Items	Theta-Epsilon	Items	Theta-Epsilon
1	0.57	16	0.53	31	0.45
2	0.43	17	0.77	32	0.63
3	0.66	18	0.61	33	0.66
4	0.50	19	0.36	34	0.38
5	0.68	20	0.54	35	0.40
6	0.54	21	0.41		
7	0.61	22	0.60		
8	0.21	23	0.38		
9	0.47	24	0.49		
10	0.45	25	0.40		
11	0.43	26	0.77		
12	0.58	27	0.47		
13	0.37	28	0.33		
14	0.29	29	0.81		
15	0.36	30	0.24		

*Diagonally Weighted Least Squares Estimation

Table 26. Standardized Mean Differences Across Analyses

Statistical Analysis	Mean Difference	p-value	Std. Mean Difference
T-tests			
SWS1	0.23	0.46	0.06
SWS2	1.61	0.00	0.32
SWS3	1.58	0.00	0.31
SWS4	1.67	0.00	0.47
SWS5	0.21	0.72	0.03
MGCFA			
SWS1	0.02	0.81	0.06
SWS2	0.16	0.01	0.36
SWS3	0.31	0.01	1.00
SWS4	0.16	0.01	0.62
SWS5	0.01	0.94	0.03
MIMIC			
SWS1	0.04	0.11	0.13
SWS2	0.12	0.01	0.27
SWS3	0.12	0.01	0.39
SWS4	0.14	0.01	0.54
SWS5	0.01	0.83	0.03

Table 27. Hypotheses Results of Measurement Invariance

Model	Hypothesis	Hypothesis Test	Hypothesis Name	Model Fit	Difference Test	Description
1	H_{form}	Overall fit	Configural invariance	**Good**	---	The test of whether the 5-factor model of the SWS holds for both groups
2	H_{Λ}	2-1	Metric invariance	**Good**	M2 Fits As **Good** As M1	The strength of the relationship between factors and items are the same for both groups
3	$H_{\Lambda,v}$	3-2	Intercept invariance	**Good**	M3 Fits **Worse** Than M2	The latent means indicated by the items are different for both groups
3a	$H_{\Lambda,v}$	3a-2	Intercept invariance (*Partial*)	**Good**	M3a Fits As **Good** As M2	The latent means indicated by the items are the same for both groups
4	$H_{\Lambda,\Theta(\delta)}$	4-2	Residual variance invariance (*Partial*)	**Good**	M4 Fits As **Good** As M2	The items have the same internal consistency for both groups
5	$H_{\Lambda,v,\ \Theta(\delta),\kappa}$	5-3	Latent mean invariance (*Partial*)	**Good**	M5 Fits **Worse** Than M3a	The latent means are not the same across both groups

CHAPTER V

DISCUSSION

The purpose of this study was to examine the psychometric properties of a newly developed instrument called the Superwoman Schema (SWS) using SEM techniques. The SWS instrument has five subscales with 35 items: subscale 1 (obligation to present an image of strength), subscale 2 (obligation to suppress emotions), subscale 3 (resistance to being vulnerable), subscale 4 (intense motivation to succeed), and subscale 5 (obligation to help others). A multigroup confirmatory factor analysis (CFA) was used to examine measurement invariance in the SWS instrument across two groups: young women (18-39) and middle-aged women (40-65). A multiple indicators multiple causes (MIMIC) model was used to further examine the group differences between the two groups using the SWS instrument. The conclusions from these SEM techniques were compared to see if both the multigroup CFA and the MIMIC model produced similar findings. This chapter provides a brief review of the models examined, the summary of findings, study limitations and recommendations.

Overview of Study

A brief overview of the study analyses are described here and a summary of the findings will follow. There were two types of invariance examined in this study: measurement level invariance and structural level invariance. The measurement level invariance included configural invariance, metric invariance, residual variance invariance

and intercept invariance. The first hierarchical step was the configural invariance. Configural invariance established the equivalence of factor structures for the young women and middle-aged women. The number of factors and loading patterns were the same across groups – no equality constraints were imposed on the parameters (Byrne, 2008).

The second hierarchical step addressed metric invariance, which is focused on the invariant operation of the items on the factor loadings (Byrne, 2008). In this study, one of the group's factor loading parameters were freely estimated, and for the other group the factor loading estimates were constrained equal to those of group one. If the model fit well while factor loading parameters remained constrained equal while simultaneously testing for invariance of additional parameters, remaining hierarchical steps were tested. However, if non-invariance was established in relation to certain factor loadings in this study, partial measurement invariance was used to examine the subsequent invariance tests (Byrne, 2008).

The third hierarchical step addressed intercept invariance. This step provides a researcher the opportunity to test subgroup latent factor mean differences if intercept invariance holds across groups (Byrne, 2008). In this current study, intercept parameters were freely estimated in one group and in another group intercept estimates were constrained equal to group one.

The fourth hierarchical step was the residual variance invariance. The residuals of the regression equations for each item are equivalent across groups in the residual variance invariance model (Schmitt & Kuljanin, 2008). In this study, group one error

variances were freely estimated and group two parameters were constrained to equal group one.

The final hierarchical step was the latent mean invariance. This step required the mean level of the construct to be the same across groups. In this study, one mean was set to 0 and the significance of the other group parameter represented a test of the difference of latent means (Schmitt & Kuljanin, 2008).

A MIMIC model was used to continue analyzing the factor structure of the SWS instrument across the young women and middle-aged women groups in this study. The same hypothesized model used in the multigroup CFA was used for the mimic model. The MIMIC model deals with groups by using a binary indicator. Factors are regressed onto that indicator. The findings of the MIMIC model were compared to the other analyses of the study. The summary of findings for both the multigroup CFA and the MIMIC model are described next.

Summary of Findings

Does the five factor structure proposed by the Superwoman Schema instrument adequately describe survey responses from women in the intended populations? The five factor structure proposed by the SWS instrument did adequately describe the survey responses from the sample of this study. The hypothesized model of the SWS was acceptable according to the overall model fit. The responses from the participants of this study demonstrated that there were five subscales that describe the attributes of women depicting characteristics of a superwoman: an obligation to present an image of strength, an obligation to suppress emotions, resistance to being vulnerable, intense motivation to

succeed, and an obligation to help others. The young women in this study typically scored significantly higher on the SWS subscales compared to the middle-aged women. The young women scored higher on subscales 2 (obligation to suppress emotions), 3 (resistance to being vulnerable), and 4 (intense motivation to succeed). Conversely, the young women and middle-aged women scored similarly on subscale 1 (obligation to present an image of strength) and subscale 5 (obligation to help others) resulting in a non-significant difference across the groups. These findings support the similar findings of mean differences found in approaches to analyses conducted in this study (Table 29). The standardized mean difference for SWS subscale 1 and subscale 5 were not significant in the observed score t-test analyses, multigroup CFA, or mimic model. The young women and middle-aged women endorsed similar characteristics of strength in the SWS subscale 1 such as "I have to be strong," "I feel obligated to present an image at work", or "I have to be strong because I am a woman." Furthermore, the young women and middle-aged women endorsed the items similarly in SWS subscale 5. For example, both groups of women in this study similarly agreed in responses that she takes on too many roles and responsibilities especially when she's already overwhelmed, she neglects her own needs and gratifications, feels guilty when she take out time for herself, and feels obligated to take care of others.

Based on the findings of this current study (no difference between groups on SWS subscale 1 and subscale 5), one can assume that no matter the age women feel obligated or pressured to be a superwoman who takes care of all her needs plus the needs of others. Group differences between the young and middle-aged women across the three SWS

subscales (obligation to suppress emotions, resistance to being vulnerable, and an intense motivation to succeed) suggest several substantive research questions that could be further examined using the SWS as an instrument. Also, researchers could further examine why women neglect their needs and feel obligated to present an image of strength regardless of age differences.

Is the superwoman schema instrument invariant across both young and middle-aged women in the intended population? Overall, the SWS instrument is invariant across both young and middle-aged women in this current study. More specifically, measurement invariance models were examined hierarchically to discover the degree of invariance across groups in the SWS instrument. The configural invariance, metric invariance, intercept invariance, residual variance invariance, and latent mean invariance models were tested in hierarchical fashion.

