Gender Invariance of Behavior and Symptom Identification Scale Factor Structure

Author: Thomaskutty B. Idiculla

Persistent link: http://hdl.handle.net/2345/1815

This work is posted on eScholarship@BC, Boston College University Libraries.

Boston College Electronic Thesis or Dissertation, 2008

Copyright is held by the author, with all rights reserved, unless otherwise noted.
GENDER INVARIANCE OF BEHAVIOR AND SYMPTOM IDENTIFICATION SCALE

FACTOR STRUCTURE

A dissertation

By

THOMASKUTTY IDICULLA

Submitted in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

December, 2008
Gender Invariance of Behavior and Symptom Identification Scale Factor Structure

Thomaskutty Idiculla, Ph.D.

Thomas O’Hare, Ph.D., Doctoral Committee Chair

Abstract

The Behavior and Symptom Identification Scale 24 (BASIS-24) is a psychiatric outcome measure used for inpatient and outpatient populations. This 24-item measure comprises six subscales: depression/functioning; interpersonal relationships; self-harm; emotional lability; psychosis; and substance abuse. Earlier studies examined the reliability and validity of the BASIS-24, but none empirically examined its factor structure across gender. The purpose of this study was therefore to assess the construct validity of the BASIS-24 six-factor model and find evidence of configural, metric, strong and strict factorial invariance across gender. The sample consisted of 1398 psychiatric inpatients that completed BASIS-24 at admission and discharge at 11 facilities nation-wide. Confirmatory factor analyses were used to test measurement invariance of the BASIS-24 six-factor model across males and females.

The single confirmatory factor analysis showed the original six-factor model of BASIS-24 provided an acceptable fit to the male sample at admission (RMSEA=0.058, SRMR=0.070, CFI=0.975, NNFI=0.971 and GFI=0.977) and at discharge (RMSEA=0.059, SRMR=0.078, CFI=0.977, NNFI=0.972, and GFI=0.969). The goodness-of-fit indices for the female group at admission (RMSEA=0.055, SRMR=0.067, CFI=0.980, NNFI=0.976, and GFI=0.983), and at discharge (RMSEA=0.055, SRMR=0.079, CFI=0.98, NNFI=0.977, and GFI=0.971) also revealed that the six factor model fit reasonably well to the data. The goodness–of–fit indices between the unconstrained and constrained models showed that all four multi-group models were
equivalent for both male and female samples at admission and discharge in terms of goodness-of-fit examined through the ΔCFI and that all of them show an acceptable fit to the data. The decrease in CFI was <0.008 for admission sample and <0.003 for discharge sample and both fell below the 0.01 cut-off. This indicates that the configural, metric, as well as the strong and strict factorial invariance of BASIS-24 exist across males and females.

The two important contributions of the present study are: 1) BASIS-24 can be used as a reliable and valid symptom measurement tool in assessing psychiatric inpatient populations which can compare quantitative differences in the magnitude of patient symptoms and functioning across genders; 2) the current study provides an example of useful statistical methodology for examining specific questions related to factorial invariance of the BASIS-24 instrument across gender. Implications of social work practice and research are discussed.
Acknowledgments

I would like to express my gratitude to Dr. Thomas O'Hare, the Chairperson of my committee, for his continued support and guidance over the years of my work and decisions. His positive, and often humorous, outlook on life has been a constant source of inspiration during this project. I would also express my thanks to other committee members, Dr. Ce Shen, Dr. Paul Kline, and Dr. Susan Eisen for their patience, willingness to see me through to the completion of this project. I am thankful to Dr. Ce Shen for his statistical expertise and generous availability. I am thankful to Dr. Kline for his instructive comments. I would like to thank Dr. Eisen whose research and commitment for patient centered evaluation sparked an interest in me and provided me with the sample for this study.

My family has been through all of this time with me. I extend a special thanks to my Dad and Mom, Chacko and Aleyama Idiculla, for their constant encouragement for my higher studies. I thank my loving wife, Suja, for her patience, support, and encouragement throughout my doctoral studies. I am also thankful to my four children, Miriam, Steve, Jeremy, and Aaron. I love them with my all my heart and I thank them for their love and understanding when I needed to study.

During my doctoral program, I was faced with several challenges, but I thank God for the strength, motivation, and intelligence He has given me to finish this project successfully.
Table of Contents

CHAPTER I: INTRODUCTION

Purpose of the Study

Gender Difference in Mental Health Assessment and Care
  Biological Differences and Gender
  Socio-economic Status and Gender
  Gender Differences in Help-Seeking Behavior
  Culture and Gender
  Health Care Provider Bias and Gender

Gender and Mental Health Theories

Research in Support of Mental Health Theories on Gender

Psychometric Properties of BASIS-24
  Depression
  Self-Harm
  Substance Abuse
  Psychosis
  Interpersonal Relationships
  Emotional Lability

Measurement Equivalence
  Factor Invariance
  Configural Factor Invariance (Pattern invariance)
  Metric factor invariance
  Strong factorial (scalar or item) invariance
  Strict factorial (complete or item residual) invariance

Patient Assessment and Evidence Based Practice in Social Work

CHAPTER II: METHODOLOGY

Sample

Study Design

Measures
  Behavior and Symptom Identification Scale (BASIS-24)
  Demographic Characteristics

Data Collection Procedure
List of Tables and Figures

Tables

Table 1: Study sample by gender (paired for admission and discharge) ........................................... 33
Table 2: Clinical and Demographic Characteristics ............................................................................ 49
Table 3: Descriptive Statistics of subscales for BASIS-24 by Gender .............................................. 53
Table 4: One-way analysis of variance of BASIS-24 Subscale and Overall Score by Gender ....... 55
Table 5: Standardized Internal Consistency Reliability (Cronbach’s Alpha) Coefficients .......... 56
Table 6: BASIS-24 Overall and Subscale Correlations at Admission................................................ 57
Table 7: BASIS-24 Overall and Subscale Correlations at Discharge................................................. 58
Table 8: Single group confirmatory factor analysis of the BASIS-24 by Gender ............................ 59
Table 9: Multi Group Confirmatory Factor Analysis by Gender ....................................................... 63

Figures

Figure 1: A conceptual six-factor model for BASIS-24 ................................................................. 38
CHAPTER I: INTRODUCTION

The value of assessing mental health treatment outcomes and quality of care has become more relevant to clinicians, patients and payers over the past two decades. Quality is a notion that is often discussed but difficult to operationalize. A shift of substantial portions of the population into integrated managed care organizations has contributed to a greater emphasis on accountability, quality, and outcomes of evidence-based care. In response to the demand for quality of care information, mental health professionals have pursued a wide range of initiatives designed to measure treatment outcomes in order to assess the performance of mental health delivery systems. As well as the evidence-based model of outcomes, the patient-centered care model has been increasingly important in clinical practice. In this context, self-reported measures based on patient functioning and symptoms are considered important in assessing outcomes and regarded as key indicators of quality of care. While the phrase patient functioning and symptoms is found frequently in mental health research literature, there has been relatively little exploration of the constructs of functioning and symptoms across gender in combination with existing gender differences.

A major challenge in conducting assessments in diverse populations is the possibility that measures developed for a given construct in one particular group may not be assessing the same construct in the same manner in other groups (Van de Vijver, F. & Leung, 1997). Given the growing minority population, as well as findings of sex-differences, cross-gender research is relevant to clinicians and social science academia. Inquiry regarding appropriate and effective methodology for cross-gender study is therefore important. Snowden (Snowden, 1996) proposes that confirmation of a general factor structure in diverse samples strengthens an instrument’s
validity as well as the likelihood that there is a common underlying theoretical framework for the interpretation of items (Chow, Snowden, & McConnell, 2001).

Purpose of the Study

BASIS-24 is a widely used behavioral health assessment tool. BASIS-24 is designed for use by mental health providers, researchers, purchasers of mental health services, accreditation agencies, and internal quality assurance departments to measure the change in consumer self-reported symptom and problem difficulty over the course of treatment. BASIS–24 is copyrighted by McLean Hospital and requires a site license to use. Development of BASIS-24 has been described in detail in previous publications and here a brief profile of the tool is included (S. V. Eisen, Gerena, Ranganathan, & Idiculla, 2006). In this context, the purposes of the present study were twofold: (a) to test the factorial structure of BASIS-24 separately for males and females; and b) to test for measurement invariance across gender.

The first objective of this study investigated the following questions: are there differences between the mean, standard deviation, skewness, and kurtosis for male and female BASIS-24 scores; does BASIS-24 have adequate internal consistency across gender; are there differences between the internal consistency reliability of BASIS-24 for men and women; will BASIS-24 subscales and items significantly correlate with their overall BASIS-24 scores for both men and women; and will item-total and subscale total correlations of BASIS-24 be similar across gender.

The second objective was to examine the factor structure and gender invariance of BASIS-24 scale using confirmatory factory analysis. More specifically, to examine whether the male and female datasets revealed the following six factors: (a) Depression/Functioning, (b) Interpersonal, (c) Self-Harm, (d) Emotional Lability, (e) Psychosis, and (f) Substance Abuse. In this regard, following four hypotheses were tested.
1. The factor structure of BASIS-24 is invariant across gender groups. Both male and female groups associate the same subscales of BASIS-24 items.

2. The factor loadings of BASIS-24 are invariant across gender. The strength of the relationship based on the factor loadings between each item and its underlying construct is the same for both groups.

3. The cross group differences and relationships among constructs (covariances) are the same across males and females.

4. BASIS-24 items have an invariant pattern of unique variances or residual or error variances across gender. In other words, BASIS-24 items reflect the same quality as measures of the underlying construct.

The analysis will further test the null hypotheses that there are no significant differences between male and female test results as such, and that all measurable differences are indeed gender differences, instead of differences due to the measure’s unsuitability to measure characteristics across gender; and there are no significant differences between the test results for each gender group included in the study per se and all significant differences are a function of gender differences rather than the measure’s psychometric properties. If the above hypotheses, and thus previous research, are confirmed, the analysis would prove and establish measurement equivalence between BASIS-24 for gender and other psychometric measures. This study is organized into the introduction, review of literature related to gender and mental health, BASIS-24 subscales, theoretical framework, research findings, presentation of psychometric properties of BASIS-24, method of study including the research questions, the sample and study design, the instrumentation, results and discussion.
Gender Difference in Mental Health Assessment and Care

Although existing psychometric measures do not generally distinguish between the gender groups, several measures have been found to be sensitive to gender differences, such as the Hopkins Symptom Checklist (HSC-46) (Casper, Belanoff, & Offer, 1996). Gender differences can result in different recommendations for men and women in terms of treatment. In fact, past research has shown that women have different needs regarding mental health care, as certain problems affect female patients more than or differently from men. Such issues include, but are not limited to, rape, domestic violence, and childhood sexual abuse. Generally and globally speaking, women are at a higher risk of experiencing sexual violence in childhood, adolescence or adulthood (World Health Organization, 2002). The fact that many women suffer from pre-menstrual syndrome involving mood swings potentially results in more diagnoses of depression by mental health care professionals in women than in men. Other life-events exclusively affect women: after pregnancy, women have been shown to have an increased risk for psychological distress, which can manifest itself as post-partum depression or post-partum psychosis, both of which are well documented in the literature. Ante-natal and post-natal depression has been associated with poor financial and social support as well as marital disharmony (World Health Organization, 2002).

Biological Differences and Gender

It has been reported that hormonal differences between men and women result in a different response to medical/psychiatric drugs (Ramsay, Welch, & Youard, 2001). For a long time, the differences in the side-effects of psychiatric medication as well as hormonal differences have not been considered in the medical/psychiatric treatment of women, which points to the fact that women have been underrepresented in both medical and clinical research (Kohen, 2001).
Women have also been found to be at an increased risk for affective and neurotic disorders after undergoing hysterectomy or diagnosis of infertility (World Health Organization, 2002). It has been said that extensive research has been done on women’s reproductive functioning and its effects on mental health, while there is insufficient literature on the reproductive biology and mental health of men (World Health Organization, 2002).

**Socio-economic Status and Gender**

Women are also at a socioeconomic disadvantage putting them at a higher risk for mental illness. For instance, H. Cooper, Arber, & Fee (H. Cooper, Arber, & Fee, 1999) showed that the social environment had a greater impact on women’s health than on men’s. Due to the double burden on working mothers, a lot of women have entered the workforce on a part-time basis (Ramsay et al., 2001). Although paid work provides a source of income and social status constitutes a protective factor with regard to mental health, working part-time often prevents women from receiving benefits and fully developing their careers with the prospect of holding higher positions and achieving better job security (Gosling, Johnson, & McCrae, J. et al., 1997).

**Gender Differences in Help-Seeking Behavior**

Women appear to be more aware of having a mental disorder than men when experiencing comparable levels of distress (World Health Organization, 2002), and might therefore contact mental health care providers at an earlier stage of perceived emotional distress. For instance, females were found to contact outpatient facilities earlier more frequently than men (World Health Organization, 2002), pointing to a difference in help-seeking behavior which sets females apart from male help-seeking behavior. Ramsay et al. (Ramsay et al., 2001) mention the possibility that men and women may have a different way of expressing distress, and that women
address mental health issues to their PCP, whereas men tend to address these issues after referral to a specialist (Goldberg & Huxley, 1992). Further, women might be reluctant to contact mental health care providers out of fear that their children will be taken away from them (Ramsay et al., 2001).

Culture and Gender

Culture and gender play important roles in influencing mental health (Katsurada & Sugihara, 1999). Mental health constructs can be viewed and defined differently from different cultures as culture shapes people’s beliefs, attitudes and behavior through its norms, customs, and socialization practices. Similarly, gender is an important variable that is influenced by culture and it affects and influence mental health. However, a major issue in the use of psychological measures is the possible variance resulting from differences across different groups such as gender. The surgeon general's report (U.S. Department of Health and Human Services, 2006) emphasized the importance of measurement invariance to mental health services and research.

Health Care Provider Bias and Gender

Women are also more frequently diagnosed with depression as a consequence of a gender-bias on the part of health care providers. According to studies from the US and Germany, elderly women were diagnosed with depression more often than elderly men (World Health Organization, 2002), although both groups were presenting with the same severity of symptoms (World Health Organization, 2002). Potts et al. (1991) (as cited in Ramsay et al., (Ramsay et al., 2001)) found evidence of gender bias in their study, as women were more often diagnosed with depression than men. Nash and Chrisler (Nash & Chrisler, 1997) examined the likelihood that their participants would “diagnose” women they knew with Pre-Menstrual
Dysphoric Disorder (PMDD) after reading the diagnostic criteria for it and found that participants were more likely to see PMDD in other women they knew than the control group, which was assigned to read the diagnostic criteria for Episodic Dysphoric Disorder (EDD), equivalent to those of PMDD without the menstrual-cycle references. They found evidence of a gender bias in the diagnosis of psychiatric disorders, which attaches a negative label to the female menstrual cycle. However, gender bias might also be due to the fact that factors contributing to psychiatric symptoms are not taken into account by the patient, the respective measure, or the mental health professional. Klonoff, Landrine, and Campbell (Klonoff, Landrine, & Campbell, 2000) refer to research by Myers (Myers et al., 2002) who found that stressors specific to women (such as experiencing sexist treatment or being the target of sexually derogatory language) are responsible for the higher rate of psychiatric symptoms in women that has been reported in the literature.

The measures currently available to assess mental illness do not generally discriminate male from female patients, and might therefore not be adequately equipped to both detect and serve patients with gender-specific needs. In fact, there is evidence that certain measures are not suitable screening instruments in a female population, such as the CAGE questionnaire on problem drinking and alcohol dependence (O'Hare & Tran, 1997). Adler et al. (1990) (as cited in Ramsay et al., 2001 (Ramsay et al., 2001)) conducted a vignette study of personality disorder diagnoses and found that men were more likely to be diagnosed with narcissistic personality disorder, while women were diagnosed more often with histrionic personality disorder, irrespective of the clinician’s gender. This could point to genuine gender differences or the fact that some of the measures used to assess personality disorder were not sensitive to gender differences. Nash and Chrisler (Nash & Chrisler, 1997) showed that, even when biological...
differences are taken into account, women are still more likely to be diagnosed with a psychiatric disorder, as the available measures are likely to be a reflection of our cultural beliefs and practices.

Gender and Mental Health Theories

Social science research in the field of gender and mental health has implemented widely accepted theories that try to explain how gender roles render men and women unequally vulnerable to psychological disorders. Some of these theories will be outlined in the following, along with research findings lending support to the presented theories. Further, the measurement equivalence and psychometric properties of BASIS-24 will be discussed in relation to gender.

Gender theories try to explain how men and women form their sexual identities based on the biological differences of their bodies as well as prevailing societal gender norms, and how they relate to each other in society. Psychological gender theories try to link these issues to psychological problems and mental illnesses that disproportionately affect either men or women. One such theory is Objectification Theory (Fredrickson & Roberts, 1997), which attempts to explain women’s mental health risks by their experience of internalizing an observer’s view as their own view of themselves and their bodies. Theorists who subscribe to Objectification Theory see sexual objectification, which treats women as objects, as a form of gender oppression. When women are objectified, their bodies are being separated from their person. Objectification is manifested in the male “objectifying gaze” (Fredrickson & Roberts, 1997) p. 176), through which the female body or body parts are scrutinized, evaluated, and commented upon. As a young female grows up, she will realize that her body is expected to become increasingly available to others, an expectation that is fostered by the female stereotype that women be open, friendly, warm, caring, and unselfish. Proponents of this theory argue that
women internalize external objectification at some point in their lives and subjugate themselves to the “objectifying gaze” (Fredrickson & Roberts, 1997) p. 176), leading to “self-objectification” (Fredrickson & Roberts, 1997) p. 179) and self-monitoring behavior. This, in turn, renders women to be more self-conscious and causes them to divert their attention from a focus on internal processes to an external focus on outward physical appearance. The mental health consequences of self-objectification are reported to be an increase in shame and anxiety, and a decrease in peak motivational states as well as awareness of internal physical states (Fredrickson & Roberts, 1997). Lewis (1989) and Lewis (1992) (as cited in Fredrickson & Roberts, 1997) state that shame is a powerful emotion often caused by falling short of societal expectations, thus regulating behavior to maintain certain societal standards. Frederickson and Roberts (Fredrickson & Roberts, 1997) further argue that women develop a sense of uncertainty with regard to when and where their bodies might be subject to objectification, which makes them more vulnerable to develop higher levels of anxiety. The above authors suggest that objectification harms women’s creativity and motivation, in that frequent comments on their physical appearance interrupt their experiences of “flow” (Csikszentmihalyi, 1975) and thus interrupt their peak motivational states. Further, Frederickson & Roberts (Fredrickson & Roberts, 1997) state that, due to self-objectification, women are at a higher risk of losing touch with their inner physical experiences, and that, taking higher levels of shame and anxiety into account, as well as impaired creativity and lower sensitivity for internal physical experiences, women could be at a higher risk for mental health problems, such as depression, sexual dysfunction, and eating disorders.

Jack (Jack, 1991) developed a theory which explains women’s tendency to develop depressive symptoms in relationships. According to Jack (Jack, 1991), women are prone to
developing relationship schema that cause them to silence their selves (i.e. deny their needs, feelings and wishes, or sacrifice their needs for those of others) due to traditional societal gender norms. Jack’s Silencing-the-Self Theory is based on attachment theory, as well as self-in-relation theory, which holds that women place great importance on relationships because of the way they are socialized: unlike men, women do not disidentify with their mothers in the process of individuation and gender identity development, thus placing less importance on individuality and more importance on relationships. This is not to say that women do not become independent or need to be in relationships, rather, it seems to be the case that women become independent within the mother-daughter relationship and continue to place more importance on relationships with other people. Further, it has been shown that women tend to become depressed not over the absence of relationships, but the quality of their relationships (Carr, Gilroy, & Sherman, 1996). Based on these theories, the assumption is that women, due to traditional gender role expectations as well as the importance of relationships to women’s sense of self (Jack, 1991), tend to hold on to relationships that are unrewarding, or sacrifice their needs in functional relationships, making them more vulnerable to experiencing symptoms of depression.

Westcott (Westcott, 1989) examines gender relations theories of a gendered self-identity by Chowdorow (Chowdorow, 1978) and the Stone Center for Developmental Services at Wellesley College. According to Chowdorow (Chowdorow, 1978), males and females grow up to develop a different self-identity due to their mother’s unequal treatment. While females do not fully separate from their mothers in order to develop a gendered self-identity characterized by the ability to engage in and maintain relationships, males develop a sense of self based on separateness and individuality. Consequently, women continue to rely on others for self-validation. Westkott (Westkott, 1998) mentions Horney’s (Horney, 1937) work, which describes
women’s need to receive validation of feelings and self from others as potentially “compulsive”, weakening women’s sense of self by implanting in them the understanding that they have to fulfill other people’s expectations at the cost of fulfilling their own needs. According to Horney (Horney, 1950), women inevitably internalize an ideal self that they are bound to fall short of, leading to constant anger and criticism directed at the self, which Horney described as an “ongoing intrapsychic process”. Other theories have tried to explain how traditional gender relations lead to specific physical health risks for men and women. According to Sabo (Sabo, 1999), “emphasized femininity is constructed in reciprocal and subordinated relation to hegemonic masculinity in ways that reinforce masculine power and male-dominated hierarchies within varying institutional settings”. Sabo then argues that traditional gender relations result in a “negative gendered health synergy” that affects both men and women alike.

Research in Support of Mental Health Theories on Gender

In support of Fredrickson & Roberts (Fredrickson & Roberts, 1997) Objectification Theory, Stapley and Haviland (1989) (as cited in Fredrickson & Roberts, 1997 (Fredrickson & Roberts, 1997)) found evidence that women report feeling shame more often than men, thus pointing towards the possibility that women might regulate and adapt their behavior more to the prevailing norms than men. In a study of how self-objectification influences the prevalence of self-harm in college women, Muehlenkamp, Swanson, and Brausch (Muehlenkamp, Swanson, & Brausch, 2005) demonstrated that the relationship between self-objectification and self-harm is mediated by negative body regard and depression, thus lending indirect support to self-objectification theory. It can therefore be assumed that objectification puts women at an elevated risk of self-harming behavior. Klonoff et al. (Klonoff et al., 2000) showed that the total score on the Hopkins Symptom Checklist (HSCL-58) was significantly higher in female participants
when compared to men. Their findings lend indirect support to Objectification theory in that women with a higher rate on the Schedule of Sexist Events (SSE) had higher scores on the HSCL, while women who reported less frequent sexual harassment events scored lower on the HSCL, and did not significantly differ from the male participants in terms of psychiatric symptoms.

Thompson (Thompson, 1995) found evidence in support of the correlation between self-silencing behavior and depressiveness in women as postulated by Jack (Jack, 1991). Further, women’s depressiveness was associated with their own as well as their partner’s relationship dissatisfaction. By contrast, men’s level of depression was unrelated to their self-silencing and their partners’ relationship dissatisfaction. These findings therefore support the theory that relationships play a central role in the development and identity of women. Carr et al. (Carr et al., 1996) found significant differences between Caucasian and African American women in the effect of self-silencing on the level of depressiveness. The results of their study suggested that self-silencing was a significant predictor of depression in their Caucasian, but not in their African American sample, although both samples engaged in self-silencing behavior (in relationships with an intimate partner) to the same degree. In regards to their findings, Carr et al. (Carr et al., 1996) suggest that the two ethnic groups might have a different idea of romantic relationships due to cultural differences (making the African-American sample less vulnerable or more resilient to depression due to silencing than their Caucasian counterparts).

Westcott’s (Westcott, 1989) theory on the discrepancy between the internalized ideal and real self in women was associated with depression (Waters, Keefe, & Straumann, 2004) as well as depression and anxiety (Veale, Kindermann, Riley, & Lambrou, 2003). Self-discrepancy was shown to be associated with chronic lower back pain (Waters et al., 2004), while the emotional
and affective consequences of self-discrepancy, such as; stress, anger, depression and anxiety, have been linked to heart disease (Hill, Weber, & Werner, 2006).

In support for his theory, Sabo (Sabo, 1999) refers to research by Eisler and Blalock (Eisler & Blalock, 1991), Good and Mintz (Good & Mintz, 1989), Oliver and Toner (Oliver & Toner, 1990), as well as Sharpe and Heppner (Sharpe & Heppner, 1991), who showed that men who hold traditional beliefs in masculinity were at an elevated risk of depression as well as psychological stress. Sabo quotes research showing that traditional beliefs in manhood are associated with unintentional injury, homicide, and suicide (the three leading causes of death among American males between 15 and 34 years old) (Morbidity and Mortality Weekly Report, 1994, as cited in Sabo, 1999). (Pleck, Sonnenstein, & Ku, 1994) Pleck, Sonnenstein, and Ku, (1994) (as cited in Sabo, 1999), (Sabo, 1999) found an association between men’s risk behavior; i.e. problems at school, police involvement, increased sexual activity and promiscuity, as well as use of manipulation or force to have sex with a female, and traditional gender beliefs. Consequently, male risk behavior based on traditional gender beliefs severely impairs the safety and health of women, resulting in women being the target of male aggression, sexual violence, unplanned pregnancy, and the risk of contracting STDs (Sabo, 1998, as cited in Sabo, 1999) (Sabo, 1999). Sabo thus sees traditional masculinity and femininity at opposite ends of a continuum that carries (mental) health risks for men and women alike. However, the overall risk for women seems to be higher, as male risk behaviors ensuing internalized traditional male role beliefs have stronger repercussions on women than vice versa, and affect women in addition to the traditional female role expectations that potentially impair women’s mental and physical health.
Psychometric Properties of BASIS-24

In order to assess whether or not BASIS-24 is a suitable outcome measure to use among male and female psychiatric patients, it is important to take a look at its psychometric properties as assessed in previous research. The following paragraphs will provide an overview of the overall psychometric properties of BASIS-24 subscales for depression, self-harm, substance-abuse, psychosis, interpersonal relationships, and emotional lability.

*Depression*

According to the DSM IV, a diagnosis of “Major Depressive Syndrome” is warranted in the presence of at least five of the following symptoms (and in the absence of mania and psychotic disorders): depressed mood; markedly diminished interest or pleasure in all or almost all activities; significant weight loss or weight gain when not dieting; decrease or increase in appetite; insomnia or hypersonmia; psychomotor agitation or retardation; fatigue or loss of energy; feelings of worthlessness or excessive or inappropriate guilt; diminished ability to think, concentrate or indecisiveness; recurrent thoughts of death or suicide without a specific plan, or a suicide attempt or a specific plan for committing suicide. Major Depressive Disorder can be classified as mild, moderate or severe.

The BASIS-24 depression domain is essentially a conglomeration of the daily/role functioning and depression/anxiety domains. The six items of the Basis-24 domains are visually grouped separately on the instrument, with the “managing your day-to-day life”, “coping with problems in your life”, and “concentrating” items from the daily/role functioning in one group, and “feel confident in yourself”, “feel sad and depressed”, and “feel nervous” grouped with the item “think about ending your life”, which is part of the self-harm domain. Eisen, Normand, Belanger, Spiro, and Esch (S. V. Eisen, Normand, Belanger, Spiro, & Esch, 2004) reported high
Review of the literature seems to suggest that the overall prevalence of depression is highest in the oldest old, while certain types of depression are more likely to occur in specific age groups, i.e. major depression in the middle-aged and minor depression in the elderly. Being female and middle-aged could thus be risk factors for experiencing a more severe form of depression, while being female and elderly are risk factors for developing a milder form of depression, but also for committing suicide. Further, being male and elderly seems to predict an increased risk for completing suicide.

Self-Harm

Self-harm (SH) is deliberate injury to one's own body without apparent suicidal intent. This injury may be aimed at relieving otherwise unbearable emotions, sensations of unreality and numbness, or for other reasons (Self-harm.2007). It can take the form of cutting, taking overdoses of tablets or medicines, punching oneself, throwing one’s body against something, pulling out hair or eyelashes, scratching, picking or tearing at one's skin causing sores and scarring, burning, inhaling or sniffing harmful substances (What is self harm? 2007). Self-harm is sometimes associated with personality disorders (especially Borderline Personality Disorder), substance abuse, eating disorders, PTSD, major depression, and anxiety disorders (Klonsky, Oltmanns, & Turkheimer, 2003).

The BASIS-24 self-harm domain contains two questions; “think about ending your life”, formerly in the depression/anxiety domain, and “think about hurting yourself” from the impulsiveness domain. Eisen et al. (S. V. Eisen et al., 2004) found excellent internal consistency reliability for this subscale ranging from 0.88 for inpatients to 0.91 for outpatients.
Gender Invariance of BASIS-24

reliability for this subscale, with a Cronbach alpha value of 0.89 for inpatients and 0.86 for outpatients. Test-retest reliability ranged from 0.96 for inpatients to 0.89 for outpatients.

As for the functions of self-harm, past research has found Deliberate self harm (DSH) to serve the purpose of emotional self-regulation (Farber, 2000), preventing the loss of an attachment figure (Rosen et al. 1997, in Farber, 2000) (Farber, 2000), release of intense emotions, such as anger (Harris, 2000), reducing anxiety (Klonsky et al., 2003), and self-soothing (increased emission of endorphins after experiencing pain) (Winchel & Stanley, 1991).

Relevant research suggests no significant differences between males and females with regard to the overall prevalence of self-harm. Self-harm in males after experiencing trauma was, however, found to be more pronounced and more violent than in females, who engaged in less violent forms of self-harm (Farber, 2000). Further, men were found to be more likely to commit suicide after self-harm (J. Cooper et al., 2005).

There was a difference between men and women as well as between Blacks and Whites in the factors causing self-harming behavior. Self-objectification and negative body regard were identified as related to DSH in women. Females stated problems within the family as reasons to self-harm, while males stated issues outside the family (Keeley, O'Sullivan, & Corcoran, 2003). Klonsky et al. (Klonsky et al., 2003) seem to indicate that there might be a difference between males and females with regard to the emotional states that self-harmers seek to resolve, with males trying to resolve anxiety more than females. Social stress was found to play a significant role in the etiology of DSH in Black males (Goddard, Subotsky, & Fombonne, 1996).

Substance Abuse

The DSM IV TR defines substance dependence as continued use of a substance (drug of abuse, medication, or toxin) despite the presence of groups of symptoms indicating significant
impairment due to substance use, characterized by developing tolerance, experiencing withdrawal symptoms, and engaging in compulsive drug-seeking behavior. Substance abuse is characterized by continued substance use despite experiencing adverse consequences for a period of more than 12 months in the absence of tolerance, withdrawal, and compulsive drug-seeking behavior.

BASIS-24 has four items in the Substance Abuse domain. They include a question about having an “urge to drink alcohol or take street drugs”, as well as “did anyone talk to you about your drinking or drug use”, “did you try to hide your drinking or drug use”, and “did you have problems from your drinking or drug use”. The results of Eisen et al. (S. V. Eisen et al., 2004) show the new subscale to have good internal consistency (0.88 inpatients, 0.82 outpatients) and excellent test-retest reliability (0.91 for both levels of care).

Studies have shown that being married, living with children and being employed were protective factors against substance abuse for both males and females (SAMHSA, 2004). Divorced women were the most likely to abuse substances, while there was no significant difference in substance abuse between divorced, and single, never before married men. Men were more likely overall to abuse substances than women, while reverse findings were found among Mexican American women for substances other than alcohol. Further, there were significant differences in substance abuse in women belonging to different ethnic groups, with African American women being the most likely to abuse substances.

**Psychosis**

The diagnostic criteria for schizophrenia according to DSM IV TR are the presence of two of the following symptoms for the majority of the time during a one-month period: delusions, hallucinations, disorganized speech, grossly disorganized or catatonic behavior, and
negative symptoms, such as flat affect. (Other psychotic disorders may use some or all of the above diagnostic criteria plus additional criteria that allow for a differential diagnosis).

The psychosis domain is comprised of symptoms correlated with psychotic disorders. BASIS-24 has four questions: “think you had special powers”, “hear voices or see things”, “think people were watching you”, and “think people were against you”. Eisen et al. (S. V. Eisen et al., 2004) obtained internal consistency values of 0.77 for both inpatients and outpatients, and test-retest reliability values of 0.83 for inpatients and 0.95 for outpatients with the psychosis subscale of BASIS-24.