Configural and metric invariance were fully supported. Intercept invariance was only partially supported. That is, some item means differed between groups by more than would be predicted by differences on the latent variables. Modification indices were examined to determine which parameters contributed to the lack of fit in the full intercept invariance model. The modification indices suggested freely estimating the following parameters: item 11 (*It's hard for me to accept help from others*), item 14 (*Asking for help is difficult for me*), item 21 (*I take on roles and responsibilities when I am already overwhelmed*), item 22 (*I take on too many responsibilities*), item 25 (*When others ask for help, I say yes when I should say no*), item 27 (*I neglect the things that bring me joy*), and item 32 (*The only way for me to be successful is to work hard*). Freely estimating

these parameters resulted in a well-fitting partial intercept invariance model compared to metric invariance. Moving forward, the rest of the invariant models were examined using partial invariance freeing the 7 items mentioned above. Further studies might include follow-up interviews to better understand how women in these age groups interpret the meaning of these items. Next, partial residual variance invariance testing revealed that the groups in this study had the same item residual variances after freeing the same 7 items in the partial intercept invariance. Lastly, the partial latent mean invariance was not supported. Partial latent mean invariance demonstrated that the young and middle-aged women have different means across groups which was expected considering the significant differences in the analyses for this study in SWS subscales 2 through subscale 4 (Table 29).

Does using a MIMIC model lead to conclusions that are similar to those conclusions reached using a multigroup CFA? The mimic model did lead to conclusions similar to those conclusions reached in the multigroup CFA. This is shown in Table 29 where the mean differences (gammas) reveal that SWS subscale 2, subscale 3, and subscale 4 were significantly different in this study populations. Like mentioned above, SWS subscale 1 and subscale 5 was not significantly different in the mimic model as well as the other analyses in Table 29. The results from the MIMIC model further supports that the SWS instrument can be used in future studies to examine the differences in young and middle-aged women because the results from the MIMIC model are similar to the multigroup CFA with minor violations of the item intercepts. The findings of this study can broadly be interpreted as showing that researchers may use the SWS instrument with young and

middle-aged women samples without having too much concern about the differences in the quality of construct measurement, even if those studies use simple linear sum scores (observed score analysis results were also consistent).

Overall, the two groups were similar in their responses to the SWS instrument across the study analyses. The young women and the middle-aged women were different in responses for SWS subscale 2, subscale 3, and subscale 4. And both groups were similar in responses for SWS subscale 1 and subscale 5. The young women scored higher on the SWS subscale 2, subscale 3, and subscale 4 compared to the middle-aged women. One possibility from this current study is that young women may feel more obligated to suppress their emotions (subscale 2), resist vulnerability (subscale 3), and succeed despite having challenges in life (subscale 4) compared to the middle-aged women. For example, young women scored higher obligation to suppress emotions subscale (“I display my emotions in privacy”), resistance to being vulnerable subscale (“I wait until I am overwhelmed to ask for help"), and intense motivation to succeed subscale (“It is very important for me to be the best at the things I do”).

Based on the study findings, African American women of various ages are similar in feeling obligated to present an image of strength (subscale 1) and obligated to help others (subscale 5). Both groups of women scored higher on obligation to present an image of strength subscale and obligation to help others subscale compared to the other subscales. The young and middle-aged women in this study are quite similar when it comes to taking care of others particularly their families and neglecting to take care of their own

needs. So, it's not surprising that both groups were not different when it came to feeling obligated to present an image of strength and help others.

"The only way for me to be successful is to work hard" is an intense motivation to succeed subscale item (32), which revealed the biggest difference (0.44) between the young women and middle-aged women in the study. One can assume that young women are in the stage of their lives where building a career and life is most important for them compared to middle-aged women who are either retired or have established a career for themselves (Carstensen, 1995; Charles et al., 2001; Phillips et al., 2006). Middle-aged women have different views about life and are more comfortable in accepting who they are no matter what anyone says about them (Carstensen,1995; Charles & Carstensen, 2007; Freund & Baltes, 2002). Instead, young women are at the stage where learning and building careers for themselves are a priority, which corresponds to the socioemotional selectivity theory discussed in the chapter two (Carstensen, 1995; Carstensen et. al, 1999). Based on the socioemotional selectivity theory, the young women in this study are more than likely in the knowledge-related goals stage which reference individuals in this stage seeking to gain knowledge, develop new social relationships and build careers (Carstensen, 1995; Carstensen et. al; Charles et al., 2001). The middle-aged women in this current study could resemble characteristics of the emotion-related goals stage where they are most concerned about intimate relationships with close family and friends. This research focused on the factor structure and not necessarily the substantive interpretation of the SWS. Although assumptions of the findings are briefly discussed, the results do

seem to provide clear avenues for further substantive research on the constructs of interest.

Study Limitations

There were a number of missing values in the data for the variable of interest, age. The missing values were not used in the analyses of this study (listwise deletion). There may be better ways of handling missing data in the future when there are fewer other conditions to consider. Originally, the sample of data collected included 674 women. With the number of cases missing for age, the sample reduced from 674 women to 561 women. The ability to generalize to the population may be limited because majority of the sample included women from mostly the southeastern region of the United States even though small numbers (29%) of women were represented from the western region of the United States.

Another limitation is due to one of the data collection methods: online survey administration. Majority of the study's sample was administered the SWS online. Most of the missing values for age came from the online administration of the SWS instrument. Participants who took the SWS online may have dropped out of the study due to a number of reasons such as test fatigue, computer and/or technology issues, etc. The mode of survey administration in this study could have contributed to confounding effects of the study results as well. Additional analyses are needed to further examine how the mode of survey administration could affect the findings of a study.

Partial invariance leads to consequences of using an instrument to compare groups because the violations of invariance and the size of the violations are unknown if

the violations are not assessed. The impact of partial invariance can be evaluated on accuracy of selection on the basis of a composite of the instrument whose factor structure is being studied (Millsap & Kwok, 2004). The accuracy of selection is beyond the scope of this study (see Millsap & Kwok, 2004 for more details). When using a partial invariance model, a researcher must be cognizant in how he or she uses the instrument. Once partial invariance is used, a researcher has an option to use the intact instrument, modify the instrument, or not use the instrument in comparison across populations. Ultimately, the researcher must rely on professional judgment when using an instrument that demonstrates partial invariance. The main key is to make sure the instrument is used as its intended purpose. Despite the limitations mentioned above, this study extends to the growing literature on superwoman characteristics and how it affects the lives of women in their professional and personal life which sometimes results in stress and health issues.

Recommendations

Future studies on the SWS are recommended to examine other populations to continue analyzing the measurement invariance of the instrument, possibly examining women over the age of 65 to see if the SWS instrument factorial structure functions in similar ways as the young and middle-aged women in this current study. Further analyses of various group differences would contribute to the use of the SWS instrument across different race and religion groups, working and non-working women, or college students (undergraduate vs. graduate or full-time vs. part-time). Also, item response theory (IRT) could be used to continue analyzing the quality of the SWS items in depth.

In terms of validity research, the SWS instrument could be compared to other related instruments such as the Strong Black Woman Scale (SBWS) and the Strong Black Women Cultural Construct Scale. This study is the first to test population heterogeneity in the SWS instrument and it is with great hopes this research expands the discussion on the application of the superwoman in research.

This paper contributes to the ongoing discussion on superwoman and provides the opportunity for the SWS instrument to be used by other researchers with confidence knowing that the SWS instrument measures the characteristics of a superwoman using five subscales and that although some item intercepts differ, the subscales are measured similarly across the studied age groups. The superwoman concept continues to be a popular term used in mainstream media such as newspapers, magazines, social media, and even in pop culture music. The word superwoman has changed the way women think about themselves and how they play a role in society. The superwoman term has evolved over years and it will be interesting to see how the term will continue to impact the lives of women across the country.

REFERENCES

Akaike, Hirotugu (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19 (6): 716–723 doi:10.1109/TAC.1974.1100705.