The latest version of the DSM (IV TR) includes information about gender differences in the onset of schizophrenia: Whereas the age-at-onset is between 25 and the mid-thirties for women, it is between 18 and 25 years for men. Further, there is a second risk time for women to develop schizophrenia later in life (3%-10% of the women develop schizophrenia after 40+) whereas later onset of schizophrenia in men is reportedly rare. As for the genetic component of schizophrenia, more women with schizophrenia were found to have relatives with a similar diagnosis than men. Men on the other hand were found to have more family members with schizotypal and schizoid personality traits than women. While no evidence of gender-based differences has been found for the overall prevalence of schizophrenia among men and women; men tend to suffer from a type of schizophrenia with an earlier onset (Kohen, 2001); (World Health Organization, 2002), poorer adjustment prior to onset, as well as poor outcome (Ramsay et al., 2001). Further, men were found to be at a higher risk for a neurodevelopmental form of schizophrenia (Castle & Murray, 1991, as cited in Ramsay et al., 2001) (Ramsay et al., 2001), resulting from perinatal insults to the central nervous system (O’Callaghan et al., 1992, as cited in Kohen (Kohen, 2001), while women are more likely to have a strong genetic component
Whereas men tend to have poor premorbid adjustment (Castle & Murray, 1991, as cited in Ramsay et al., 2001), women were found to have better premorbid competence (both sexual and social), while there was evidence of a more benign and remitting course of illness (Flor-Henry, 1990, as cited in Kohen, 2001). Perry, Moore, and Braff (Perry, Moore, & Braff, 1995) used two different measures to assess schizophrenia symptoms in male and female patients, specifically thought disorder and social competence. Males scored higher on the thought disorder scale (the Ego Impairment Index). Thought disorder was highly correlated with social competence, and males scored higher on the Social Competency Index, suggesting that the males in this sample were significantly less socially competent than their female counterparts.

**Interpersonal Relationships**

Interpersonal relationships have been described by Weiss (1969, 1974), (as cited in Duck & Cohen Silver, 1990) as fulfilling six needs: opportunity for being nurturant (feeling needed by others), attachment (emotional closeness), social integration (a sense of belonging to a group who share similar interests, concerns, and activities), reassurance of worth (recognition of competence, skill, and value by others), guidance (advice or information), and reliable assistance (persons who can be counted on for tangible assistance). Berg and Piner (1990, as cited in Duck & Cohen Silver, 1990) further divide social support through interpersonal relationships into emotional support, informational support, instrumental support, and companionship. They describe social networks by the characteristics of network range/size, support network, and network density. Different types of social support fulfill different types of social needs: the social need of attachment is
satisfied by emotional support, guidance by informational support, reliable assistance by instrumental support, and social integration/reassurance of self-worth by companionship.

The relationships domain of BASIS-24 covers a person’s perception of the quality of their interpersonal experience. It has five items that request client judgment of ability in this area of functioning. The five items include: get along with people in your family, get along with people outside your family, get along well in social situations, feel close to another person, feel like you had someone to turn to if you needed help. These items require assessment of comfort in family relationships, relations with others, and feeling close to others. In Eisen’s et al., (S. V. Eisen et al., 2004) analysis of this subscale, reliability was very good: internal consistency reliability as indicated by Cronbach alpha was good at 0.81 for inpatients, and 0.84 for outpatients, while the test-retest reliability value (intraclass correlation coefficient) reached 0.81 for the inpatient sample and 0.93 for outpatients.

Berg and Piner (1990, as cited in Duck & Cohen Silver, 1990 (Duck & Cohen Silver, 1990)) outlined gender differences in experiencing loneliness in interpersonal relationships as well as differences between masculinity and femininity across gender. In a review of literature, Berg and Piner (1990, as cited in Duck & Cohen Silver, 1990 (Duck & Cohen Silver, 1990)) report that there is a relationship between loneliness and social network size. This association was stronger for men than women. This was supported by Lewin’s (1986, as cited in Duck & Cohen Silver, 1990 (Duck & Cohen Silver, 1990)) reanalysis of Stokes’ (1985, as cited in Duck & Cohen Silver, 1990 (Duck & Cohen Silver, 1990)) data, which showed a significant correlation between network size, density, number of confidantes, and loneliness for men but not for women. Wittenberg and Reis (1986, as cited in Duck & Cohen Silver, 1990 (Duck & Cohen Silver, 1990)) differentiate between relationship-initiating skills and relationship-maintaining
skills. They assign the former to “instrumental aspects of masculinity” and the latter to femininity. They found that lonely people who were higher in masculinity than femininity were more likely to suffer from a deficiency in the quantitative aspects of social contacts (e.g. frequency and length), while those higher in femininity were suffering from a lack of qualitative aspects (e.g. intimacy). Berg and Piner (1990, as cited in Duck & Cohen Silver, 1990) mention Jones’ et al. (1980, as cited in Duck & Cohen Silver, 1990) findings that women are more vulnerable to loneliness resulting from insufficient skills to maintain relationships.

**Emotional Lability**

The Royal College of Psychiatrists describes emotional lability as “an excessive emotional response to a minor stimulus, which is generally brief in duration.”

BASIS-24 contains three items that correspond to the emotional lability domain. These questions are; “have thoughts racing through your head”, “have mood swings” and “feel short-tempered”. Internal consistency values for this subscale from Eisen et al. (S. V. Eisen et al., 2004) were 0.75 for inpatients and 0.78 for outpatients; test-retest reliability was 0.96 for inpatients, and 0.89 for outpatients. Ott, Tate, Gordon, and Heindel (Ott, Tate, Gordon, & Heindel, 1996) found gender differences in emotional lability in Alzheimer’s patients. Using six defined behavior factors, they found that men were more apathetic and presenting with vegetative signs, whereas women displayed more reclusiveness and emotional lability. Epstein, J. N. et al. (Epstein, J. N. et al., 2000) found no gender differences in emotional lability in the parents of children with ADHD. Thus it seems that emotional lability occurs comorbidly with other disorders.
Measurement Equivalence

The following section describes the framework determining the importance of measurement equivalence for BASIS-24 across gender. Relevant literature related to measurement equivalence is introduced and theoretical underpinnings are highlighted. It is important to recognize that conducting assessments with a diverse population involves administering measures that were originally developed using only a subset of the population. Equivalence is not an intrinsic property of an instrument (Vandenberg, 2002); (Vandenberg & Lance, 2000); (Van de Vijver, F. & Leung, 1997).

Conceptual equivalence is included in virtually all models, peculiarly suggesting that the concept of equivalence does demonstrate conceptual equivalence: it has the same meaning across cultures and subpopulations. Also consistent across approaches is the recognition of scalar equivalence at the top of the hierarchy, and signifies that derived scores from an instrument measuring a particular construct apply to the same degree, intensity, or magnitude across groups (Choi & Harachi, 2002); (Miyamoto et al., 2001). Tran (Tran, 1997) refers to Hui and Triandis (Hui & Triandis, 1985) in his outline of cross-cultural equivalence. Conceptual equivalence is included, as well as metric and structural. Metric is described as concerning the ‘similarity of the relationships between observed items and their respective latent concepts or factors’ (p. 501): factor patterns and factor loadings correspond. Structural equivalence evaluates how the causes behind a research concept compare across groups.

Miyamoto et al. (Miyamoto et al., 2001) note five types of measurement equivalence that were each found frequently in their literature search: construct, scalar, functional, conceptual, and linguistic. Here, conceptual concerns, whether the concept being studied has similar meanings across groups, and construct equivalence pertains to the degree to which an instrument
measures the same construct across groups. Scores having similar predictors and outcomes, or comparable validity coefficients across groups demonstrate functional equivalence for an instrument. Linguistic equivalence is relevant when translations of instruments are involved.

Few studies have adopted the model purported by Van de Vijver, F. and Leung (Van de Vijver, F. & Leung, 1997). Here, levels of equivalence are deemed construct, measurement unit, and scalar. Achieving the level of construct equivalence, the same construct is measured but not necessarily operationalized in the same way across groups. In measurement unit equivalence, measurable by factor analysis, the measurement units are identical but the origins of the scales could differ. Scores are comparable on an interval level but not at the level of ratio. Scalar equivalence at the top of the hierarchy refers to whether a particular score on a measure represents the same degree, intensity or magnitude of the construct across groups. This is attained when scores on an instrument have the same interval scale across groups. Differences on an interval scale are measured at ratio level.

In the interest of investigating the quality of BASIS-24 as an outcome tool for a diverse psychiatric population, equivalence is important to measure. Sederer, Dickey, and Eisen (Sederer, Dickey, & Eisen, 1997) honor clinical relevance, usefulness, sensitivity to change, and cultural sensitivity in their list of the qualities of an ideal outcome assessment tool. Confirming measurement equivalence is conducted for the purpose of clarifying how appropriate the instrument is across categories of clients. Measuring equivalence involves operationalizing this variability and then striving for objectivity and consistency in both method and interpretation. The first step requires evaluating for equivalence and identifying how and where variance may be present. Subsequently, investigation of how to modify the instrument and reduce bias, and
further evaluation for equivalence complements the process of establishing the tool to be culturally sensitive.

Tran (Tran, 1997) describes conceptual equivalence as being somewhat subjective and advocates using both experts and clients, or respondents, from different groups to evaluate conceptual contents in an instrument (Mishler, 1986). Construct equivalence can be measured using multidimensional scaling.

Several options are available to examine the internal structures of a measure across groups including computing reliability coefficients (Cronbach’s alphas), using exploratory and confirmatory factor analysis (EFA and CFA) or conducting an item response theory (IRT) analysis. IRT, also called latent trait theory, is appropriate when there are circumstances that threaten the quality of measurement tools that aren’t directly addressable through Classical Test Theory (CTT) approaches such as the calculation of reliability coefficients. A CFA has the advantages of being less biased than reliability coefficients by low numbers of items in the measure and of providing fit indices that are more readily comparable across groups.

It can be more difficult to determine whether an instrument has attained scalar equivalence. Salzberger et al. (Salzberger, Sinkovics, & Schlegelmilch, 1999) determined that factor analysis is unable to affirm scalar equivalence and allow for comparisons across cultures, as bias affecting all stimuli of an instrument in generally the same way will not be detected. In attaining this level, they recommend extending the factor analysis model by including item intercepts, where an intercept is the constant of the regression of the observed score on the score of the latent variable. The intercept of one item is fixed, and then the intercepts of all other items are estimated relative to this value. When cultural aspects are affecting an item’s difficulty, the fit of the scalar invariance model decreases: the same observed score in two cultures will show a
difference in levels of the latent score between the two cultures. Tran (Tran, 1997) suggests using path analysis or structural equation modeling analysis for structural equivalence; structural equivalence encompassing a similar idea to scalar.

As argued by Bollen (K. A. Bollen, 1989), measurement equivalence is defined through invariance of factor correlation, invariance of factor loadings, and invariance of measurement error. However, even in substantive research, the assumption that there is measurement equivalence was merely implied when comparing groups and most often not tested at all (B. Byrne, 1989). A number of statistical methods to evaluate measurement equivalence have been described in the literature. Most commonly, they are item response theory (IRT) and confirmatory factor analysis (CFA) (Raju, N. S., Laffitte, L. J, Byrne, B. M, 2002). According to Raju et al. (Raju, N. S., Laffitte, L. J, Byrne, B. M, 2002), CFA provides a more streamlined method for the analysis of factor invariance. Moreover, Mullen argued that, “Multi Group LISREL [a CFA software] is, in general, the preferred approach for diagnosing measurement equivalence” (Mullen, 1995), p. 590). Therefore, this study employed CFA as the statistical method to determine the effects of gender on measurement equivalence of BASIS-24.

**Factor Invariance**

The above review of literature shows that often four levels of factor invariance were used to test the validity of an instrument. The baseline model should have no invariance constraints whereby the hierarchy of testing the instrument for invariance should begin with the least restrictive model, where there are not any invariance constraints on the estimated parameters between the groups being compared (M. J. Marsh & Hawkins, 1994); (B. Byrne, 1989).

Furthermore, to test models for comparison, “It is useful to test the same pattern of fixed and free parameters for all groups in the a posteriori baseline model” (M. J. Marsh & Hawkins, 1994).
Testing the order of the hierarchical constraints imposed on the model began with a totally free model with no constraints on the factor loadings, the factor correlations and variances, or the uniquenesses. The second model constrained only the factor loadings but did not restrict the factor correlations, variances or the uniquenesses. The third model restricted the factor loadings, factor correlations, factor variances and held them invariant but did not restrict the uniquenesses. The fourth and final model restricted variance across factor loadings, factor correlations, factor variances and also uniquenesses; this model is also known as the total invariant model.

Multi-group confirmatory factor analysis (MGCFA) is known as the best method testing measurement invariance because it provides the chi-square statistic and goodness-of-fit indices, emphasizing a priori model testing (Lim & Ployhart, 2004). Four types of measurement invariance tests were used: configural or pattern factorial invariance, weak metric invariance, strong metric invariance, and strict or complete metric invariance (Meredith, 1993). These tests were independently conducted and compared to determine the best fit model to the data. The parameter estimates for all models were obtained with maximum likelihood estimation by using LISREL 8.80 (Joreskog & Sorbom, 2006). The adequacy of model fit was assessed by the chi-square statistic, comparative fit index (CFI), and the root mean square error of approximation (RMSEA), as multiple fit indices are recommended for the evaluation of overall goodness-of-fit of models (Kline, 2005). Following convention (Werner Wothke, SmallWaters Corp., 2000), values higher than 0.95 for CFI indicate an acceptable model, and a model with a value lower than 0.08 for RMSEA was accepted.

Configural Factor Invariance (Pattern invariance)

Configural Factor Invariance is the least constrained model and is used to determine if BASIS-24 subscales can be used for both male and females with producing equality between the
two genders (Gregorich, 2006). Configural invariance is defined as the same factor loading pattern across groups, but no equality constraints. In this model, group differences in parameters may be attributable to group differences in elements of all parameter matrices. This model is typically the first model used because there are no constraints and is used as a stepping stone to build upon by slowly imputing constraints to check the equality of all three parameters by using weak factorial invariance, strong factorial invariance and strict factor invariance.

Based on the baseline model, initial constraints on the relationship among the error terms (i.e., error terms were not correlated) were gradually freed, and baseline models were established for each gender group (B. Byrne, 1998). Once baseline models were established for all the groups, multi-group confirmatory factor analyses (Joreskog & Sorbom, 1989) were then performed to assess the factor structure invariance across gender groups. Different constraints, which represented invariance in different aspects of the factor structure across groups, were imposed in a series of multisampling confirmatory factor analyses. Often goodness of fit of the model was examined on the basis of a low chi-square value which implies high correspondence between the matrices indicating a good fit of the model to the data. Because the chi-square statistic is highly sensitive to sample size, model fit was evaluated by examination of the root mean square error of approximation (RMSEA), and the comparative fit index (CFI) in accordance with convention (Hoyle & Panter, 1995). A value of 0.90 as a reasonable minimum for model acceptances, and the chi-square difference tests ($\Delta \chi^2$).

**Metric factor invariance**

With metric or weak invariance, the factor loadings for male and female are constrained to determine the equality of measurement between genders with factor correlation and error variances. This means the factors have the same meanings across groups. Establishing
weak/metric invariance is a prerequisite for cross-group comparison (K. A. Bollen, 1989). In this model, weak invariance requires that the same BASIS-24 subscales load into the same factors with the same values of loadings across gender, but the factor correlation and the measurement errors are free to vary across groups.

*Strong factorial (scalar or item) invariance*

This model examines whether the group mean comparisons are meaningful. This model examines the cultural norms that may cause differences in responses in one group compared to the other (Gregorich, 2006). Such differences are represented by the item intercepts or regression of items with their related factors. In this model, factorial invariance constrains factor loadings and intercepts to verify if the factor model of factor correlation is equal and measures the same between genders. As with weak factorial invariance, this specification implies that the measurement of the latent variables is the same across groups. Furthermore, the invariance in the intercepts in the mean structure allows for evaluating mean differences in latent variables across groups. Thus under strong factorial invariance, the group differences in covariances among observed variables and in means of observed variables are attributable to group differences in covariances and means on latent variables.

*Strict factorial (complete or item residual) invariance*

This model extends the previous model by invoking the additional constraint and unique variances across groups. Strict factor invariance checks the equality of all parameters; factor loadings, intercepts, and error or residual variances. This model is used to establish that all of the variances are similar and measure the same skill areas between males and females. This model is a highly constrained model.
In the context of BASIS-24, measurement equivalence is of critical importance. As argued, BASIS-24 is developed to assess patient outcomes and to improve treatment planning. The previously stated research findings outline why BASIS-24 data cannot simply be assumed to be equivalent across gender. As outlined, the purpose of this study is to determine if there is gender bias reflected in observed ratings on the BASIS-24 instrument in each of the six subscales. In examining mean scores, standard deviation, skewness, kurtosis and using CFA to explore the differences, if any, this study is designed to investigate the effects of gender on measurement equivalence of BASIS-24.

Patient Assessment and Evidence Based Practice in Social Work

Patient assessment is one of the basic steps in evidence based practice. Systematic patient assessment and outcomes are now part of the case management process in mental health service settings. The monitoring of patient’s functioning is essential to the process of mental health treatment (Kinzie & Manson, 1987). Especially with the emergence of third party payers for mental health services, patient assessment, and mental health outcomes becomes a necessary part of mental health services. Patient level and aggregate reporting is required for Joint Commission accreditation, state regulations, and managed health care systems. Broad interest has been growing among social workers, nurses and psychologists, in the field of psychiatry around patient assessment and outcomes. Much of the interest generated has been due to multiple factors including deinstitutionalization of mental health care, pharmacology, reduced length of stay in hospitalization, Joint Commission accreditation requirements and the professional recognition of evidence based care in clinical practice. Measuring progress has become a part of case management and social workers are primarily involved in patient assessment. However, the problem is then what instrument is best suited to accomplish this task. Social workers and nurses
are very much involved in the patient evaluation and treatment planning process. However, the research tradition of most social workers in practice is historically about case studies, qualitative analysis, and basic surveys dominate.

The NASW Code of Ethics states that “Social workers should monitor and evaluate policies, the implementation of programs, and practice interventions; social workers should promote and facilitate evaluation and research to contribute to the development of knowledge; social workers should critically examine and keep current with emerging knowledge relevant to social work and fully use evaluation and research evidence in their professional practice” (NASW, 2008), 5.02, (a-c). Evidence-based practice is based on evaluation research that highlights interventions that have been found to be effective (Cournoyer, 2004). Evidence-Based Practice is a thoughtful integration of the best available evidence, coupled with clinical expertise. An evidence-based practice model is a “the mindful and systematic identification, analysis, evaluation, and synthesis of evidence of practice effectiveness as a primary part of an integrative and collaborative process concerning the selection and application of service to members of target client groups.” (Cournoyer, 2004), p. 4). Establishing evidence-base practice in social work involves testing of existing assessment tools for its psychometric properties. Also, it is crucial to establish mechanisms to measure the way services and treatment are delivered as a first step toward documenting and monitoring outcomes from changing financial and organizational arrangements (Mechanic, 1996).

Several assessments instruments are now available for measuring the wide variety of symptoms presented by patients as well as general patient functioning. Some are administered within a clinical interview, others as a self-report. Outcome instruments typically target overall well being, patient satisfaction with care, symptoms, functioning, or functioning as well as
symptoms (Sederer et al., 1997). Sederer et al. (Sederer et al., 1997) outline a list of qualities that the model quality assessment instrument should possess. The instrument, ideally, should strive to be: clinically relevant, useful and timely; sensitive to change; culturally sensitive; low burden; low cost; involves the patient; built into standard operating procedures, and meets their own continuous quality improvement efforts as well as the expectations and demands of accreditors and regulators such as the Joint Commission on Accreditation of Healthcare Organizations, or JCAHO. Traditionally mental health assessment and practices were mostly focused on personality traits, substance abuse, suicidal thoughts, and violence to self or others, and diagnosis leading to a treatment plan. Clinical social workers and nurses do the bulk of mental health work, and the profession is having a hard time catching up with recent changes happening in patient assessment. Such trends include increasing emphasis placed on demonstrating the incremental validity of mental health assessment instruments (Butcher et al. 1995, Kuncel et al. 2001).

The BASIS-24 is one of a handful of instruments that measure both functioning and symptoms and the use of the BASIS-24 among psychiatric inpatients is significantly growing (S. V. Eisen et al., 2006). The Outcome Questionnaire (OQ-45) (Lambert et al. 1998, 1999), a brief and cost-effective measure, was designed to track treatment progress and outcome for patients with a wide variety of diagnoses. The Addiction Severity Index (ASI), measures aspects of a patient's life, as reported by a patient to a trained interviewer, related to pathological substance use. The Depression Outcomes Module (DOM) solicits both patient and clinician input regarding types of care, symptomatic and functional outcomes, and patient characteristics. The self-report instrument, SF-36, and the Quality of Life Interview (QOLI) similarly examine symptoms and functioning, as well as related factors, in the broader context of health-related quality of life. The Beck Depression Index and BPRS gauge more specifically patient symptoms. The BDI is a
21-item instrument designed for depressive symptoms that is completed by the patient. Eighteen items measuring symptoms of psychosis, depression, and anxiety comprise the Brief Psychiatric Rating Scale (BPRS). This scale includes patient and clinician evaluations. The SCL-90 and its short form, BSI (Brief Symptom Inventory) are other self-report instruments that assess symptomatology.

Though social workers perceive the need for evidence based practice and culturally appropriate instruments to guide evidence based practice, the profession lacks evaluative research on outcomes of methods (Braye & Preston-Shoot, 2007). Now implementation of evidence-based practice makes it important for the social work profession to accurately document mental health services, evaluation and outcomes. By identifying patients’ risk including substance abuse and suicidal thoughts, the BASIS-24 has the potential to be useful to social work educators, practitioners, and researchers who deal with this vulnerable population.
CHAPTER II: METHODOLOGY

Sample

The data came from Mental Health Services Evaluation Department of McLean Hospital, as a part of a field test study of 6000 psychiatric patients (S. V. Eisen et al., 2006). The original field test sample consisted of 6000 psychiatric patients admitted at 28 mental health and/or substance abuse programs throughout the country. Among those, 2656 individuals were admitted at 14 inpatient care facilities and the rest were admitted at outpatient facilities. Psychiatric patients who are (over the age of 18) inpatient admissions who arrived for treatment at one of the participating sites during the study period (May, 2001 through June, 2002) were eligible for inclusion in the study. Both mental health and substance abuse treatment centers were included in the study (which includes both mental health and substance abuse domains) among diverse programs and populations exhibiting a broad range of mental health problems. All four of the major geographic census regions (northeast, south, mid-west and west) were represented.

Table 1: Study sample by gender (paired for admission and discharge)

<table>
<thead>
<tr>
<th>Time Point</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male - Admission &amp; Discharge</td>
<td>773</td>
<td>55%</td>
</tr>
<tr>
<td>Female - Admission &amp; Discharge</td>
<td>625</td>
<td>45%</td>
</tr>
<tr>
<td>Total</td>
<td>1398</td>
<td>100%</td>
</tr>
</tbody>
</table>

Time 2 (discharge) BASIS-24 assessments were available for 1,398 of the inpatients (53%). Thus the final sample consisted of 1,398 inpatients that completed BASIS-24 at
admission and discharge. The study sample by gender by admission and discharge are presented in Table 1.

The concept of power in statistical theory is defined as the probability of rejecting the null hypothesis given that the null hypothesis is false (Jackson, 2003). In confirmatory factor analysis, the null hypothesis is defined by the specification of fixed and free elements in relevant parameter matrices of the model equations. The specification of fixed and free elements represents the researchers' initial hypothesis concerning the putative direct and/or indirect effects among the latent variables. The null hypothesis is assessed by forming a discrepancy function between the model-implied set of moments (mean vector and/or covariance matrix) and the sample moments. Confirmatory factor analysis often drives the sample size of a validation study given that when CFA has sufficient power, most classic psychometrics have sufficient power.

There are three methods to estimate the power for the confirmatory factor analysis (Jackson, 2003): 1) the easiest and least accurate method. In this method, one needs 10 participants per free parameter in the CFA. In a unidimensional model, the number of free parameters is often equal to 2 x the number test items (Bentler & Chou, 1987). Bentler & Chou (1987) recommended at least 10 participants per free parameter; 2) medium difficulty and accuracy based on RMSEA and formulas from (MacCallum, Browne, & Sugawara, Jun 1996). This method balances between free parameters and sample size; and 3) difficult and exacting method based on Monte Carlo simulation of power estimates based on formulas from (Muthén & Muthén, 2002). As there were 773 male and 623 female respondents, it was determined that model misfit would be more indicative of the inadequacy of the model, as opposed to the lack of a sufficient sample size to estimate the model. In addition, it should be noted that the sample
size in these analyses was very similar to the sample size used in Eisen et al. (S. V. Eisen et al., 2004).

Study Design

This study used structural equation modeling (SEM) methods to test the factorial invariance for inpatient psychiatric sample. The study hypotheses and research questions are formulated and examined using multi-group structural equation modeling design with pre and post samples of male (n=775) and female (n=623) psychiatric patients to validate parameter population estimates.

Measures

Behavior and Symptom Identification Scale (BASIS-24)

The BASIS-24 instrument is a twenty-four item patient self-report questionnaire designed to assess treatment outcomes by measuring symptoms and functional difficulties experienced by individuals seeking mental health services. The original tool, Behavior and Symptom Identification Scale (BASIS–32) was developed in the early 1980s to meet the need for a brief but comprehensive mental health status measure that would be useful in assessing the outcomes of mental health treatment from the consumer’s point of view. It is a measure of self-reported difficulty in the major symptom and functioning domains that lead to the need for mental health services (S. V. Eisen, Dill, & Grob, 1994).

The new BASIS-24 survey cuts across diagnoses, recognizing the wide range of symptoms and problems that occur across the diagnostic spectrum. BASIS-24 is designed to measure outcome for a broad range of treatments and services encompassing many theoretical orientations. Scores can be computed for the overall BASIS-24, as well as for six domains:
Depression and Functioning, Interpersonal Relationships, Psychosis, Substance Abuse, Emotional Lability, and Self-Harm. Development of BASIS-24 has been described in detail in previous publications and a brief profile of the tool is included in this section (S. V. Eisen et al., 2006).

Among the 24 items, each has 5 ordered response options reporting either the level of difficulty experienced (no difficulty to extreme difficulty), or the frequency with which a symptom or problem has occurred (none of the time to all of the time). Respondents answer each question in terms of how they have been during the past week. For example, “During the past week, how much of the time did you feel sad or depressed”. BASIS-24 is administered at the beginning of a treatment episode, with repeat assessments obtained at desired intervals to assess change during or following treatment. All items are answered on a 5-point scale with different sets of response options tailored to particular sets of questions. The survey was written at a 5th grade reading level in order to maximize the number of individuals who are able complete it. Shorter than its predecessor BASIS-32 yet more comprehensive, BASIS-24 cuts across diagnoses by identifying a wide range of symptoms and problems that occur across the diagnostic spectrum. Validated and found reliable in inpatient, residential, and outpatient settings, BASIS-24 assesses treatment outcomes from the patient perspective.

**Demographic Characteristics**

Age, gender, ethnicity, race, education and marital status were obtained by patient self-report questions appended to BASIS-24. Patient clinical information such as primary diagnosis, comorbidity, GAF scores were taken in the patient characteristics form completed by the clinicians.
Data Collection Procedure

BASIS-24 was administered twice, upon admission, and in the 24-hour period before discharge (for inpatients). It was administered by program staff within the context of continuous quality improvement (CQI) programs as part of routine outcomes monitoring. Verbal consent was obtained from all participants. This data collection process was approved by the Institutional Review Board of the grantee institution and by each participating site. Demographic characteristics, admission and discharge dates, payer, and DSM-IV psychiatric diagnoses including Global Assessment of Functioning (GAF) ratings were extracted from medical records or administrative databases.

Operational Hypothesis

The present study examines the factor structure and gender invariance of BASIS-24. In this section, the hypotheses are presented along with the conceptual framework within which the analyses were conducted. The following four operational hypotheses were tested:

Hypothesis 1

The factor structure of BASIS-24 is invariant across gender groups. Both male and female groups associate the same subscales of BASIS-24 items. The constructs are manifested in the same way across groups. Based on the prior research of BASIS-24 factor structure by Eisen al. (S. V. Eisen et al., 2004), it was hypothesized that the six-factor structure of BASIS-24 subscales would be replicated across gender groups at admission and discharge. This was tested by endorsing all BASIS-24 subscales into same factor patterns across gender, but allowing the magnitude of factor loadings, factor correlation and the measurement errors to vary across gender.
Figure 1: A conceptual six-factor model for BASIS-24

Latent variables are represented by circles; scale items are represented by squares.
**Hypothesis 2**

The factor loadings of BASIS-24 subscales are invariant across gender. The strength of the relationship based on the factor loadings between each item and its underlying construct is the same for both groups. In addition, the factor loadings are constrained to be equal across groups, but no other equality constraints are imposed. This was tested by loading the same BASIS-24 subscales into the same factors with the same values of loadings across the genders, but the factor correlation and the measurement errors are free to vary across groups.

**Hypotheses 3**

The cross differences and relationships among BASIS-24 subscales (covariances) are the same across gender groups. This was examined by whether the group comparisons are meaningful. Invariance of both factor loadings and intercepts is required to verify if the factor model of factor correlation is equal and measures the same between genders. Thus under strong factorial invariance, the group differences in covariances among observed variables and in means of observed variables are attributable to group differences in covariances and means on latent variables. The third hypothesis tested was a more restrictive model with equal factors.

**Hypothesis 4**

BASIS-24 items have an invariant pattern of unique variances or residual or error variances. This checks the equality of all parameters; factor loadings, intercepts, and error or residual variances. This is used to establish that all of the variances are similar and measure the same skill areas between males and females.
Data Analysis

Data analyses proceeded in three steps: first, the tests of data quality (examination of the rate of missing data and floor and ceiling effects) and tests of internal consistency reliability; and second, confirmatory factor analysis (CFA) for male and female groups at admission and discharge; and third, multiple group confirmatory factor analysis to test the gender invariance of BASIS-24.

Descriptive statistics

Item frequency distributions were generated to assess distribution of the male and female sample distribution. Preliminary analysis was done to test data quality, which examined the rate of missing data, floor and ceiling effects, summarized the mean, standard deviation, skewness, kurtosis, and examined the internal consistency reliability, and construct validity of BASIS-24. High rates of missing data can identify items that are confusing, difficult to answer or inapplicable to respondents. Extensive floor and ceiling effects can indicate insensitivity of the instrument to individual differences in symptom levels at the extreme ends of the continuum, or inapplicability of items to the sample. Reliability and validity analyses were conducted separately for male and female samples. Cronbach’s alpha analyses were computed to assess internal consistency reliability of each subscale and the total scale.

Confirmatory Factor Analysis

Structural equation modeling (SEM) was developed from work in econometrics (K. A. Bollen, 2002). SEM is an enormously flexible technique and it is possible to use a structural equation modeling approach to carry out direct equivalents of many analyses, including analysis of variance (ANOVA), correlation, multiple regression, multivariate analysis of variance, and
multivariate regression. Confirmatory factor analysis (CFA) is one of the most preferred statistical methods used by social scientists to examine the factorial invariance across groups such as gender, ethnicity and diagnostic groups (H. W. Marsh, 1994); ((Pedhazur & Schmelkin, 1991). CFA is generally based on a strong theoretical foundation that allows the researcher to specify the factor model a priori (Pedhazur & Schmelkin, 1991). For the present study, LISREL 8.8 structural equation modeling program was used to perform all single and multi-group confirmatory factor analyses (Joreskog & Sorbom, 2006).

**Single-Group Confirmatory Factor Analyses**

Based on the prior research of BASIS-24 factor structure by Eisen al. (S. V. Eisen et al., 2004), it was hypothesized that the six-factor structure of BASIS-24 (Depression, Emotional Lability, etc.) were replicated across gender groups at admission and discharge. To establish a well-fitting baseline model for each group separately prior to testing for factorial invariance, separate analysis was conducted for: a) the entire sample at admission; b) the entire sample at discharge; c) the male sample at admission; and d) the female sample at discharge. This procedure allows for model testing and specification searches (MacCallum, Jul 1986) to be carried out on one subsample (the test sample) and for model cross-validation to be performed on the second subsample. Often, error variances and their covariances were also tested for invariance.

These tests were independently conducted and compared to determine the best fit model to the data. Because the data were ordinal, the diagonally weighted least squares (DWLS) method with Polychoric correlation matrices for each group was run. DWLS also required an estimate of the asymptomatic covariance matrix of the sample correlations for each group. Both the Polychoric correlation matrices and the asymptomatic covariance matrices were calculated.
with Joreskog and Sorbom’s LISREL software (Joreskog & Sorbom, 2006). Although the standard ML method has been shown to be fairly robust under violation of the multivariate normality assumption, the DWLS method is preferred with ordinal level data (Joreskog & Sorbom, 2006).

*Multiple-group Confirmatory Factor Analysis*

Multi-group confirmatory factor analysis was used to test four types of measurement invariance models of BASIS-24 across gender. MGCFA is known as the best method for testing measurement invariance because it provides the chi-square statistic and goodness-of-fit indices, emphasizing a priori model testing (Lim & Ployhart, 2004). The four types of measurement invariance tests used were: configural or pattern factorial invariance, weak metric invariance, strong metric invariance, and strict or complete metric invariance (Meredith, 1993).