Almeida, D. M., & Horn, M. C. (2004). Is daily life more stressful during middle adulthood? In O. G. Brim, C. D. Ryff & R. C. Kessler (Eds.), *How healthy are we? A national study of well-being at midlife* (pp. 425–451). Chicago: University of Chicago Press.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. *Standards for educational and psychological tests*. Washington, D.C.: American Psychological Association, 1990.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. *Standards for educational and psychological tests*. Washington, D.C.: American Psychological Association, 1999.

Barton, R. M., Andrew, M. D., & Schwab, R. L. (1994). Factorial validity and reliability of a survey to assess the teaching effectiveness of graduates of teacher education programs. *Educational and Psychological Measurement*, 54: 218-226.

Beauboeuf-Lafontant, T. (2007). You have to show strength: An exploration of gender, race and depression. *Gender & Society*, 21(1), 28-51.

Bradely, K. L., Bagnell, A. L., & Brannen, C. L. (2010). Factorial validity of the Center for Epidemiological Studies Depression 10 in adolescents. *Issues in Mental Health Nursing*, 31:408-412.

Black, A. & Peacock, N. (2011). Pleasing the masses: Messages for daily life management in African American women's popular media sources. *American Journal of Public Health*, 101(1), 144-150.

Bollen, K. A. (1989). A new incremental fit index for general structural models. *Sociological Methods & Research*, 17, 303-316.

Browne, M. W. (1984). Asymptotic distribution free methods in analysis of covariance structures. *British Journal of Mathematical and Statistical Psychology*, 37, 62-83.

Browne, F. B., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds), *Testing Structural Equation Models*. (pp. 136-162). Newbury Park, CA: Sage.

Byrne, B. M. (1994). Testing for the factorial validity, replication, and invariance of a measuring instrument: A paradigmatic application based on the Maslach Burnout Inventory. Multivariate Behavioral Research, 29(3), 289-311.

Byrne, B. M. (1998). *Structural equation modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications, and programming*. Mahwah, NJ: Erlbaum.

Byrne, B. M. (2008). Testing for multigroup equivalence of a measuring instrument: A walk through the process. Psicothema, 20(4), 872-88.

Byrne, B. M. (2010). Testing for measurement and structural equivalence in large-scale cross-cultural studies: Addressing the issue of nonequivalence. *International Journal of Testing*, 10, 107-132.

Byrne, B. M. & Campbell, T. L. (1999). Cross-cultural comparisons and the presumption of equivalent measurement and theoretical structure: A look beneath the surface. *Journal of Cross-Cultural Psychology*, 30, 555-574.

Byrne, B. M., Shavelson, R., & Muthén, B. O. (1989). Testing for equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin, 105*, 456-466.

Carstensen, L. L. (1995). Evidence for a life-span theory of socioemotional selectivity. Current Directions in Psychological Science, 4 (5), 151–156.

Carstensen, L. L., Isaacowitz, D. M., & Charles, S. T. (1999). Taking time seriously: A theory of socioemotional selectivity. *American Psychologist*, 54 (3), 165–181.

Charles, S. T., & Carstensen, L. L. (2007). Emotion regulation and aging. In J. J. Gross (Ed.), *Handbook of Emotion Regulation*. New York: Guilford Press.

Charles, S.T., Reynolds, C.A., Gatz, M. Age-related differences and change in positive and negative affect over 23 years. *Journal of Personality and Social Psychology*, Vol 80(1), 136-151.

Crago, M., Yates, A., Fleischer, C., Segerstorm, B., & Gray, N. (1996). The superwoman ideal and other risk factors for eating disturbance in adolescent girls. *Sex Roles*, 35, 801-810.

Crawford, J. R., & Henry, J. D. (2004). The Positive and Negative Affect Schedule (PANAS): Construct validity, measurement properties and normative data in a large non-clinical sample. *British Journal of Clinical Psychology, 43,* 245-265.

Charles, S. T., Reynolds, C. A., & Gatz, M. (2001). Age-related differences and change in positive and negative affect over 23 years. *Journal of Personality and Social Psychology, 80,* 136–151.

Cheung, G. W. (2007). Testing equivalence in the structure, means, and variances of higher-order constructs with Structural Equation Modeling. *Organizational Research Methods,* 11, 593-614.

Cheung, G.W. & Rensvold, R. B. (2002). Evaluating Goodness-of-Fit Indexes for Testing Measurement Invariance. Structural Equation Modeling, 9, 233-255.

Cohen, A., Kim, S. & Wollack, J. (1996). An investigation of the likelihood ratio test for detection of differential item functioning. *Applied Psychological Measurement,* 20(15), 15-26.

Cronbach, L. J. (1971). Test validation. In R. L. Thorndike (Ed.), Educational measurement, 2nd ed. (pp. 443–507). Washington, DC: American Council on Education.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. Psychological Bulletin, 52, 281–302.

Finch, H. (2005). The MIMIC model as a method for detecting DIF: Comparison with Mantel-Haenszel, SIBTEST, and the IRT Likelihood Ratio. *Applied Psychological Measurement,* 29, 278-295.

Flora, D. B. & Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. Psychological Methods, 9 (4), 466-491.

Freund, A. M., & Baltes, P. B. (2002). Life-management strategies of selection, optimization and compensation: Measurement by self-report and construct validity. *Journal of Personality and Social Psychology,* 82, 642–662.

Frisbie, D. (2005). Measurement 101: Some fundamentals revisited. Educational Measurement: Issues and Practice, 24, 21-28.

Hambelton, R. & Swaminathan, H. (1985). *Item Response Theory: Principles & Applications.* Boston: Kluwer Publishing.

Hamilton-Mason, J., Hall, J. C., & Everett, J. (2009). And some of us are braver: Stress and coping among African American women. *Journal of Human Behavior in the Social Environment*, 19, 463-482.

Hamin, D. A. (2008). Strong Black Woman cultural construct: Revision and validation. *Dissertation Abstracts International*, 71(4). Retrieved October 22, 2010, from ProQuest Digital Dissertations database. (Publication No. AAT 3405845).

Hart, K. & Kenny, M. (1997). Adherence to the super woman ideal and eating disorder symptoms among college women. *Sex Roles*, 36(7/8), 461-478.

Herrera, R. S. & DelCampo, R. L. (1995). Beyond the superwoman syndrome: Work satistification and family functioning among working-class, Mexican American Women. *Hispanic Journal of Behavioral Sciences*, 17(1), 49-60.

Hayduk, L. A. (1987). *Structural equation modeling with LISREL: Essentials and advances*. Baltimore: John Hopkins University Press.

Hoyle, R. H. (2012). *Handbook of Structural Equation Modeling*. New York, NY: The Guilford Press.

Hu, L., & Benlter, P. (1995). Evaluating model fit. In R. H. Hoyle (Ed.), *Structural equation modeling: Issues, concepts, and applications* (pp. 76-99). Newbury Park, CA: Sage.

Hu, L., Bentler, P. M., & Kano, Y. (1992). Can test statistics in covariance structure analysis be trusted? *Psychological Bulletin*, 112, 351-362.

Hull, D. M., Beaujean, A. A., Worell, F. C., & Verdisco, A. E. (2010). An item-level examination of the factorial validity of NEO Five-Factor Inventory scores. *Educational and Psychological Measurement*, 70(6), 1021-1041.

IBM Corp. Released (2011). IBM SPSS Statistics for Windows, Version 20.0. Armonk, NY: IBM Corp.

Jacques, H. (2008). Young women constructing identities: Multiple subject positions and the superwoman ideal. *Dissertation Abstracts International*, 69(12). Retrieved October 22, 2010, from ProQuest Digital Dissertations database. (Publication No. AAT NR44356).

Jöreskog, K. G. (1990). New developments in LISREL: Analysis of ordinal variables using polychoric correlations and weighted least squares. Quality & Quantity, 24, 387-404.

Jöreskog, K. G. (1994). On the estimation of polychoric correlations and their asymptotic covariance matrix. *Psychometrika,59*, 381-389.

Jöreskog, K. G., & Goldberger, A. S. (1975). Estimation of a model with multiple indicators and multiple causes of a single latent variable. *Journal of the American Statistical Association*, 70, 631-639.