Configural Invariance investigates whether the respondents from different groups employ the same conceptual framework to different evaluation tools (S. Cheung, 2002). In MGCFA, constraining the number of factors and the pattern of the free and fixed loadings to be the same across groups, test configural invariance. Failure to demonstrate configural invariance indicates that different constructs were measured across groups. It is the least constrained or non-variant multi-group model and indicates that the factor structure underlying item responses is equivalent across different group (Gregoric, 2006). In this model, all BASIS-24 subscales should be endorsed into same factor patterns across gender, but allowing the magnitude of factor loadings, factor correlation and the measurement errors to vary across gender. This model is important and the first model used because there are no constraints and is used as a stepping stone to build upon by slowly imputing constraints to check the equality of all three sets of parameters by using weak factorial invariance, strong factorial invariance and strict factor invariance. This model
tested the hypothesis that an a priori pattern of free and fixed factor loadings imposed on the measures was equivalent across groups (Horn & McArdle, 1992). Thus the first hypothesis tested was that the six factor structure is invariant across gender groups. If the data does not suggest the rejection of this hypothesis, testing of the intermediate nested models is unnecessary.

In weak factor invariance model, in addition to the configural constraints, the factor loadings are constrained to be equal across groups, but no other equality constraints are imposed. With weak invariance, the factor loadings for male and female are constrained to determine the equality of measurement between genders with factor correlations and error variances. This means the factors have the same meanings across groups. Establishing weak/metric invariance is a prerequisite for cross-group comparison (K. A. Bollen, 1989). Weak invariance requires that the same BASIS-24 subscales load into the same factors with the same values of loadings across the genders, but the factor correlations and the measurement errors are free to vary across groups.

Strong factorial (scalar or item) invariance model examines whether the group comparisons are meaningful. Invariant factor loadings are not enough to compare scale scores across groups. Invariance of both factor loadings and intercepts is required. In this model, factorial invariance constrains factor loadings and intercepts to verify if the factor model of factor correlations is equal and measures the same between genders. As with weak factorial invariance, this specification implies that the measurement of the latent variables is the same across groups. Furthermore, the invariance in the intercepts in the mean structure allows for evaluating mean differences in latent variables across groups. Thus under strong factorial invariance, the group differences in covariances among observed variables and in means of
observed variables are attributable to group differences in covariances and means on latent variables.

Strict factorial (complete or item residual) invariance model extends the previous model by invoking the additional constraint and unique variances across groups. Strict factorial invariance checks the equality of all parameters; factor loadings, intercepts, and error or residual variances. This model is used to establish that all of the variances are similar and measure the same skill areas between males and females. This model is a highly constrained model and may often not hold in practice. In fact, there is reason to expect that it would not hold, even if strong factorial invariance does hold. Even if all populations come from a common parent population with given error variances, it would be expected that error variances would vary from one subpopulation to another.

Model Fit

Assessment of the goodness of model-data fit was conventionally based on $\chi^2$ test and a group of descriptive goodness-of-fit indices such as root mean square error of approximation (RMSEA), its confidence interval, standardized root mean square residual (SRMR), and comparative fit index (CFI). However, the $\chi^2$ test is sensitive to sample size, and the hypothesized model is likely to be rejected when the sample size is large, even though the discrepancy between the sample covariance matrix and model-predicted covariance matrix may be small or trivial (Fan, Thompson, & Wang, 1999) and clinically significant. Accordingly, other additional fit indices were used to supplement the $\chi^2$ statistic, the root mean squared error of approximation (RMSEA) (Steiger, 1990), goodness-of-fit index (GFI) and adjusted goodness-of-fit index (AGFI) (Joreskog & Sorbom, 1989), normed fit index (NFI), non-normed fit index (NNFI), and comparative fit index (CFI) (Bentler, 1990).
For single group confirmatory factor analyses, the model fit was examined using RMSEA, CFI, and SRMR. RMSEA values of less than .06 indicate a good fit and values as high as .08 indicate a reasonable fit (L. Hu & Bentler, 1999). CFI was also reported to complement RMSEA (S. Cheung, 2002). Although a value of .90 for CFI has served as a rule-of-thumb lower limit cutoff of acceptable fit, a value of .95 is expected of models considered to be well fitting (L. Hu & Bentler, 1999). The SRMR is another residual based index of fit. A SRMR between 0 and 0.05 indicates a good fit and between 0.05 and 0.10 an acceptable fit ((Walter, Schermelleh, Cremer, Tashiro, & Cremer, 2003) (Schermelleh-Engel, Moosbrugger, & Müller, 2003). For the single group analysis, a RMSEA value of at least 0.08 is preferred and values less than 0.05 are considered good. Also CFI values greater than 0.95 were preferred and values near 0.90 were considered acceptable. Finally, a $\chi^2$ test of model fit is taken, where rejection of the null hypothesis suggests inadequate model fit. However, the Chi Square is sensitive to small deviations of fit in large samples, and significant results should necessarily be interpreted with caution. Therefore, a significant Chi Square alone is not used to reject a given model. In sum, a number of fit indices are used to assess global fit. No single index of fit is taken as evidence of model misfit. Rather, agreement among the indices was examined to determine model fit. Evidence for model misfit is considered when the majority of indices suggested the model poorly fit the data.

For multi-group confirmatory factor analysis, the model was generally assessed by the $\Delta \chi^2$ value between two nested models. A chi-square test provided an indication of whether there was a statistically significant difference in fit between the constrained and unconstrained models. A statistically non-significant difference between the two models would be evidence of factorial invariance across groups. Researchers have argued that this $\Delta \chi^2$ value is as sensitive to sample
size and nonnormality as the $\chi^2$ statistic itself, rendering it an unrealistic criterion on which to base evidence of invariance (G. W. Cheung & Rensvold, 2002). Increasingly, two alternative criteria have been used to argue for evidence of invariance: the multi-group model exhibiting an adequate fit to the data and the delta comparative fit index ($\Delta$CFI) values between models being negligible (B. M. Byrne, 2006). Many recent authors have considered CFI as the most appropriate index for examining cross-group differences in invariance testing and have recommended that $\Delta$CFI should not exceed .01 because their studies showed that decreases in CFI of .01 or less were within the range of sampling error (B. M. Byrne, 2006; G. W. Cheung & Rensvold, 2002).

In the present study gender invariance tests for all four models were performed in the following hierarchical order: configural invariance, weak invariance, strong invariance, and strict invariance. Therefore, two criteria was used to compare the nested multi-group models: a) a non-significant chi-square difference test; and b) a $\Delta$CFI value less than or -.01 was used to make final decisions about whether weak, strong, and strict measurement invariance models hold.
CHAPTER III: RESULTS

The primary purpose of this study was to test the factorial structure of BASIS-24 across gender. The first part of this chapter summarizes the descriptive statistics based on the following questions: are there differences between the mean, standard deviation, skewness, and kurtosis for male and female BASIS-24 scores; does BASIS-24 have adequate internal consistency across gender; and are there differences between the internal consistency reliability of BASIS-24 for men and women.

In the second section, the results from the main analyses are presented, as they are related to the confirmatory factor analysis (CFA) and multi group confirmatory factor analysis (MGCFA) in order to examine the factor structure and gender invariance of the BASIS-24 scale. Confirmatory factor analysis was conducted to examine whether the male and female datasets revealed the following six factors: a) Depression/Functioning, (b) Interpersonal, (c) Self-Harm, (d) Emotional Lability, (e) Psychosis, and (f) Substance Abuse. In this regard, results of the four following operational hypotheses are included:

1. The six factor structure of BASIS-24 is invariant across gender groups. Both male and female groups associate the same subscales of BASIS-24 items (configural invariance). The constructs are manifested in the same way across groups.

2. The factor loadings of BASIS-24 are invariant across gender. The strength of the relationship based on the factor loadings between each item and its underlying construct is the same for both groups (weak or metric invariance).

3. The group differences of BASIS-24 among constructs (covariances) are the same across groups (strong invariance).
4. BASIS-24 items have an invariant pattern of unique variances or residual or error variances. In other words, BASIS-24 items assess the same underlying construct (strict or scalar invariance) for men and women.

Descriptive statistics

Sample Characteristics

The study sample consists of 1,398 inpatients that completed BASIS-24 at admission and discharge. Table 1 shows the social and demographic and clinical characteristics of the study sample by gender. Half of the participating facilities in this study were state or county psychiatric hospitals (49%), followed by private psychiatric hospitals (38%). Most were in suburban settings (61%). The majority of respondents were male (55%), and the rest (45%) were female. For both genders, the most common age range of respondents was 35-44 (31%), followed by 25-34 (22%) and 45-54 (21%) respectively. There were more female white participants (66%) than male (58%), but more black/African American males (32%) than females (23%). The female population was more educated than the male population, with 31% of females having some college degree compared to 21% of males. Despite a slightly higher level of education females reported slightly higher unemployment rates within the past 30 days, (68% of females were unemployed versus 63% of males). Comorbid diagnoses were similar for males and females, with over two thirds of the sample with at least one comorbid psychiatric diagnosis.

Preliminary analysis was done to test data quality, which examined the rate of missing data, floor and ceiling effects, summarized the mean, standard deviation, skewness, kurtosis, and
examined the internal consistency reliability, and construct validity of BASIS-24. These data are reported at admission and discharge on Table 2.

Table 2: Clinical and Demographic Characteristics

<table>
<thead>
<tr>
<th>Patient Characteristics</th>
<th>Male (n=773) (%)</th>
<th>Female (n=625) (%)</th>
<th>Total (n=1398) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18-24</td>
<td>16.4</td>
<td>13.9</td>
<td>15.3</td>
</tr>
<tr>
<td>25-34</td>
<td>23.9</td>
<td>22.2</td>
<td>23.2</td>
</tr>
<tr>
<td>35-44</td>
<td>31.6</td>
<td>31.2</td>
<td>31.4</td>
</tr>
<tr>
<td>45-54</td>
<td>20.3</td>
<td>23.0</td>
<td>21.5</td>
</tr>
<tr>
<td>55-64</td>
<td>5.4</td>
<td>6.4</td>
<td>5.9</td>
</tr>
<tr>
<td>65+</td>
<td>2.3</td>
<td>3.2</td>
<td>2.7</td>
</tr>
<tr>
<td>Race and Ethnicity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>American Indian/Alaskan Native</td>
<td>1.4</td>
<td>1.0</td>
<td>1.2</td>
</tr>
<tr>
<td>Asian/Pacific Islander</td>
<td>1.2</td>
<td>0.8</td>
<td>1.0</td>
</tr>
<tr>
<td>Black/African-American</td>
<td>31.6</td>
<td>22.8</td>
<td>27.7</td>
</tr>
<tr>
<td>White/Caucasian</td>
<td>58.1</td>
<td>66.4</td>
<td>61.8</td>
</tr>
<tr>
<td>Multiracial/Other</td>
<td>3.5</td>
<td>5.1</td>
<td>4.2</td>
</tr>
<tr>
<td>Hispanic</td>
<td>4.2</td>
<td>3.9</td>
<td>4.0</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8th Grade or Less</td>
<td>6.6</td>
<td>7.2</td>
<td>6.9</td>
</tr>
<tr>
<td>Some High School</td>
<td>23.2</td>
<td>14.1</td>
<td>19.1</td>
</tr>
<tr>
<td>High School Graduate/ GED</td>
<td>35.0</td>
<td>28.6</td>
<td>32.2</td>
</tr>
<tr>
<td>Some College</td>
<td>21.2</td>
<td>30.6</td>
<td>25.4</td>
</tr>
<tr>
<td>4-year College Graduate</td>
<td>13.9</td>
<td>19.6</td>
<td>16.4</td>
</tr>
<tr>
<td>Marital Status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>17.3</td>
<td>22.0</td>
<td>19.4</td>
</tr>
<tr>
<td>Separated</td>
<td>9.8</td>
<td>9.3</td>
<td>9.6</td>
</tr>
<tr>
<td>Divorced</td>
<td>18.0</td>
<td>23.6</td>
<td>20.5</td>
</tr>
<tr>
<td>Widowed</td>
<td>1.4</td>
<td>5.6</td>
<td>3.3</td>
</tr>
<tr>
<td>Never Married</td>
<td>53.4</td>
<td>39.6</td>
<td>47.2</td>
</tr>
<tr>
<td>Social Support</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spouse or partner</td>
<td>17.7</td>
<td>23.1</td>
<td>20.4</td>
</tr>
<tr>
<td>Other family (parents, children)</td>
<td>41.6</td>
<td>38.4</td>
<td>40.0</td>
</tr>
<tr>
<td>Friends/roommates</td>
<td>11.8</td>
<td>14.2</td>
<td>13.0</td>
</tr>
<tr>
<td>Community/church</td>
<td>7.0</td>
<td>5.5</td>
<td>6.2</td>
</tr>
<tr>
<td>Other</td>
<td>7.0</td>
<td>5.7</td>
<td>6.3</td>
</tr>
<tr>
<td>No one</td>
<td>14.9</td>
<td>13.1</td>
<td>14.1</td>
</tr>
<tr>
<td>Residence/Living Situation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apartment or house</td>
<td>75.3</td>
<td>83.8</td>
<td>79.1</td>
</tr>
<tr>
<td>Supervised housing: halfway house, etc</td>
<td>6.0</td>
<td>3.9</td>
<td>5.1</td>
</tr>
<tr>
<td>School dormitory</td>
<td>0.1</td>
<td>0.4</td>
<td>0.6</td>
</tr>
<tr>
<td>Patient Characteristics</td>
<td>Male (n=773) (%)</td>
<td>Female (n=625) (%)</td>
<td>Total (n=1398) (%)</td>
</tr>
<tr>
<td>-----------------------------------------</td>
<td>------------------</td>
<td>--------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>Hospital or detox center</td>
<td>4.3</td>
<td>5.1</td>
<td>4.7</td>
</tr>
<tr>
<td>Nursing home/assisted living</td>
<td>0.9</td>
<td>0.8</td>
<td>0.9</td>
</tr>
<tr>
<td>Shelter/street</td>
<td>6.8</td>
<td>1.4</td>
<td>4.3</td>
</tr>
<tr>
<td>Jail/prison</td>
<td>0.8</td>
<td>2.0</td>
<td>1.3</td>
</tr>
<tr>
<td>Other</td>
<td>5.1</td>
<td>2.7</td>
<td>4.1</td>
</tr>
<tr>
<td><strong>Employed in the past 30 days</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>63.1</td>
<td>68.4</td>
<td>65.4</td>
</tr>
<tr>
<td>Yes</td>
<td>36.9</td>
<td>31.6</td>
<td>34.6</td>
</tr>
<tr>
<td><strong>Volunteer in the past 30 days</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>90.9</td>
<td>91.0</td>
<td>91.0</td>
</tr>
<tr>
<td>Yes</td>
<td>9.1</td>
<td>9.0</td>
<td>9.0</td>
</tr>
<tr>
<td><strong>Student in the past 30 days</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>94.2</td>
<td>93.3</td>
<td>93.8</td>
</tr>
<tr>
<td>Yes</td>
<td>5.8</td>
<td>6.7</td>
<td>6.2</td>
</tr>
<tr>
<td><strong>Disability Benefits</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>60.3</td>
<td>61.2</td>
<td>60.7</td>
</tr>
<tr>
<td>Yes for medical reasons</td>
<td>12.5</td>
<td>9.6</td>
<td>11.2</td>
</tr>
<tr>
<td>Yes for psychiatric reasons</td>
<td>26.2</td>
<td>28.8</td>
<td>24.4</td>
</tr>
<tr>
<td>Yes for substance abuse</td>
<td>1.0</td>
<td>0.3</td>
<td>0.7</td>
</tr>
<tr>
<td><strong>Primary Payer</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self pay</td>
<td>6.6</td>
<td>4.7</td>
<td>5.8</td>
</tr>
<tr>
<td>BC/BS</td>
<td>1.3</td>
<td>1.5</td>
<td>1.4</td>
</tr>
<tr>
<td>Medicaid</td>
<td>15.3</td>
<td>16.5</td>
<td>15.8</td>
</tr>
<tr>
<td>Medicare</td>
<td>22.2</td>
<td>17.3</td>
<td>20.1</td>
</tr>
<tr>
<td>Commercial</td>
<td>21.1</td>
<td>34.3</td>
<td>26.9</td>
</tr>
<tr>
<td>Uninsured</td>
<td>33.4</td>
<td>25.6</td>
<td>30.0</td>
</tr>
<tr>
<td><strong>Region</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>North</td>
<td>41.8</td>
<td>47.5</td>
<td>44.3</td>
</tr>
<tr>
<td>South</td>
<td>56.9</td>
<td>49.9</td>
<td>53.8</td>
</tr>
<tr>
<td>West</td>
<td>1.3</td>
<td>2.6</td>
<td>1.9</td>
</tr>
<tr>
<td><strong>Geographic Setting</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>21.0</td>
<td>14.2</td>
<td>18.0</td>
</tr>
<tr>
<td>Suburban</td>
<td>56.4</td>
<td>66.9</td>
<td>61.1</td>
</tr>
<tr>
<td>Rural</td>
<td>22.6</td>
<td>18.9</td>
<td>21.0</td>
</tr>
<tr>
<td><strong>Facility Type</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>State or County Psychiatric Hospital</td>
<td>53.2</td>
<td>44.6</td>
<td>49.4</td>
</tr>
<tr>
<td>Private Psychiatric Hospital</td>
<td>31.7</td>
<td>44.8</td>
<td>37.6</td>
</tr>
<tr>
<td>Psychiatric Services in General Hospital</td>
<td>4.3</td>
<td>5.8</td>
<td>4.9</td>
</tr>
<tr>
<td>Outpatient Mental Health Center</td>
<td>6.7</td>
<td>3.2</td>
<td>5.2</td>
</tr>
<tr>
<td>Other</td>
<td>4.1</td>
<td>1.6</td>
<td>3.0</td>
</tr>
<tr>
<td>Patient Characteristics</td>
<td>Male (n=773) (%)</td>
<td>Female (n=625) (%)</td>
<td>Total (n=1398) (%)</td>
</tr>
<tr>
<td>-------------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td><strong>Primary Diagnosis</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schiz/Schizoaff</td>
<td>26.3</td>
<td>23.9</td>
<td>25.2</td>
</tr>
<tr>
<td>Bipolar, Manic, Mixed</td>
<td>9.9</td>
<td>19.5</td>
<td>14.1</td>
</tr>
<tr>
<td>Major/Minor Depression</td>
<td>21.8</td>
<td>30.1</td>
<td>25.4</td>
</tr>
<tr>
<td>Alcohol/Drug</td>
<td>32.9</td>
<td>18.1</td>
<td>26.4</td>
</tr>
<tr>
<td>Dissociative, Anxiety, PTSD</td>
<td>1.7</td>
<td>3.6</td>
<td>2.5</td>
</tr>
<tr>
<td>Other</td>
<td>7.4</td>
<td>4.9</td>
<td>6.3</td>
</tr>
<tr>
<td><strong>Comorbidity – Medical</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>13.1</td>
<td>17.8</td>
<td>15.2</td>
</tr>
<tr>
<td><strong>Comorbidity – Substance Abuse</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>27.2</td>
<td>17.8</td>
<td>26.0</td>
</tr>
<tr>
<td><strong>Comorbidity - Personality Disorder</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>11.4</td>
<td>16.0</td>
<td>3.4</td>
</tr>
<tr>
<td><strong>Comorbidity – Other</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>33.1</td>
<td>36.3</td>
<td>34.5</td>
</tr>
<tr>
<td><strong>Comorbidity Index</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zero</td>
<td>35.8</td>
<td>33.1</td>
<td>34.6</td>
</tr>
<tr>
<td>One</td>
<td>44.9</td>
<td>41.9</td>
<td>43.6</td>
</tr>
<tr>
<td>Two</td>
<td>18.0</td>
<td>22.2</td>
<td>19.9</td>
</tr>
<tr>
<td>Three</td>
<td>1.3</td>
<td>2.7</td>
<td>1.9</td>
</tr>
</tbody>
</table>