Jöreskog, K.G. & Sörbom, D. (1993). *LISREL 8*: *User's reference guide*. Chicago, IL: Scientific Software International, Inc.

Jöreskog, K.G. & Sörbom, D. (2006a). *LISREL 8.8 for Windows* [Computer software]. Lincolnwood, IL: Scientific Software International, Inc.

Jöreskog, K.G. & Sörbom, D. (2006b). *LISREL 8.8*: *User's reference guide*. Chicago, IL: Scientific Software International, Inc.

Kane, M. (1992). An argument-based approach to validation. Psychological Bulletin, 112, 527–535.

Kane, M. (2006). Validation. In R. Brennan (Ed.), Educational measurement, 4th ed. (pp. 17–64),Westport, CT: American Council on Education and Praeger.

Kane, M. (2009). Validating the interpretations and uses of test scores. In R. Lissitz (Ed.), The Concept of Validity: Revisions, New Directions, and Applications (pp. 39-64). Charlotte, NC: Information Age Publishing, Inc.

Kane, M. (2012). Validating score interpretations and uses: Messick Lecture, Language Testing Research Colloquium, Cambridge, April 2010. Language Testing, 29(1), 3-17.

Kanfer, R. & Ackerman, P. (2004). Aging, adult development and motivation. *Academy of Management Review*, 29 (3), 440-458.

Kaplan, D. (2009). Structural equation modeling: Foundations and extensions. Thousand Oaks, California: SAGE Publications, Inc.

Kirby, J.B. & Bollen, K.A. (2009). Using instrumental variables to evaluate model specification in latent variable structural equation models. Sociologcial Methdology, 39, 327-355.

Kline, R. (2005). *Principles and practice of structural equation modeling*. New York, NY: The Guildford Press.

Lissitz, R. W. & Samuelsen, K. (2007). A suggested change in terminology and emphasis regarding validity and education. Educational Researcher, 36, 437- 448.

Little, T. D. (1997). Mean and covariance structures (MACS) analyses of cross-cultural data: Practical and theoretical issues. *Multivariate Behavioral Research, 32,* 53-76.

Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Hillside, NJ: Erlbaum.

MacCallum, R. C. & Austin, J. T. (2000). Applications of Structural Equation Modeling in psychological research. Annual Review of Psychology, 51, 201-226.

Marsh, H. W. (1990). Confirmatory Factor Analysis of multitrait-multimethod data: The construct validation of multidimensional self-concept response. *Journal of Personality*, 58(4), 661-692.

Meade, A. W. & Lautenschlager, G. J. (2004, April). *Same Question, Different Answers: CFA and Two IRT Approaches to Measurement Invariance.* Symposium presented at the 19th Annual Conference of the Society for Industrial and Organizational Psychology, Chicago, IL.

Meredith,W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58,* 525-543.

Messick, S. (1989). Validity. In R. Linn (Ed.), *Educational measurement* (3rd ed). Washington, DC: American Council on Education/Macmillan.

Messick, S. (1995). Standards of validity and the validity of standards in performance assessment. *Educational Measurement: Issues and Practice,* 14, 5-8.

Mullings, L. (2006). Resistance and resilience: The Sojourner Syndrome and the social context of reproduction in Central Harlem. In A. J. Schulz & L. Mullings (Eds.), *Gender, race, class, and health* (pp. 345-370). San Francisco: Jossey-Bass.

Muthén, B. (1989). Latent variable modeling in heterogeneous populations. *Psychometrika*, 54(4), 557-585.

Phillips, L. H., Henry, J. D., Hosie, J. A., & Milne, A. B. (2006). Age, anger regulation and well-being. *Aging and Mental Health*, 10, 250–256.

Raju, N. (1988). The area between two item characteristic curves. *Psychometrika*, 53, 495–502.

Raju, N.(1990). Determining the significance of estimated signed and unsigned areas between two item response functions. *Applied Psychological Measurement*, 14, 197–207.

Raju, N., Laffitte, L., & Byrne, B. (2002). Measurement equivalence: A comparison of methods based on confirmatory factor analysis and item response theory. *Journal of Applied Psychology*, 87(3), 517-529.

Raju, N., van der Linden,W., & Fleer, P. (1995). An IRT-based internal measure of test bias with applications for differential item functioning. *Applied Psychological Measurement*, 19, 353–368.

Rauschenbach, C. & Hertel, G. (2011). Age difference in strain and emotional reactivity to stressors in professional careers. *Stress and Health*, 27, 48-60.

Reise, S., Widaman, K., & Pugh, R. (1993). Confirmatory factor analysis and item response theory: Two approaches for exploring measurement invariance. *Psychological Bulletin*, 114(3), 552-566.

Reynolds, M. R., Keith, T. Z., Ridley, K. P., & Patel, P. G. (2008). Sex differences in latent general and broad cognitive abilities for children and youth: Evidence from higher-order MG-MACS and MIMIC models. *Intelligence*, 36, 236-260.

Rivera, P., & Satorra, A. (2002). Analysing group differences: A comparison of SEM approaches. In G. A. Marcoulides & I. Moustaki (Eds.), *Latent variable and latent structure models* (pp. 86-104). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Romero, R. E. (2000). The icon of the strong Black woman: The paradox of strength. In L. C. Jackson & B. Greene (Eds.), *Psychotherapy with African American women: Innovations in psychodynamic perspectives and practice* (pp. 225-238). New York: Guilford Press.

Rosén, M.(1995). Gender differences in structure, means and variances of hierarchically ordered ability dimensions. *Learning and Instruction*, 5, 37-62.

Rubio, D. M., Berg-Wegerb, M., Tebb, S. S., & Rauchc, S. M. (2005). Validating a measure across groups. *Journal of Social Service Research*, 29, 53-67.

Santrock, J. (2009). *Topical Life-Span Development* (4th ed.). New York: McGraw-Hill.

Satorra, A., & Bentler, P. M. (1988). Scaling corrections for chi-square statistics in covariance structure analysis. *ASA 1988 Proceedings of the Business and*

Eeonomic Statistics ,Section (308-313). Alexandria, VA: American Statistical Association.

Satorra, A., & Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. von Eye & C. C. Clogg (Eds.), *Latent variable analysis: Applications to developmental research* (pp. 399-419). Thousand Oaks, CA: Sage.

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, 66, 507-514.

Schickedanz, J. A., Schickedanz, D. I., Forsyth, P. D., Forsyth, G. A. (2001). *Understanding Children and Adolescents*. Needham Heights, MA: Allyn & Bacon.

Schmitt, N. & Kuljanin,G. (2008). Measurement invariance: Review of practice and implications. *Human Resource Management Review*, 18, 210-222.

Settles, I. H., Pratt-Hyatt, J. S., & Buchanan, N. T. (2008). Through the lens of race: Black and White women's perceptions of womanhood. *Psychology of Women Quarterly*, 32, 454-468.

Sireci, S. (2007). On validity theory and test validation. *Educational Researcher,* 36(8), 477-481.

Sireci, S. & Parker, P. (2006). Validity on Trial: Psychometric and Legal Conceptualizations of Validity. *Educational Measurement: Issues and Practices*, 25 (3), 27-34.

Stark, S., Chernyshenko, O., & Drasgow, F. (2006). Detecting differential item functioning with confirmatory factor analysis. *Journal of Applied Psychology*, 91(6), 1292-1306.

Thissen, D., Steinberg, L., & Wainer, H. (1988). Use of item response theory in the study of group differences in trace lines. In H. Wainer & H. I. Braun (Eds.), *Test validity* (pp. 147–169). Hillsdale, NJ: Erlbaum.

Thomas, A. J., Witherspoon, K. M., & Speight, S. L. (2004). Toward the development of the stereotypic roles for Black woman scale. *Journal of Black Psychology,* 30(3), 426-442.

Thompson, P. C. (2003). Strong Black Woman Scale: Construction and validation. *Dissertation Abstracts International,* 64(7), 3545B. Retrieved October 22, 2010, from ProQuest Digital Dissertations database. (Publication No. AAT 3097501).

Thompson, M. S., & Green, S. B. (2006). Evaluating between-group differences in latent variable means. In G. R. Hancock, & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (pp. 119−170). Greenwich, CT: Information Age.