**Missing values**

Item frequency distributions of BASIS-24 were generated to assess rates of missing data for each domain at admission and discharge. High rates of missing data can identify items that are confusing, difficult to answer, or inapplicable to respondents (S. V. Eisen et al., 2006). The rate of missing data for each item ranged from <1% to 3.8% for both admission and discharge samples with no major variation among males and females. The highest missing data rate of 3.8% occurred for no more than two items for any of the male and females.

**Floor and ceiling effects**

Extensive floor and ceiling effects can indicate insensitivity of the instrument to individual differences in symptom levels at the extreme ends of the continuum, or inapplicability.
of items to the sample (S. V. Eisen et al., 2006). For each subscale, floor effects (worst possible functioning) occurred for no more than 5% of admission and discharge within each male/female sample. Ceiling effects (best possible functioning) were infrequent for common domains such as depression and functioning, occurring for up to 3.6% of admission, and up to 9.4% of discharge for both genders. Ceiling effects were more common for infrequently occurring domains such as self-harm, with 41.6% of the total sample at admission and 68.7% at discharge reporting no thoughts of self-harm during the past week. As expected, the discharge sample generally had higher rates of ceiling effects and admission sample had higher rates of floor effects with relatively little variation among male and female groups.

**Skewness and Kurtosis**

Before conducting any statistical analyses, the distribution of BASIS-24 data were examined for departures from normality by examining standardized skewness and kurtosis estimates. This testing was important for confirmatory factor analysis. Because the structural equation model procedures used in this study can produce distorted results when the normality assumption is severely violated (Curran, West, & Finch, 1996), the normality of each subscale was investigated in terms of skewness and kurtosis. Table 2 presents the results of the estimates of skewness and kurtosis, for the six subscales and total score on BASIS-24 admission and discharge scores. According to the guidelines of severe non-normality (i.e., skew>2; kurtosis>7) proposed by (West, Finch, & Curran, 1995), the normality assumption of all the variables were met.
Table 3: Descriptive Statistics of subscales for BASIS-24 by Gender

<table>
<thead>
<tr>
<th></th>
<th>Admission</th>
<th></th>
<th>Discharge</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male (n=773)</td>
<td>Female (n=625)</td>
<td>Total (n=1398)</td>
<td>Male (n=773)</td>
<td>Female (n=625)</td>
<td>Total (n=1398)</td>
</tr>
<tr>
<td><strong>Depression</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>2.11 (1.13)</td>
<td>2.36 (1.12)</td>
<td>2.22 (1.13)</td>
<td>1.10 (0.86)</td>
<td>1.26 (0.89)</td>
<td>1.18 (0.88)</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.19</td>
<td>-0.49</td>
<td>-0.32</td>
<td>0.81</td>
<td>0.63</td>
<td>0.73</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>0.18</td>
<td>-0.78</td>
<td>-0.96</td>
<td>0.25</td>
<td>-0.08</td>
<td>0.07</td>
</tr>
<tr>
<td>% Floor</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>% Ceiling</td>
<td>3.8</td>
<td>3.4</td>
<td>3.6</td>
<td>10.3</td>
<td>8.3</td>
<td>9.4</td>
</tr>
<tr>
<td><strong>Interpersonal</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>1.80 (1.07)</td>
<td>1.71 (1.05)</td>
<td>1.76 (1.06)</td>
<td>1.29 (1.05)</td>
<td>1.26 (1.02)</td>
<td>1.28 (1.04)</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.16</td>
<td>0.24</td>
<td>0.20</td>
<td>0.78</td>
<td>0.89</td>
<td>0.83</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>0.18</td>
<td>-0.70</td>
<td>-0.77</td>
<td>-0.05</td>
<td>0.28</td>
<td>0.08</td>
</tr>
<tr>
<td>% Floor</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>% Ceiling</td>
<td>5.8</td>
<td>7.00</td>
<td>6.4</td>
<td>14.4</td>
<td>13.4</td>
<td>13.9</td>
</tr>
<tr>
<td><strong>Self Harm</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>1.07 (1.24)</td>
<td>1.24 (1.27)</td>
<td>1.15 (1.25)</td>
<td>0.36 (0.71)</td>
<td>0.46 (0.78)</td>
<td>.405 (0.74)</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.90</td>
<td>0.63</td>
<td>0.77</td>
<td>2.42</td>
<td>1.94</td>
<td>2.18</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>0.18</td>
<td>-0.83</td>
<td>-0.61</td>
<td>6.63</td>
<td>3.88</td>
<td>5.18</td>
</tr>
<tr>
<td>% Floor</td>
<td>4.70</td>
<td>4.60</td>
<td>4.60</td>
<td>0.80</td>
<td>0.80</td>
<td>0.80</td>
</tr>
<tr>
<td>% Ceiling</td>
<td>44.0</td>
<td>38.7</td>
<td>41.6</td>
<td>70.9</td>
<td>65.9</td>
<td>68.7</td>
</tr>
<tr>
<td><strong>Emotional Lability</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>1.86 (1.11)</td>
<td>2.09 (1.13)</td>
<td>1.96 (1.13)</td>
<td>1.22 (0.93)</td>
<td>1.32 (0.96)</td>
<td>1.26 (0.94)</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.02</td>
<td>-0.21</td>
<td>-0.10</td>
<td>0.46</td>
<td>0.38</td>
<td>0.43</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>0.18</td>
<td>-0.87</td>
<td>-0.89</td>
<td>-0.36</td>
<td>-0.35</td>
<td>-0.36</td>
</tr>
<tr>
<td>% Floor</td>
<td>3.90</td>
<td>3.20</td>
<td>3.6</td>
<td>0.6</td>
<td>1.10</td>
<td>0.90</td>
</tr>
<tr>
<td>% Ceiling</td>
<td>8.70</td>
<td>6.10</td>
<td>7.5</td>
<td>16.0</td>
<td>14.9</td>
<td>15.5</td>
</tr>
<tr>
<td><strong>Psychosis</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>1.09 (1.16)</td>
<td>1.13 (1.12)</td>
<td>1.11 (1.15)</td>
<td>0.65 (0.88)</td>
<td>0.64(0.86)</td>
<td>0.65 (0.87)</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.84</td>
<td>0.73</td>
<td>0.79</td>
<td>1.43</td>
<td>1.41</td>
<td>1.42</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>0.18</td>
<td>-0.59</td>
<td>-0.56</td>
<td>1.47</td>
<td>1.46</td>
<td>1.46</td>
</tr>
<tr>
<td>% Floor</td>
<td>1.20</td>
<td>1.00</td>
<td>1.1</td>
<td>0.40</td>
<td>0.50</td>
<td>0.40</td>
</tr>
<tr>
<td>% Ceiling</td>
<td>31.70</td>
<td>29.6</td>
<td>30.8</td>
<td>47.5</td>
<td>48.2</td>
<td>47.8</td>
</tr>
<tr>
<td><strong>Substance Abuse</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>1.43 (1.30)</td>
<td>0.98 (1.23)</td>
<td>1.23 (1.29)</td>
<td>0.91 (0.97)</td>
<td>0.63 (0.92)</td>
<td>0.79 (0.96)</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.34</td>
<td>1.01</td>
<td>0.61</td>
<td>0.84</td>
<td>1.50</td>
<td>1.10</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>0.18</td>
<td>-0.33</td>
<td>-1.04</td>
<td>-0.20</td>
<td>1.46</td>
<td>0.30</td>
</tr>
<tr>
<td>% Floor</td>
<td>2.3</td>
<td>2.1</td>
<td>2.2</td>
<td>0.3</td>
<td>0.8</td>
<td>0.5</td>
</tr>
<tr>
<td>% Ceiling</td>
<td>30.4</td>
<td>45.6</td>
<td>37.2</td>
<td>37.0</td>
<td>54.2</td>
<td>44.7</td>
</tr>
</tbody>
</table>
**Mean and Standard Deviation**

At admission, BASIS-24 overall scores for the entire sample ranged from 0 to 4, with a mean of 1.85 and a standard deviation (SD) of 0.83. The mean and SD values for the male subsample were 1.79 and 0.83. The corresponding values for females were 1.92 and 0.83. Consistent with other prior BASIS-32 and BASIS-24 studies, the respondents reported moderate degree of difficulty in each of the six areas. The greatest amount of difficulty was reported on subscales of depression and emotional lability, and less difficulty was reported for self-harm and psychosis. As expected, the female sample consistently reported significantly higher symptom/problem levels than the male population in the overall summary score, as well as in the domains of depression, self-harm, and psychosis (Table 3). Men, however, reported higher levels of difficulty for interpersonal and substance abuse subscales.

In order to examine whether the sample differed by gender, a one-way analysis of variance (ANOVAs) was computed. The results showed significant differences on 4 subscales, depression, self-harm, emotional lability and substance abuse as well as the overall score. Though the F value and the level of significance reduced considerably at discharge for most of the subscales and the overall score, the trend was inconsistent in the domain emotional lability. This can be either due to the measurement error or a real difference between the male and female population.

**Reliability**

For each BASIS-24 scale and the overall scale, internal consistency reliability estimates (Cronbach’s alpha), and item-total correlations were computed for male, female and total samples. These analyses were done separately for admission and discharge samples to determine
the instrument’s psychometric properties and potential utility for outcome assessment. As shown in Table 5, the reliability coefficients ranged from 0.73 to 0.89.

The overall internal consistency of admission test scores for each subscale was comparable to that obtained in the original validation studies (S. V. Eisen et al., 2006). Most of the coefficients were also above the 0.80 cutoff that is typically considered acceptable for general research purposes ((Henson, 2001); (Loo, 2001). Internal consistency within each of the gender subgroups examined herein was also largely acceptable. Internal consistency reliability (Cronbach’s alpha) coefficients exceeded 0.70 for all six domains and for male and female for the admission and discharge samples (Table 5).

Table 4: One-way analysis of variance of BASIS-24 Subscale and Overall Score by Gender

<table>
<thead>
<tr>
<th>Subscale</th>
<th>Admission Score</th>
<th>F value</th>
<th>Discharge Score</th>
<th>F value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depression</td>
<td>Male</td>
<td>2.11</td>
<td>17.30***</td>
<td>1.10</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>2.36</td>
<td></td>
<td>1.26</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>2.22</td>
<td></td>
<td>1.17</td>
</tr>
<tr>
<td>Interpersonal</td>
<td>Male</td>
<td>1.80</td>
<td>2.44</td>
<td>1.29</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>1.71</td>
<td></td>
<td>1.26</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>1.76</td>
<td></td>
<td>1.28</td>
</tr>
<tr>
<td>Self Harm</td>
<td>Male</td>
<td>1.07</td>
<td>5.90*</td>
<td>.36</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>1.24</td>
<td></td>
<td>.46</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>1.15</td>
<td></td>
<td>.40</td>
</tr>
<tr>
<td>Emotional</td>
<td>Male</td>
<td>1.86</td>
<td>14.62***</td>
<td>1.22</td>
</tr>
<tr>
<td>Lability</td>
<td>Female</td>
<td>2.09</td>
<td></td>
<td>1.32</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>1.96</td>
<td></td>
<td>1.26</td>
</tr>
<tr>
<td>Psychosis</td>
<td>Male</td>
<td>1.09</td>
<td>0.37</td>
<td>.65</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>1.13</td>
<td></td>
<td>.64</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>1.11</td>
<td></td>
<td>.65</td>
</tr>
<tr>
<td>Substance</td>
<td>Male</td>
<td>1.43</td>
<td>44.34***</td>
<td>.91</td>
</tr>
<tr>
<td>Abuse</td>
<td>Female</td>
<td>.98</td>
<td></td>
<td>.63</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>1.23</td>
<td></td>
<td>.79</td>
</tr>
<tr>
<td>Overall</td>
<td>Male</td>
<td>1.79</td>
<td>8.49***</td>
<td>1.03</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>1.92</td>
<td></td>
<td>1.10</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>1.85</td>
<td></td>
<td>1.06</td>
</tr>
</tbody>
</table>
The internal consistencies for the six subscales ranged from 0.73 to 0.89 for the entire sample, with ranges of 0.89 to 0.73 for males and 0.89 to 0.77 for females. The emotional lability subscale at discharge yielded somewhat moderate internal consistency, \( \alpha = 0.73 \), as it did in a study by (Capaldi & Rothbart, 1992). Considering the large sample sizes and limited number of items for each subscale, these reliability estimates are considered from being adequate to quite good. All subscales were deemed acceptable for research purposes, which were above 0.70 (Henson, 2001).

Table 5: Standardized Internal Consistency Reliability (Cronbach’s Alpha) Coefficients

<table>
<thead>
<tr>
<th></th>
<th>Admission</th>
<th>Discharge</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male (n=773)</td>
<td>Female (n=625)</td>
</tr>
<tr>
<td>Depression</td>
<td>.870</td>
<td>.883</td>
</tr>
<tr>
<td>Interpersonal</td>
<td>.817</td>
<td>.824</td>
</tr>
<tr>
<td>Self Harm</td>
<td>.887</td>
<td>.881</td>
</tr>
<tr>
<td>Emotional Lability</td>
<td>.757</td>
<td>.783</td>
</tr>
<tr>
<td>Psychosis</td>
<td>.780</td>
<td>.778</td>
</tr>
<tr>
<td>Substance Abuse</td>
<td>.861</td>
<td>.891</td>
</tr>
<tr>
<td>Overall</td>
<td>.879</td>
<td>.888</td>
</tr>
</tbody>
</table>

Correlation

The correlation matrix among the subscale scores of BASIS-24 is presented in Tables 6 (admission) and 7 (discharge) for both males and females. Factor correlations provide information on the relationship between each subscale and the overall score. As can be seen in Tables 5 and 6, correlations among the subscales ranged from 0.005 to 0.62. Most of the
correlations between factors were positive and significant, although some were quite small.

These correlations demonstrate a logical pattern of relationships. For example, the largest correlations between factors centered on the relationships between depression and emotional lability. No correlations among constructs the exceeded the moderate range, however, indicating that the factors were tapping related, but conceptually distinct, dimensions of functioning.

Table 6: BASIS-24 Overall and Subscale Correlations at Admission

<table>
<thead>
<tr>
<th>Admission</th>
<th>Gender</th>
<th>Depression</th>
<th>Interpersonal</th>
<th>Self-Harm</th>
<th>Emotional Lability</th>
<th>Psychosis</th>
<th>Substance Abuse</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depression</td>
<td>Female</td>
<td>-</td>
<td>0.318**</td>
<td>0.559**</td>
<td>0.628**</td>
<td>0.256**</td>
<td>0.273**</td>
<td>0.923**</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>0.319**</td>
<td>-</td>
<td>0.248**</td>
<td>0.22**</td>
<td>0.205**</td>
<td>0.05**</td>
<td>0.486**</td>
</tr>
<tr>
<td>Interpersonal</td>
<td>Female</td>
<td>0.318**</td>
<td>-</td>
<td>0.283**</td>
<td>0.23**</td>
<td>0.258**</td>
<td>0.049</td>
<td>0.493**</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>0.568**</td>
<td>0.248**</td>
<td>-</td>
<td>0.429**</td>
<td>0.292**</td>
<td>0.037</td>
<td>0.686**</td>
</tr>
<tr>
<td>Self Harm</td>
<td>Female</td>
<td>0.559**</td>
<td>0.283**</td>
<td>-</td>
<td>0.461**</td>
<td>0.206**</td>
<td>0.157**</td>
<td>0.685**</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>0.536**</td>
<td>0.22**</td>
<td>0.429**</td>
<td>-</td>
<td>0.434**</td>
<td>0.273**</td>
<td>0.714**</td>
</tr>
<tr>
<td>Emotional</td>
<td>Female</td>
<td>0.628**</td>
<td>0.237**</td>
<td>0.461**</td>
<td>-</td>
<td>0.346**</td>
<td>0.347**</td>
<td>0.768**</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>0.265**</td>
<td>0.205**</td>
<td>0.292**</td>
<td>0.434**</td>
<td>-</td>
<td>0.28**</td>
<td>0.452**</td>
</tr>
<tr>
<td>Psychosis</td>
<td>Female</td>
<td>0.256**</td>
<td>0.258**</td>
<td>0.206**</td>
<td>0.346**</td>
<td>-</td>
<td>0.005**</td>
<td>0.419**</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>0.267**</td>
<td>0.055**</td>
<td>0.037**</td>
<td>0.273**</td>
<td>0.028</td>
<td>-</td>
<td>0.325**</td>
</tr>
<tr>
<td>Substance</td>
<td>Female</td>
<td>0.273**</td>
<td>0.049</td>
<td>0.157**</td>
<td>0.347**</td>
<td>0.005</td>
<td>-</td>
<td>0.354**</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>0.929**</td>
<td>0.486</td>
<td>0.686</td>
<td>0.714</td>
<td>0.452**</td>
<td>0.325**</td>
<td>-</td>
</tr>
<tr>
<td>Overall</td>
<td>Female</td>
<td>0.932**</td>
<td>0.493</td>
<td>0.685**</td>
<td>0.768**</td>
<td>0.419</td>
<td>0.354**</td>
<td>-</td>
</tr>
</tbody>
</table>

**statistically significant; p < .001
Table 7: BASIS-24 Overall and Subscale Correlations* at Discharge

<table>
<thead>
<tr>
<th>Discharge</th>
<th>Gender</th>
<th>Depression</th>
<th>Interpersonal</th>
<th>Self-Harm</th>
<th>Emotional Lability</th>
<th>Psychosis</th>
<th>Substance Abuse</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depression</td>
<td>Male</td>
<td>-</td>
<td>0.333**</td>
<td>0.512**</td>
<td>0.575**</td>
<td>0.313</td>
<td>0.258**</td>
<td>0.921**</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>-</td>
<td>0.275**</td>
<td>0.531**</td>
<td>0.622**</td>
<td>0.314</td>
<td>0.202**</td>
<td>0.929**</td>
</tr>
<tr>
<td>Interpersonal</td>
<td>Male</td>
<td>0.333**</td>
<td>-</td>
<td>0.202**</td>
<td>0.231**</td>
<td>0.176**</td>
<td>0.088*</td>
<td>0.53**</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>0.275**</td>
<td>-</td>
<td>0.149**</td>
<td>0.224**</td>
<td>0.201**</td>
<td>0.053</td>
<td>0.476**</td>
</tr>
<tr>
<td>Self Harm</td>
<td>Male</td>
<td>0.512**</td>
<td>0.202**</td>
<td>-</td>
<td>0.399**</td>
<td>0.351**</td>
<td>0.146**</td>
<td>0.622**</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>0.531**</td>
<td>0.149**</td>
<td>-</td>
<td>0.384**</td>
<td>0.324**</td>
<td>0.11**</td>
<td>0.618**</td>
</tr>
<tr>
<td>Emotional Lability</td>
<td>Male</td>
<td>0.575**</td>
<td>0.231**</td>
<td>0.399**</td>
<td>-</td>
<td>0.38**</td>
<td>0.314**</td>
<td>0.731**</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>0.622**</td>
<td>0.224**</td>
<td>0.384**</td>
<td>-</td>
<td>0.389**</td>
<td>0.25**</td>
<td>0.757**</td>
</tr>
<tr>
<td>Psychosis</td>
<td>Male</td>
<td>0.313**</td>
<td>0.176**</td>
<td>0.351**</td>
<td>0.38**</td>
<td>-</td>
<td>0.06**</td>
<td>0.48**</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>0.314**</td>
<td>0.201**</td>
<td>0.324**</td>
<td>0.389**</td>
<td>-</td>
<td>0.07</td>
<td>0.483**</td>
</tr>
<tr>
<td>Substance Abuse</td>
<td>Male</td>
<td>0.258**</td>
<td>0.088*</td>
<td>0.146**</td>
<td>0.314**</td>
<td>0.06</td>
<td>-</td>
<td>0.351**</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>0.202**</td>
<td>0.053**</td>
<td>0.11**</td>
<td>0.25**</td>
<td>0.07</td>
<td>-</td>
<td>0.3**</td>
</tr>
<tr>
<td>Overall</td>
<td>Male</td>
<td>0.053**</td>
<td>0.11**</td>
<td>0.622**</td>
<td>0.731</td>
<td>0.48**</td>
<td>0.351**</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>0.929**</td>
<td>0.476**</td>
<td>0.618**</td>
<td>0.757**</td>
<td>0.483**</td>
<td>0.3**</td>
<td>-</td>
</tr>
</tbody>
</table>

*statistically significant; p < .001
Single-Group Confirmatory Factor Analyses of BASIS-24 by Gender

This section examined the factor structure of BASIS-24 by gender. Based on the prior research of BASIS-24 factor structure by Eisen et al (S. V. Eisen et al., 2004), it was hypothesized that the six-factor structure of BASIS-24 would be replicated across gender groups at admission and discharge. To establish a well-fitting baseline model for each group separately prior to testing for factorial invariance, separate analysis was conducted for: a) the entire sample at admission; b) the entire sample at discharge; c) the male sample at admission; and d) the female sample at discharge.

Table 8: Single group confirmatory factor analysis of the BASIS-24 by Gender

<table>
<thead>
<tr>
<th>Model</th>
<th>df</th>
<th>$\chi^2$</th>
<th>RMSEA</th>
<th>SRMR</th>
<th>CFI</th>
<th>NNFI</th>
<th>GFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1a</td>
<td>237</td>
<td>1333.964</td>
<td>0.058</td>
<td>0.065</td>
<td>0.976</td>
<td>0.973</td>
<td>0.982</td>
</tr>
<tr>
<td>Entire sample (Admission)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 1d</td>
<td>237</td>
<td>1344.984</td>
<td>0.058</td>
<td>0.074</td>
<td>0.978</td>
<td>0.974</td>
<td>0.973</td>
</tr>
<tr>
<td>Entire sample (Discharge)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 2a</td>
<td>237</td>
<td>862.309</td>
<td>0.058</td>
<td>0.070</td>
<td>0.975</td>
<td>0.971</td>
<td>0.977</td>
</tr>
<tr>
<td>Male sample (Admission)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 2d</td>
<td>237</td>
<td>884.858</td>
<td>0.059</td>
<td>0.078</td>
<td>0.977</td>
<td>0.973</td>
<td>0.969</td>
</tr>
<tr>
<td>Male sample (Discharge)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 3a</td>
<td>237</td>
<td>679.151</td>
<td>0.055</td>
<td>0.067</td>
<td>0.980</td>
<td>0.976</td>
<td>0.983</td>
</tr>
<tr>
<td>Female sample (Admission)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 3d</td>
<td>237</td>
<td>689.038</td>
<td>0.055</td>
<td>0.079</td>
<td>0.980</td>
<td>0.977</td>
<td>0.971</td>
</tr>
<tr>
<td>Female sample (Discharge)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

RSMSEA, Root Mean Square Error of Approximation; SRMR, standardized root mean square residual; CFI, Comparative Fit Index; GFI, Goodness of Fit Index; NNFI: Non-Normed Fit Index or TLI: Tucker Lewis Index

RMSEA <0.05, SRMR <0.05, CFI >.95, NNFI/TLI >0.95 indicate a good fit. RMSEA <0.080, SRMR<0.100, and CFI>0.90, NNFI/TLI>0.90 indicate an acceptable fit (L. Hu & Bentler, 1999, Schermelleh-Engel, Moosbrugger, & Müller, 2003; Steenkamp,Jan-Benedict E.M. 1998)
Entire sample (Model 1a & 1d)

Table 8 shows that the hypothesized six factor model fit for BASIS-24 reasonably well to the entire sample data at admission and discharge. Separate analyses were conducted for the combined sample of male and females at admission and discharge. Though the initial fit of the hypothesized model for the entire sample at admission was poor from a statistical standpoint considering the chi-square statistic ($\chi^2=1333.964$, df=237, $P<.01$) and the relative likelihood ratio ($\chi^2/df=5.629$), the other fit indices (RMSEA=0.058, SRMR=0.065, NNFI=0.973, CFI=0.976, and GFI=0.982), all indicated an acceptable fit of the model from a practical perspective. Similar analysis was conducted for the discharge sample and indicated the model fit was adequate ($\chi^2=1344.984$, df=237, $P<0.001$) and the relative likelihood ratio ($\chi^2/df=5.625$), the other fit indices (RMSEA=0.058, SRMR=0.074, NNFI=0.974, CFI=0.978, and GFI=0.973).

Male sample (Model 2a & 2d)

Analysis of the male sample at admission and discharge revealed that the six factor model was a reasonable fit to the data. All chi-square values were significant ($p<0.00$) for the male sample at admission and discharge. The female group had slightly lower relative likelihood ratios than the male group, although the differences were not statistically significant. The goodness-of-fit indices, RMSEA, SRMR, NNFI, and GFI all exceeded 0.90, a preferred criterion for a close fit. The confirmatory factor analysis showed the original six-factor model of BASIS-24 fit provided a modest to excellent fit to the male group at admission (RMSEA=0.058, SRMR=0.070, NNFI=0.971, CFI=0.975, and GFI=0.977) with the relative likelihood ratio ($\chi^2/df=3.638$). Similar analysis were conducted for the discharge sample and the model was adequate (RMSEA=0.059, SRMR=0.078, NNFI=0.973, CFI=0.977, and GFI=0.969) with the relative likelihood ratio ($\chi^2/df=3.734$).
Female sample (Model 3a & 3d)

The goodness-of-fit indices for the female group at admission (RMSEA=0.055, SRMR=0.067, NNFI=0.976, CFI=0.980, and GFI=0.983), and at discharge (RMSEA=0.055, SRMR=0.079, NNFI=0.977, CFI=0.980, and GFI=0.971) also revealed that the six factor model fit reasonably well to the data. The female group had slightly lower relative likelihood ratios ($\chi^2/df=2.86$ at admission and $\chi^2/df=2.907$ at discharge) than the male group, although the differences were not statistically significant. The goodness-of-fit indices, RMSEA, NNFI, CFI, and GFI all exceeded 0.90, a preferred criterion for a close fit.

In summary, evidence from the above confirmatory factor analyses showed the original six-factor model of BASIS-24 fit provided a modest to excellent fit to the entire sample of inpatient psychiatric patients at admission and discharge as a whole, as well as across genders.

Multiple-Group Confirmatory Factor Analysis of BASIS-24 by Gender

Although the original model had reasonable fit, the factorial invariance of the model across genders was explored in accordance with the original question. It was similarly hypothesized that the items of BASIS-24 would be found to demonstrate factorial invariance across gender and that the six factor structure would fit well. The next step in the analysis tested whether the six-factor model was structurally invariant across male and females. Multi-group confirmatory factor analysis was used to test four types of measurement invariance models of BASIS-24 across gender. The four types of measurement invariance tests used were configural or pattern factorial invariance, weak metric invariance, strong metric invariance, and strict or complete metric invariance. The factorial invariance is generally assessed by the $\Delta \chi^2$ value between two nested models. Two separate criterion were used to compare the nested multi-group models: a) a non-significant chi-square difference test; and b) a delta CFI value less than
or -0.01 was used to make final decisions about whether weak, strong, and strict measurement invariance models hold.

Model 4a and 4d - Test of Configural Factor Invariance

This model tested the hypothesis that an a priori pattern of free and fixed factor loadings imposed on the measures was equivalent across groups. Thus the first hypothesis tested whether the six factor structure is invariant across gender groups at admission and discharge. To test configural invariance of the six factor BASIS-24 model and whether it’s a good fit of the data across genders, analyses were conducted separately for admission and discharge samples. All BASIS-24 items were specified to load on the related BASIS-24 factors or subscales. The factor loadings, factor covariance and the error variances were freely estimated. The \( \chi^2 \) and descriptive values for the admission and discharge groups are presented in Table 9. According to the practical fit indices, model 4a (RMSEA=0.055, CFI=0.976, IFI=.976), and model 4d (RMSEA=.053, CFI=.980, IFI=.980), fit the data, indicating that the six BASIS-24 factors are psychometrically distinct constructs at admission and discharge. In comparing Model 4a (admission) and 4d (discharge), Model 4d fits better in terms of RMSEA and other fit indices. The goodness-of-fit indices in model 4a and 4d in Table 9 indicate that the six factor model is plausible across the admission and discharge groups. Hence, this was eligible for further tests of stricture measurement invariance.
<table>
<thead>
<tr>
<th>Model</th>
<th>df</th>
<th>$\chi^2$</th>
<th>CFI</th>
<th>IFI</th>
<th>NNFI</th>
<th>RMSEA</th>
<th>GFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 4a (Admission)</td>
<td>509</td>
<td>1572.235</td>
<td>0.976</td>
<td>0.976</td>
<td>0.974</td>
<td>0.055</td>
<td>0.980</td>
</tr>
<tr>
<td>Configural Invariance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No constraints</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 4d: (Discharge)</td>
<td>509</td>
<td>1473.337</td>
<td>0.980</td>
<td>0.980</td>
<td>0.979</td>
<td>0.053</td>
<td>0.978</td>
</tr>
<tr>
<td>Configural Invariance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No constraints</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 5a: (Admission)</td>
<td>527</td>
<td>1853.643</td>
<td>0.970</td>
<td>0.970</td>
<td>0.969</td>
<td>0.061</td>
<td>0.946</td>
</tr>
<tr>
<td>Weak Invariance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Factor loadings constrained</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 5d: (Discharge)</td>
<td>527</td>
<td>1625.315</td>
<td>0.977</td>
<td>0.977</td>
<td>0.976</td>
<td>0.055</td>
<td>0.957</td>
</tr>
<tr>
<td>Weak Invariance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Factor loadings constrained</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 6a: (Admission)</td>
<td>532</td>
<td>1936.388</td>
<td>0.968</td>
<td>0.968</td>
<td>0.967</td>
<td>0.062</td>
<td>0.958</td>
</tr>
<tr>
<td>Strong Invariance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Factor loadings &amp; variances and covariances constrained</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 6d: (Discharge)</td>
<td>532</td>
<td>1633.873</td>
<td>0.977</td>
<td>0.977</td>
<td>0.977</td>
<td>0.055</td>
<td>0.957</td>
</tr>
<tr>
<td>Strong Invariance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Factor loadings &amp; variances and covariances constrained</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 7a: (Admission)</td>
<td>537</td>
<td>1969.593</td>
<td>0.968</td>
<td>0.968</td>
<td>0.967</td>
<td>0.062</td>
<td>0.958</td>
</tr>
<tr>
<td>Strict Invariance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Factor loadings &amp; variances, and covariances error variance constrained</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 7d: (Discharge)</td>
<td>537</td>
<td>1651.918</td>
<td>0.977</td>
<td>0.977</td>
<td>0.976</td>
<td>0.055</td>
<td>0.957</td>
</tr>
<tr>
<td>Strict Invariance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Factor loadings &amp; variances, and covariances error variance constrained</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model Comparison (Admission Sample)</td>
<td>Δdf</td>
<td>Δχ²</td>
<td>ΔCFI</td>
<td>ΔIFI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------------------------</td>
<td>-----</td>
<td>-------</td>
<td>------</td>
<td>------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 4a vs. 5a: (Admission)</td>
<td>18</td>
<td>281.408</td>
<td>-0.006</td>
<td>-0.006</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 4d vs. 5d (Discharge)</td>
<td>18</td>
<td>151.979</td>
<td>-0.003</td>
<td>-0.003</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 4a vs. 6a: (Admission)</td>
<td>23</td>
<td>364.153</td>
<td>-0.008</td>
<td>-0.008</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 4d vs. 6d (Discharge)</td>
<td>23</td>
<td>160.537</td>
<td>-0.003</td>
<td>-0.003</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 4a vs. 7a (Admission)</td>
<td>28</td>
<td>397.358</td>
<td>-0.008</td>
<td>-0.008</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 4d vs. 7d (Discharge)</td>
<td>28</td>
<td>178.582</td>
<td>-0.003</td>
<td>-0.003</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

RMSSEA, Root Mean Square Error of Approximation; CFI, Comparative Fit Index; IFI, Incremental Fit Index; GFI, Goodness of Fit Index; NNFI, Non-Normed Fit Index or TLI, Tucker Lewis Index.