Utsey, S., Giesbrecht, N., Hook, J., & Stanard, P. (2008). Cultural, sociofamilial, and psychological resources that inhibit psychological distress in African Americans Exposed to stressful life events and race-related stress. *Journal of Counseling Psychology*, 55, 49-62.

Vandenberg, R. J. (2011). Toward a further understanding of and improvement in measurement invariance methods and procedures. *Organizational Research Methods*, 25, 139-158.

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods,* 3, 4-69.

Wallace, M. (1990). *Black macho and the myth of the superwoman*. London: Verso.

Wang, W. & Shih, C. (2010). MIMIC methods for assessing differential item functioning in polytomous items. *Applied Psychological Measurement,* 34(3) 166-180.

Wicherts, J. M. & Dolan, C.V. (2010). Measurement invariance in Confirmatory Factor Analysis: An illustration using IQ test performance on minorities. *Educational Measurement: Issues and Practice*, 29(3), 39-47.

Wirth, R. J. & Edwards, M. C. (2007). Item factor analysis: Current approaches and future directions. Psychological Methods, 12(1), 58-79

Woods, C. (2009). Evaluation of MIMIC-Model methods for DIF testing with comparison to two-group analysis. *Multivariate Behavioral Research*, 44, 1-27.

Woods-Giscombé, C. L. (2010). Superwoman schema: African American women's views on stress, strength, and health. *Qualitative Health Research*, 20(5), 668-683.

Wu, P. (2010). Measurement invariance and latent mean differences of the Beck Depression Inventory II across gender groups. *Journal of Psychoeducational Assessment*, 28, 551-563.

Yuan, K., & Bentler, P. M. (1998). Normal theory based test statistics in structural equation modeling. *British Journal of Mathematical and Statistical Psychology*, *51*, 289-309.

APPENDIX A

SUPERWOMAN SCHEMA INSTRUMENT

Participant ID#_____

INSTRUCTIONS: The following is a list of items that may or may not be relevant for you. Some of the questions may sound similar, but each is important. Please read and complete each item to the best of your ability using the response scale provided.

		If you checked **TRUE**, please indicate how undesirable or disturbing this statement is for you by checking one of the boxes below.
1. I try to present an image of strength.	☐ This is **NOT TRUE** for me ☐ This is **TRUE** for me rarely ☐ This is **TRUE** for me sometimes ☐ This is **TRUE** for me all of the time	**This bothers me:** ❑ Very much ❑ Somewhat ❑ Not at all
2. I have to be strong.	☐ This is **NOT TRUE** for me ☐ This is **TRUE** for me rarely ☐ This is **TRUE** for me sometimes ☐ This is **TRUE** for me all of the time	**This bothers me:** ❑ Very much ❑ Somewhat ❑ Not at all
3. I feel obligated to present an image of strength at work	☐ This is **NOT TRUE** for me ☐ This is **TRUE** for me rarely ☐ This is **TRUE** for me sometimes ☐ This is **TRUE** for me all of the time	**This bothers me:** ❑ Very much ❑ Somewhat ❑ Not at all
4. I feel obligated to present an image for my family	☐ This is **NOT TRUE** for me ☐ This is **TRUE** for me rarely ☐ This is **TRUE** for me sometimes ☐ This is **TRUE** for me all of the time	**This bothers me:** ❑ Very much ❑ Somewhat ❑ Not at all

5. I display my emotions in privacy.	☐ This is **NOT TRUE** for me ☐ This is **TRUE** for me rarely ☐ This is **TRUE** for me sometimes ☐ This is **TRUE** for me all of the time	**This bothers me:** ❑ Very much ❑ Somewhat ❑ Not at all
6. I keep my feelings to myself.	☐ This is **NOT TRUE** for me ☐ This is **TRUE** for me rarely ☐ This is **TRUE** for me sometimes ☐ This is **TRUE** for me all of the time	**This bothers me:** ❑ Very much ❑ Somewhat ❑ Not at all
7. I keep my feelings to myself.	☐ This is **NOT TRUE** for me ☐ This is **TRUE** for me rarely ☐ This is **TRUE** for me sometimes ☐ This is **TRUE** for me all of the time	**This bothers me:** ❑ Very much ❑ Somewhat ❑ Not at all
8. My tears are a sign of weakness.	☐ This is **NOT TRUE** for me ☐ This is **TRUE** for me rarely ☐ This is **TRUE** for me sometimes ☐ This is **TRUE** for me all of the time	**This bothers me:** ❑ Very much ❑ Somewhat ❑ Not at all
9. I keep my problems bottle up inside	☐ This is **NOT TRUE** for me ☐ This is **TRUE** for me rarely ☐ This is **TRUE** for me sometimes ☐ This is **TRUE** for me all of the time	**This bothers me:** ❑ Very much ❑ Somewhat ❑ Not at all
10. Expressing emotions is difficult for me.	☐ This is **NOT TRUE** for me ☐ This is **TRUE** for me rarely ☐ This is **TRUE** for me sometimes ☐ This is **TRUE** for me all of the time	**This bothers me:** ❑ Very much ❑ Somewhat ❑ Not at all
11. It's hard for me to accept help form others.	☐ This is **NOT TRUE** for me ☐ This is **TRUE** for me rarely ☐ This is **TRUE** for me sometimes ☐ This is **TRUE** for me all of the time	**This bothers me:** ❑ Very much ❑ Somewhat ❑ Not at all
12. I have a hard time trusting others.	☐ This is **NOT TRUE** for me ☐ This is **TRUE** for me rarely ☐ This is **TRUE** for me sometimes	**This bothers me:** ❑ Very much ❑ Somewhat

	☐ This is **TRUE** for me all of the time	❑ Not at all
13. I wait until I am overwhelmed to ask for help.	☐ This is **NOT TRUE** for me ☐ This is **TRUE** for me rarely ☐ This is **TRUE** for me sometimes ☐ This is **TRUE** for me all of the time	**This bothers me:** ❑ Very much ❑ Somewhat ❑ Not at all
14. Asking for help is difficult for me.	☐ This is **NOT TRUE** for me ☐ This is **TRUE** for me rarely ☐ This is **TRUE** for me sometimes ☐ This is **TRUE** for me all of the time	**This bothers me:** ❑ Very much ❑ Somewhat ❑ Not at all
15. I resist help to prove that I can make it on my own.	☐ This is **NOT TRUE** for me ☐ This is **TRUE** for me rarely ☐ This is **TRUE** for me sometimes ☐ This is **TRUE** for me all of the time	**This bothers me:** ❑ Very much ❑ Somewhat ❑ Not at all
16. If I want things done right, I do them myself.	☐ This is **NOT TRUE** for me ☐ This is **TRUE** for me rarely ☐ This is **TRUE** for me sometimes ☐ This is **TRUE** for me all of the time	**This bothers me:** ❑ Very much ❑ Somewhat ❑ Not at all
17. I accomplish my goals with limited resources.	☐ This is **NOT TRUE** for me ☐ This is **TRUE** for me rarely ☐ This is **TRUE** for me sometimes ☐ This is **TRUE** for me all of the time	**This bothers me:** ❑ Very much ❑ Somewhat ❑ Not at all
18. It is very important to me to be the best at the things that I do.	☐ This is **NOT TRUE** for me ☐ This is **TRUE** for me rarely ☐ This is **TRUE** for me sometimes ☐ This is **TRUE** for me all of the time	**This bothers me:** ❑ Very much ❑ Somewhat ❑ Not at all
19. No matter how hard I work, I feel like I should do more.	☐ This is **NOT TRUE** for me ☐ This is **TRUE** for me rarely ☐ This is **TRUE** for me sometimes ☐ This is **TRUE** for me all of the time	**This bothers me:** ❑ Very much ❑ Somewhat ❑ Not at all
20. I put pressure on myself to achieve a certain level of	☐ This is **NOT TRUE** for me ☐ This is **TRUE** for me rarely ☐ This is **TRUE** for me sometimes	**This bothers me:** ❑ Very much

accomplishment.	☐ This is **TRUE** for me all of the time	❑ Somewhat ❑ Not at all
21. I take on roles and responsibilities when I am already overwhelmed.	☐ This is **NOT TRUE** for me ☐ This is **TRUE** for me rarely ☐ This is **TRUE** for me sometimes ☐ This is **TRUE** for me all of the time	**This bothers me:** ❑ Very much ❑ Somewhat ❑ Not at all
22. I take on too many responsibilities in my family.	☐ This is **NOT TRUE** for me ☐ This is **TRUE** for me rarely ☐ This is **TRUE** for me sometimes ☐ This is **TRUE** for me all of the time	**This bothers me:** ❑ Very much ❑ Somewhat ❑ Not at all
23. I put everyone else's needs before mine.	☐ This is **NOT TRUE** for me ☐ This is **TRUE** for me rarely ☐ This is **TRUE** for me sometimes ☐ This is **TRUE** for me all of the time	**This bothers me:** ❑ Very much ❑ Somewhat ❑ Not at all
24. I feel obligated to take care of others.	☐ This is **NOT TRUE** for me ☐ This is **TRUE** for me rarely ☐ This is **TRUE** for me sometimes ☐ This is **TRUE** for me all of the time	**This bothers me:** ❑ Very much ❑ Somewhat ❑ Not at all
25. When others ask for my help, I say yes when I should say no.	☐ This is **NOT TRUE** for me ☐ This is **TRUE** for me rarely ☐ This is **TRUE** for me sometimes ☐ This is **TRUE** for me all of the time	**This bothers me:** ❑ Very much ❑ Somewhat ❑ Not at all
26. I neglect my health. In what specific ways do you think that you neglect your health? _____	☐ This is **NOT TRUE** for me ☐ This is **TRUE** for me rarely ☐ This is **TRUE** for me sometimes ☐ This is **TRUE** for me all of the time	**This bothers me:** ❑ Very much ❑ Somewhat ❑ Not at all
27. I neglect the things that bring me joy.	☐ This is **NOT TRUE** for me ☐ This is **TRUE** for me rarely ☐ This is **TRUE** for me sometimes ☐ This is **TRUE** for me all of the time	**This bothers me:** ❑ Very much ❑ Somewhat ❑ Not at all
28. I feel guilty when I	☐ This is **NOT TRUE** for me ☐ This is **TRUE** for me rarely	**This bothers me:**

take time for myself.	☐ This is **TRUE** for me sometimes ☐ This is **TRUE** for me all of the time	❑ Very much ❑ Somewhat ❑ Not at all
29. The struggles of my ancestors require me to be strong.	☐ This is **NOT TRUE** for me ☐ This is **TRUE** for me rarely ☐ This is **TRUE** for me sometimes ☐ This is **TRUE** for me all of the time	**This bothers me:** ❑ Very much ❑ Somewhat ❑ Not at all
30. I keep my problems to myself to prevent from burdening others.	☐ This is **NOT TRUE** for me ☐ This is **TRUE** for me rarely ☐ This is **TRUE** for me sometimes ☐ This is **TRUE** for me all of the time	**This bothers me:** ❑ Very much ❑ Somewhat ❑ Not at all
31. I do things by myself without asking for help.	☐ This is **NOT TRUE** for me ☐ This is **TRUE** for me rarely ☐ This is **TRUE** for me sometimes ☐ This is **TRUE** for me all of the time	**This bothers me:** ❑ Very much ❑ Somewhat ❑ Not at all
32. The only way for me to be successful is to work hard.	☐ This is **NOT TRUE** for me ☐ This is **TRUE** for me rarely ☐ This is **TRUE** for me sometimes ☐ This is **TRUE** for me all of the time	**This bothers me:** ❑ Very much ❑ Somewhat ❑ Not at all
33. I am a perfectionist.	☐ This is **NOT TRUE** for me ☐ This is **TRUE** for me rarely ☐ This is **TRUE** for me sometimes ☐ This is **TRUE** for me all of the time	**This bothers me:** ❑ Very much ❑ Somewhat ❑ Not at all
34. There is no time for me because I am always taking care of others.	☐ This is **NOT TRUE** for me ☐ This is **TRUE** for me rarely ☐ This is **TRUE** for me sometimes ☐ This is **TRUE** for me all of the time	**This bothers me:** ❑ Very much ❑ Somewhat ❑ Not at all
35. I have to be strong because I am a woman.	☐ This is **NOT TRUE** for me ☐ This is **TRUE** for me rarely ☐ This is **TRUE** for me sometimes ☐ This is **TRUE** for me all of the time	**This bothers me:** ❑ Very much ❑ Somewhat ❑ Not at all

Participant Background Information

Participant ID # _____

Please answer each question and circle the number corresponding to the answer that fits you the best.

Age _____

How do *you* describe your race/ ethnicity? CHECK ALL THAT APPLY.

- ☐ African-American or Black
- ☐ West Indian or Caribbean (Which country/ territory: _____)
- ☐ Native American
- ☐ Latino (Which country/territory:_____)
- ☐ African (Which country:_____)
- ☐ Other (PLEASE SPECIFY:_____)

Where did you grow up? CHECK ONE.

- ☐ United States (List Specific State(s):_____)
- ☐ West Indies/ Caribbean (Which country/territory:_____)
- ☐ Africa (Which country:_____)
- ☐ Other (PLEASE SPECIFY:_____)

With what religion or religious denomination do you *most* identify?

_____

Which of the following best describes your current employment status?

- ☐ Working full-time for wages
- ☐ Working part-time for wages
- ☐ Not working but looking for a job
- ☐ Retired
- ☐ Disabled
- ☐ Homemaker (Keeping house or raising children full-time)
- ☐ Student
- ☐ Military or armed forces
- ☐ Not working and not looking for a job

SKIP NEXT QUESTION IF YOU ANSWERED 3, 4, 5, 6, 7, or 9 TO PREVIOUS QUESTION:

If you are currently working, what is your current job or occupation? PLEASE PROVIDE AS MUCH DETAIL AS POSSIBLE.

_____

SKIP NEXT QUESTION IF ANSWERED 1 OR 2 TO PREVIOUS QUESTION:

If you are unemployed, how long have you been unemployed?

_____

Are you currently a student? CHECK ONE.

☐ No
☐ Yes, part-time
☐ Yes, full-time

What is the highest level of education you have completed?
☐ No formal education
☐ Some elementary or grade school but did not graduate
☐ Elementary or Grade School
☐ Some middle school but did not graduate
☐ Graduated Middle School
☐ Some Junior High School but did not graduate
☐ Graduated Junior High School
☐ GED or High School Equivalent
☐ High School Diploma
☐ Some college but did not graduate
☐ Associate's Degree or Community College degree
☐ Bachelor's Degree
☐ Graduate Degree

What is the highest level of education your mother (the mother that raised you) completed?
☐ No formal education
☐ Some elementary or grade school but did not graduate
☐ Elementary or Grade School
☐ Some middle school but did not graduate
☐ Graduated Middle School
☐ Some Junior High School but did not graduate
☐ Graduated Junior High School
☐ GED or High School Equivalent

☐ High School Diploma
☐ Some college but did not graduate
☐ Associate's Degree or Community College degree
☐ Bachelor's Degree
☐ Graduate Degree

What is the highest level of education your father (the father that raised you) completed?
☐ No formal education
☐ Some elementary or grade school but did not graduate
☐ Elementary or Grade School
☐ Some middle school but did not graduate
☐ Graduated Middle School
☐ Some Junior High School but did not graduate
☐ Graduated Junior High School
☐ GED or High School Equivalent
☐ High School Diploma
☐ Some college but did not graduate
☐ Associate's Degree or Community College degree
☐ Bachelor's Degree
☐ Graduate Degree

Which of the following best describes your current marital or relationship status?
☐ Married
☐ Not married but living with a romantic partner
☐ Married but separated
☐ Divorced
☐ Widowed
☐ Single/never married
☐ In a romantic relationship, but not living together

Do you currently have health insurance?
☐ Yes
☐ No

SKIP NEXT QUESTION IF ANSWERED NO TO PREVIOUS QUESTION:

What kind of health insurance do you currently have?
☐ Health insurance from an employer or former employer
☐ Health insurance that you purchase directly
☐ Health insurance from a government program, such as Medicare, Medicaid (MediCal), CHAMPUS, or VA
☐ Other private health insurance
☐ Other public health insurance

Do you have a health care provider you usually go to for check-ups and other health care needs?
☐ No
☐ Yes

Have you been to the doctor for a check-up in the past year?
☐ No
☐ Yes

How many children do you have? _____

Are you a grandparent currently raising your grandchildren?
☐ No
☐ Yes

Are you raising foster children or someone else's children who are NOT your grandchildren?
☐ No
☐ Yes

How many people including you live in your household? (enter numerical value)
_____

How many children under the age of 18 live with you? (enter numerical value) _____

How many adults ages 18 and over live with you? (enter numerical value) _____

What city and state do you currently live in?
City: _____
State: _____

How long have you lived at your current residence?
☐ Less than one year
☐ 1-2 years
☐ 3-4 years
☐ 5 years or more

GENERAL HEALTH

Now, we're going to ask you some questions about your general health. Again, please take your time and answer these questions to the best of your ability.

During the past month, how would you rate your sleep quality overall?
☐ Very good
☐ Fairly good
☐ Fairly bad
☐ Very bad

How would you rate your overall physical health at the present time?
☐ Excellent
☐ Very Good
☐ Good
☐ Fair
☐ Poor

How would you rate your overall mental health at the present time?
☐ Excellent
☐ Very Good
☐ Good
☐ Fair
☐ Poor

Below is a list of health conditions. For each one, please indicate whether a doctor or health professional has ever told you
that you have that problem.

	No	**Yes**
1. Arthritis or Rheumatism		
2. Chronic back or neck problems		
3. Frequent or severe headaches		
4. Any other chronic pain		
5. Stroke		
6. Heart attack		
7. High blood pressure or hypertension		
8. Blood circulation problems such as atherosclerosis or "hardening of the arteries"		
9. Angina		
10. Other heart condition or heart disease		
11. Asthma		
12. Tuberculosis (TB)		
13. Chronic lung disease, like emphysema or COPD		
14. Any other respiratory or breathing problem		
15. Diabetes or high blood sugar		
16. Ulcer in stomach or intestines		

17. Epilepsy or seizures		
18. HIV/AIDS		
19. Cancer		
20. Liver problem or liver trouble		
21. Kidney problem or kidney trouble		
22. Kidney disease		
23. Glaucoma		
24. Osteoporosis		
25. Serious hearing problem		
26. Serious vision problem		
27. Anemia		
28. Sickle-cell disease		
29. Serious allergies or infections		
30. Fibroid tumors		

Have you ever taken medication for any of the following: high blood pressure, hypertension, stroke, heart attack, blood circulation problems, hardening of the arteries, or any other heart or blood problem?

☐ No
☐ Yes

Have you ever taken medication or under gone dialysis for kidney disease?

☐ No
☐ Yes

ADDITIONAL QUESTIONS

What are your own personal earnings, including only your own personal wages every month after taxes?

☐ Less than $500/month	☐ $5001 to $5500/month
☐ $501 to $1000/month	☐ $5501 to $6000/month
☐ $1001 to $1500/month	☐ $6001 to $6500/month
☐ $1501 to $2000/month	☐ $6501 to $7000/month
☐ $2001 to $2500/month	☐ $7001 to $7500/month
☐ $2501 to $3000/month	☐ $7501 to $8000/month
☐ $3001 to $3500/month	☐ $8001 to $8500/month
☐ $3501 to $4000/month	☐ $8501 to $9000/month
☐ $4001 to $4500/month	☐ $9001 to $9500/month
☐ $4501 to $5000/month	☐ More than $9501/month

What is the total income of all the people who live in your household after taxes? This includes all money from wages, social security, retirement, and other benefits,

government assistance programs, and all other sources of income, such as pensions, investments, child support, and alimony for all the people who live in your household, including yourself.

☐ Less than $500/month	☐ $5001 to $5500/month
☐ $501 to $1000/month	☐ $5501 to $6000/month
☐ $1001 to $1500/month	☐ $6001 to $6500/month
☐ $1501 to $2000/month	☐ $6501 to $7000/month
☐ $2001 to $2500/month	☐ $7001 to $7500/month
☐ $2501 to $3000/month	☐ $7501 to $8000/month
☐ $3001 to $3500/month	☐ $8001 to $8500/month
☐ $3501 to $4000/month	☐ $8501 to $9000/month
☐ $4001 to $4500/month	☐ $9001 to $9500/month
☐ $4501 to $5000/month	☐ More than $9501/month

How many people does this income support? _____

How difficult is it for you to meet the monthly payments on your bills?

- ☐ Extremely difficult
- ☐ Very difficult
- ☐ Somewhat difficult
- ☐ Slightly difficult
- ☐ Not difficult at all

How much do you worry that your total income will not be enough to meet your expenses and bills?

- ☐ A great deal
- ☐ A lot
- ☐ A little
- ☐ Not at all

I am the first in my family to attain the level of educational or professional success that I have achieved?

- ☐ This is not true for me
- ☐ This is true for me

I take care of food and expenses for my family members with little or no help from others?

- ☐ This is not true for me
- ☐ This is true for me

I am raising/I raised my child(ren) as a single mother while working.

☐ This is not true for me
☐ This is true for me

I am raising/I raised my child(ren) as a single mother while going to school.
☐ This is not true for me
☐ This is true for me

I am in school despite having family members besides my own children to support.
☐ I do not have the This is not true for me
☐ This is true for me

I do not have the financial support from others to accomplish my goals (e.g., go to school, start a business, buy a house).
☐ This is not true for me
☐ This is true for me

I am responsible for the care of elderly relatives or other family members (other than my own children).
☐ This is not true for me
☐ This is true for me

I am a member of groups or organizations that require my time and commitment
☐ This is not true for me
☐ This is true for me

If so, please list how many organizations._____

Please list how many hours per month you are involved with duties related to this organization._____

Has there been any recent event that has influenced your current level of stress? If so, please explain below.

_____
_____
_____
_____
_____

Thank you for your time and your interest in participating in this study!

APPENDIX B

INSTITUTIONAL REVIEW BOARD

INITIAL INSTITUTIONAL REVIEW BOARD (IRB)

OFFICE OF RESEARCH COMPLIANCE
2718 Beverly Cooper Moore and Irene Mitchell Moore
Humanities and Research Administration Bldg. PO Box 26170
Greensboro, NC 27402-6170
336.256.1482
Web site: www.uncg.edu/orc
Federalwide Assurance (FWA) #216

To: John Willse
Ed Research Methodology
246 School of Education Building

From: UNCG IRB

_____
Authorized signature on behalf of IRB

Approval Date: 5/15/2012
Expiration Date of Approval: 5/15/2013

RE: Notice of IRB Approval by Expedited Review (under 45 CFR 46.110)
Submission Type: Initial
Expedited Category: 7.Surveys/interviews/focus groups
Study #: 12-0170
Study Title: Superwoman Schema Survey Research Study

This submission has been approved by the IRB for the period indicated. It has been determined that the risk involved in this research is no more than minimal.

Study Description:

The purpose of this research is to provide a reliable and valid survey instrument to measure superwoman characteristics in women.
Investigator's Responsibilities

Federal regulations require that all research be reviewed at least annually. It is the Principal Investigator's responsibility to submit for renewal and obtain approval before

the expiration date. You may not continue any research activity beyond the expiration date without IRB approval. Failure to receive approval for continuation before the expiration date will result in automatic termination of the approval for this study on the expiration date.

Signed letters, along with stamped copies of consent forms and other recruitment materials will be scanned to you in a separate email. These consent forms must be used unless the IRB hasgiven you approval to waive this requirement.

You are required to obtain IRB approval for any changes to any aspect of this study before they can be implemented (use the modification application available at http://www.uncg.edu/orc/irb.htm). Should any adverse event or unanticipated problem involving risks to subjects or others occur it must be reported immediately to the IRB using the "Unanticipated Problem/Event" form at the same website.

CC:
Teneka Steed
ORC, (ORC), Non-IRB Review Contact

RENEWAL IRB

OFFICE OF RESEARCH COMPLIANCE
2718 Beverly Cooper Moore and Irene Mitchell Moore
Humanities and Research Administration Bldg. PO Box 26170
Greensboro, NC 27402-6170
336.256.1482
Web site: www.uncg.edu/orc
Federalwide Assurance (FWA) #216

To: John Willse
Ed Research Methodology
246 School of Education Building

From: UNCG IRB

_____
Authorized signature on behalf of IRB

Approval Date: 4/03/2013
Expiration Date of Approval: 4/02/2014

RE: Notice of IRB Approval by Expedited Review (under 45 CFR 46.110)
Submission Type: Renewal
Expedited Category: 7.Surveys/interviews/focus groups
Study #: 12-0170

Study Title: Superwoman Schema Survey Research Study

This submission has been approved by the IRB for the period indicated.

Study Description:

The purpose of this research is to provide a reliable and valid survey instrument to measure superwoman characteristics in women.

Regulatory and other findings:

- This research is closed to enrollment and remains open for data analysis only.

Investigator's Responsibilities

Federal regulations require that all research be reviewed at least annually. It is the Principal Investigator's responsibility to submit for renewal and obtain approval before the expiration date. You may not continue any research activity beyond the expiration date without IRB approval. Failure to receive approval for continuation before the expiration date will result in automatic termination of the approval for this study on the expiration date.

Signed letters, along with stamped copies of consent forms and other recruitment materials will be scanned to you in a separate email. These consent forms must be used unless the IRB hasgiven you approval to waive this requirement.

You are required to obtain IRB approval for any changes to any aspect of this study before they can be implemented (use the modification application available at http://www.uncg.edu/orc/irb.htm). Should any adverse event or unanticipated problem involving risks to subjects or others occur it must be reported immediately to the IRB using the "Unanticipated Problem/Event" form at the same website.

CC:
Teneka Steed

APPENDIX C

CONSENT FORM

UNIVERSITY OF NORTH CAROLINA AT GREENSBORO

CONSENT TO ACT AS A HUMAN PARTICIPANT

Project Title: Superwoman Schema Research Study

Principal Investigator: John Willse, PhD
UNC-Greensboro Department: Educational Research Methodology
UNC-Greensboro Phone Number: (336) 334-3435
Email Address: jtwillse@uncg.edu.

Co-Principal Investigator: Cheryl Giscombé, PhD
UNC-Chapel Hill Department: School of Nursing
UNC-Chapel Hill Phone Number: (919) 452-0116
Email Address: cheryl.giscombe@unc.edu

Student Researcher: Teneka Steed
UNC-Greensboro Department: Educational Research Methodology
UNC-Greensboro Phone Number: (980) 475-1718
Email Address: tcsteed@uncg.edu

Participant's Name: _____

What is the study about?
This is a research project. You are being asked to take part in a research study. To join the study is voluntary. You may refuse to join, or you may withdraw your consent to be in the study, for any reason, without penalty.

Stress is known to influence both psychological and physical health. It is known that stress specifically affects the health of Black women. This research study is important, because it will guide us toward decreasing or preventing the negative effect that stress has on the health of Black women. The purpose of this research study is to learn more about how Black, African, and African women experience and cope with stress and to develop a questionnaire that will assist in assessing stress.

Why are you asking me?
You are being asked to participate in this research study to help our research team develop a survey about how Black, African, and African American women experience and cope with stress in their lives. Also, you are asked to participate in this research study if you consider yourself a Black, African, or African American woman who is at least 18 years old.

What will you ask me to do if I agree to be in the study?
If you take part in this study, you will be asked to complete a questionnaire packet to help our research team understand more about how Black, African, and African American women experience and cope with stress in their lives.

Once the study is completed, the research team would like to obtain your feedback on what we found about how African American women experience and cope with stress. A summary of what was learned from the questionnaires will be sent to you via postal mail. You will be invited to communicate feedback to the research team in writing or by calling us directly. A stamped envelope will be included in the mailing so that you can write down comments for the investigator. If you would prefer to speak to the investigator by telephone, her phone number will be included in the mailing. The time required for your participation will be dependent on how you choose to participate ranging from 15 minutes to approximately 1 hour.

What are the dangers to me?
The Institutional Review Board at the University of North Carolina at Greensboro has determined that participation in this study poses minimal risk to participants. The only foreseeable direct risk from participating in this study involves the chance that thinking about the topics that are included in the questionnaire may raise issues that you may find mildly distressing or you may feel uncomfortable about disclosing how you experience and cope with stress.

The co-principal investigator, Cheryl Giscombé, has a background in psychology and mental health nursing and is sensitive about addressing these issues. You may discuss any distressing issues that arise with the research staff at any time while you are completing the questionnaire or contact Cheryl Giscombé later by phone, (919) 452-0116. A referral list for psychological services will be made available to participants who report or appear to be in serious psychological distress.

Questions, concerns or complaints about this project or benefits or risks associated with being in this study can be answered by John Willse who may be contacted at (336) 334-3435 or by email jtwillse@uncg.edu**.**

What are the possible benefits from being in this study?
Research is designed to benefit society by gaining new knowledge. You may not benefit personally from being in this research study. However, it is possible that you may receive personal satisfaction in knowing that your participation will assist researchers in understanding more about how women experience and cope with stress. This information may eventually lead researchers to understand more about how to help women prevent the negative effect of stress on health outcomes in African American women.

Will I get paid for being in the study? Will it cost me anything?
You will be entered into a drawing for a $30.00 gift card for your participation in this study. A total of four gift cards will be awarded. There are no costs to you or payments made for participating in this study

How will you keep my information confidential?
The surveys will be identified with code numbers, not names. All data will be kept secured in a locked file cabinet and on a password protected computer located in a locked office off UNCGs

campus for at least three years. Only research staff will have access to this data.
Participants *will not* be identified in any report or publication about this study. All information obtained in this study is strictly confidential unless disclosure is required by law. Although every effort will be made to keep research records private, there may be times when federal or state law requires the disclosure of such records, including personal information. This is very unlikely, but if disclosure is ever required, UNC-Greensboro will take steps allowable by law to protect the privacy of personal information. In some cases, your information in this research study could be reviewed by representatives of the University, research sponsors, or government agencies for purposes such as quality control or safety**.**

What if I want to leave the study?
You have the right to refuse to participate or to withdraw at any time, without penalty. If you do withdraw, it will not affect you in any way. If you choose to withdraw, you may request that any of your data which has been collected be destroyed unless it is in a de-identifiable state.
What about new information/changes in the study?
If significant new information relating to the study becomes available which may relate to your willingness to continue to participate, this information will be provided to you.

What if you have questions about your rights as a research participant?
All research on human volunteers is reviewed by a committee that works to protect your rights and welfare.
If you have any concerns about your rights, how you are being treated or if you have questions, want more information or have suggestions, please contact Eric Allen in the Office of Research Compliance at UNCG toll-free at (855)-251-2351.

Voluntary Consent by Participant:
By signing this consent form you are agreeing that you read, or it has been read to you, and you fully understand the contents of this document and are openly willing consent to take part in this study. All of your questions concerning this study have been answered. By signing this form, you are agreeing that you are 18 years of age or older and are agreeing to participate, or have the individual specified above as a participant participate, in this study described to you by the research team.

Signature: _____ Date: _____

Print Name: _____

APPENDIX D

RECRUITMENT FLYER

Superwoman Schema Survey Research Study

Purpose of Study: This research study aims to evaluate a survey to help researchers better understand stress and coping in African American women.

Participant Eligibility: Women of African descendent (e.g., African Americans, Africans, Blacks) who are at least 18 years old are eligible to participate in this research study.

Study Procedures: A survey will be administered to eligible participants, which will take about 20 minutes to complete.

Compensation: Eligible participants will be entered in a $30 drawing for participation in the research study.

Location: The survey will be administered online or by paper in a public setting convenient for the participant.

For more information regarding this study, please contact:

Teneka Steed (336) 419-7738 tcsteed@uncg.edu
Cheryl Giscombé (919) 452-0116 cheryl.giscombe@unc.edu

This research is conducted under the direction of John Willse, PhD jtwillse@uncg.edu

University of North Carolina at Greensboro