RMSEA <0.05, SRMR <0.05, and NNFI/TLI >0.95 indicate a good fit. RMSEA <0.080, SRMR <0.100, and NNFI/TLI >0.90 indicate an acceptable fit (Schermelleh-Engel, Moosbrugger, & Müller, 2003; Steenkamp, Jan-Benedict E.M. 1998).

Fit Index; GFI, Goodness of Fit Index; Δχ² = differences in chi-squares of models; Δdf = differences in degree of freedom of models; ΔCFI = model differences in Comparative Fit Index. ΔIFI = model differences in Bollen’s Incremental Fit Index (IFI).

CFI decrease = CFI unconstraint model – CFI constraint model; CFI decrease can have a negative value (Cheung & Rensvold, 2002). For configural invariance, if RMSEA > cut-off, then the invariance hypothesis is accepted. For all other invariance types, if CFI decrease > cut-off -0.01, the invariance hypothesis is rejected (Alpha = 0.01).
Despite the indication of good model fit by RMSEA and CFI, $\chi^2$ rejected both configural models. If the decision rule had been based on $\chi^2$, it would have suggested termination for further tests for these comparisons. However, based on the other fit indices configural invariance is achieved, supporting the hypothesis that male and female sample identified the same cognitive domains underlying BASIS-24 responses, conceptualizing the constructs in the same way. In other words, the data produced the same number of factors and each factor was defined by the same items. Parameter estimates of Model a1 and a2 for admission and discharge are presented in Table 9. The results reported shows that despite gender differences, the patients provided the same meaning to the constructs assessed by the instrument items. Configural invariance serves as a useful baseline model to compare more restrictive models. The common metric completely standardized factor loadings were all high ranging from 0.60 to .95 with a mean of 0.79. Thus the first hypothesis is tenable that the six factor structure of BASIS-24 is invariant across gender groups.

Model 5a and 5d- Test of weak or metric invariance

In this model, in addition to the configural constraints, the factor loadings are constrained to be equal across gender groups, but no other equality constraints are imposed. With weak invariance, the factor loadings for males and females are constrained to determine the equality of measurement between genders with factor correlation and error variances. To test for metric or weak factorial invariance, the factor pattern coefficients were constrained equal. These constraints increased the $\chi^2$ value from 1572.235 to 1853.643 ($p<.01$) for admission sample and 1625.314 ($P<.01$) for discharge sample gaining 18 degrees of freedom (Table 9). Because the weak invariance model (Model 5a) is nested within the baseline model (Model 4a), a chi-square difference test ($\Delta \chi^2$) and delta comparative fit index ($\Delta$CFI) test were performed. The $\chi^2$
differences, 281.408 for admission sample, and 151.979 for discharge sample were statistically significant, with 18 degrees of freedom and the model difference showed that the statistical difference was significant. The $\Delta$CFI for weak measurement invariance was computed by subtracting the CFI from the configural invariance model from that of the weak variance model. The results of the goodness-of-fit indices for the admission sample (RMSEA=.060, CFI=.971, IFI=.970, NNFI=.959) and discharge sample (RMSEA=.055, CFI=.977, IFI=.977) and particularly the $\Delta$CFI value for each model showed that all of the multi-group models were equivalent in terms of goodness of fit examined through the $\Delta$CFI and that all of them show an excellent fit to the data (Table 9). The decrease in CFI was 0.006 for the admission data and 0.003 for the discharge data and both fell below the 0.009 cut-off, indicating weak factor invariance. This model for both the admission and discharge data (5a and 5d) demonstrated reasonably good fit of indices indicating that it adequately represented the data. The equivalence of the strength of the item-factor relations demonstrated that the construct was manifested in the same away across gender. Thus, the second hypothesis underlying construct framework of BASIS-24 subscales did not differ for males and females. The factor loadings of BASIS-24 are invariant across gender. The strength of the relationship based on the factor loadings between each item and its underlying construct is the same for both groups (weak or metric invariance). If $\Delta\chi^2$ was employed for decision making for weak invariance, it would have rejected both comparisons. However CFI >0.90 is considered as the cut-off commonly referred to in the recent literature.

*Model 6a and 6d- Test of Strong factorial (scalar or item) invariance*

This model examines whether the group comparisons are meaningful. Invariant factor loadings are not enough to compare scale scores across groups. In this model (Model 6a and 6d),
factorial invariance constrains factor loadings and intercepts to verify if the factor model of factor correlations is equal between genders. As with weak factorial invariance, this specification implies that the measurement of the latent variables is the same across groups. Furthermore, the invariance in the intercepts allows for evaluating group differences in subscales across groups. Thus the third hypothesis tested was a more restrictive model (Model 6a and 6d) with equal factor loadings and intercepts.

The additional constraints increased the $\chi^2$ value from 1572.235 to 1936.388 ($p<.01$) for the admission data and from 1473.337 to 1633.873 ($p<.01$) for the discharge data, gaining 23 degrees of freedom (Table 9). Because the strong invariance model (Model 6a and 6d) is nested within the baseline model (Model 4a and 4d), a chi-square difference test ($\Delta\chi^2$) and delta comparative fit index ($\Delta$CFI) test were performed between these models. The $\chi^2$ differences, ($\chi^2=364.153$, $p<.01$) for admission, and 160.537, $p<.01$ for discharge with 23 degrees of freedom, indicated a statistically significant difference and hence strong invariance was not supported. However, the goodness of fit indices between the configural invariance model and the strong invariance model and particularly the $\Delta$CFI value for each model showed that all of the multi-group models were equivalent for both male and female samples at admission and discharge in terms of goodness of fit examined through the $\Delta$CFI, and that all of them show an excellent fit to the data (Table 9). The decrease in CFI was 0.008 at admission and 0.003 at discharge and both fell below the 0.009 cut-off, indicating strong factor invariance. Similar to the weak invariance model, the strong invariance model for admission data (6a) demonstrated reasonable good fit of indexes and the model for discharge data (5d) demonstrated acceptable fit indices, indicating that it adequately represented the data, and supported strong invariance. Thus, the third hypothesis underlying the cross cultural differences and relationships among constructs
(covariances) are the same across male and female. If $\Delta \chi^2$ was employed for decision making for strong invariance, it would have rejected both comparisons. Constraining the intercepts to be equal, the RMSEA and CFI were reduced slightly, but the CFI >0.90 is considered as the cut-off commonly referred to in the recent literature.

**Model 7a and 7d: Test of Strict factorial (complete) invariance**

This model extends the previous model by invoking the additional constraint and unique variances across groups. Strict factor invariance checks the equality of all parameters: factor loadings; intercepts; and error or residual variances. This model is used to establish that all of the variances are similar and measure the same symptom areas between males and females.

Table 9 shows the results of the strict invariance test. The additional constraints increased the $\chi^2$ value from 1572.235 to 1969.593 (p<.01) for the admission data, and from 1473.337 to 1651.918 (p<.01) for the discharge data, gaining 23 degrees of freedom (Table 9). The $\chi^2$ differences, $\chi^2=397.358$, p<.01 for admission, and 178.582, p<.01 for discharge with 28 degrees of freedom, indicates that the model difference was statistically significant and hence strong invariance was not supported. However, the goodness of fit indices between the configural invariance model and the strong invariance model and particularly the $\Delta$CFI value for each model showed that all of the multi-group models were equivalent for both male and female samples at admission and discharge in terms of goodness of fit examined through the $\Delta$CFI, and that all of them show an excellent fit to the data (Table 9). The decrease in CFI was 0.008 for admission and 0.003 for discharge and both fell below the 0.009 cut-off, indicating strong factor invariance. Similar to the strong invariance model, the strict invariance model for admission (6a) demonstrated reasonable good fit of indexes and the model for discharge (5d) demonstrated acceptable fit indexes, indicating that it adequately represented the data, and supported strict
invariance. Both admission and discharge data passed the test for strict invariance because CFI was < -01. Thus, the final hypothesis that BASIS-24 items reflect the same underlying construct across males and females was supported.

All invariance tests came to the same conclusion based on the CFI < -01 cutoff rule, though the constraints reduced the CFI slightly. Hence additional constraints did not alter the final decision (Figure 9). However, a large number of contradictory conclusions between $\Delta \chi^2$ and CFI > 0.90 were observed as in all four models. If $\Delta \chi^2$ was employed for decision making, it would have rejected all four hypotheses.
CHAPTER IV: DISCUSSION

This study tested the factorial invariance of BASIS-24 across gender in a national field test sample of 1,398 psychiatric inpatients who completed BASIS-24 at admission and discharge at 11 facilities nation-wide. The scope of the study, along with the reliability and validity data, suggest that the revised BASIS-24 instrument achieves all goals proposed in this study. The discussion section summarizes the results of the study as well as implications for social work practices, social work education, and suggestions for future research.

Summary Results

The study findings show that BASIS-24 is a valid patient self-rated measure, and the data demonstrate good fit for the male and female datasets for the six factor model: a) Depression/Functioning, (b) Interpersonal, (c) Self-Harm, (d) Emotional Lability, (e) Psychosis, and (f) Substance Abuse. Second, the current findings suggest that BASIS-24 can be used to assess patient symptoms and functioning at admission, discharge, and/or change scores for both males and females. Third, the factor invariance demonstrated significantly better results for the discharge sample than expected. Finally, the present study supports the configural, metric, as well as the strong and strict factorial invariance of BASIS-24 across both genders. More specifically, the following study hypotheses were accepted based on the current study:

1. The six factor structure of BASIS-24 is invariant across gender groups. Both male and female groups associate the same subscales of BASIS-24 items (configural invariance).

2. The factor loadings of BASIS-24 are invariant across gender. The strength of the relationship based on the factor loadings between each item and its underlying construct is the same for both groups (metric invariance).
3. The group differences and relationships among constructs (covariances) are the same across groups (strong invariance).

4. BASIS-24 items have an invariant pattern of unique variances or residual or error variances. In other words, BASIS-24 items measure the underlying construct (strict or scalar invariance).

When a measure holds configural invariance across groups, the measure can be construed in the same way across groups so that it is a valid measure for capturing a concept of interest. However, configural invariance has limits in the comparison of factor means, variances, and covariances across groups because these parameters are not necessarily invariant. Weak metric invariance, which suggests that the relationships between factors and their indictors are the same across groups, permits comparison of factor variances and covariances but does not guarantee factor means across groups. Along the same lines, comparisons of factor variances and covariances are meaningful when strong metric invariance holds. The presence of strict metric invariance is the best evidence for comparing factor means, variances, and covariances across gender because a measure holding strict metric invariance can be used without a systemic bias due to possible cultural differences (Van deVijver & Tanzer, 2004). The important finding of the present study is that the BASIS-24 can be used as a reliable and valid symptom measure in assessing psychiatric inpatient populations and can be used to compare quantitative differences in the magnitude of patient symptoms and functioning across gender.

Implications for Research

This study offers significant contributions to social work research in the area of psychiatric patient evaluation and social work research by conducting a validity study for
psychiatric patients who commonly use the BASIS-24 instrument. Though mental health services researchers have identified the underlying structure of patient assessment in psychiatric assessment, the progression of such knowledge of scientific assessment and outcomes in this regard has been modest among social workers.

This study especially answered an important research question: does the six-factor structure of the BASIS-24 hold for male and female samples?

This study has several implications for mental health service research and evaluation. First, the measurement invariance models and its results provide researchers with important psychometric properties about BASIS-24 subscales. It offers important insight into the study questions; it aids in understanding prior studies, and provides directions for future research. For example, the six BASIS-24 subscales can be used in research across gender groups. Second, the current study supports the generalizability of measurement relationships between latent variables and their manifest indicators across populations. More specifically, BASIS-24 can be helpful for comparing group differences in means and covariance matrices across genders.

Third, the current study provides an example of a useful statistical methodology for examining specific questions related to factorial invariance of the BASIS-24 instrument or similar instruments across gender. The methods outlined in this study address several questions often asked about validity, measurement, assessment, and evaluation by behavioral and social work researchers. Thus, this method for testing measurement invariance can be used for future studies that use the BASIS-24 with a variety of populations. For example, structural equation modeling is a useful technique to evaluate these measures because it identifies the most reliable and valid items for each scale. For example, to test the null hypothesis, the current study fit four nested models: first without restriction; second, with constrained factor loadings; third, with
constrained factor loadings, covariances and variances; and fourth, with constrained factor loadings, covariances, variances and error variances. The difference in CFI values of the restricted model was compared with the unconstrained model to the hypotheses.

Patient symptoms and functioning are complex constructs that are difficult to measure. Thus, the self-rated measures are not free from measurement errors. In this context, the use of structural equation modeling to test the factorial invariance of BASIS-24 is a robust method over basic statistical approaches. It enables one to estimate the measurement error present in each variable and makes the psychometric evaluation feasible in assessing psychiatric functioning and symptoms in which a significant amount of measurement error exists. Future research can build on this research by using these findings to develop the measures further.

Social workers in general are behind in using the advanced research techniques and statistical software available to analyze complex clinical and social problems (Jenson, 2008). In recent years, however, there has been an increase in the availability of statistical techniques and software used to analyze complex data. Keeping pace with relevant advances in quantitative methodology and analysis is critical to competing successfully for external funding and advancing knowledge about mental health and social problems. These study methods and the analytical procedures will add to the quantitative methodology. For example, the current study using structural equation modeling has the capability to handle complex models, multiple groups, measurement error, and observed and latent variables, and to simultaneously estimate from sample observations the population values of all variables. Also, it facilitates testing whether the hypothesized model fits or is supported by the empirical data.
*Implications for practice*

The methods of tracking changes in patient’s functioning are essential to the process of mental health treatment (Kinzie & Manson, 1987). Especially with the emergence of third party payers for mental health services, patient assessment, and mental health outcomes have become a necessary part of mental health services. Broad interest in patient assessment and outcomes has been growing among psychologists, nurses, and social workers in the field of mental health services. Much of the interest generated has been due to multiple factors including deinstitutionalization of mental health care, pharmacology, reduced length of stay in hospitalization, The Joint Commission accreditation requirements, and the professional recognition of evidence based care in clinical practice. Measuring progress has become a part of case management where social workers are primarily involved in patient assessment.

Traditionally mental health assessment and practices were mostly focused on personality traits, substance abuse, suicidal thoughts, and violence to self or others, and diagnosis leading to a treatment plan. The Global Assessment of Functioning Scale (GAF) is completed by a clinician. It evaluates psychological symptoms and occupational and social functioning. However most of the mental health providers are now using both a clinician-based patient assessment and a patient self-rated assessment. These are complementary to each another. Patient self-rated instruments score the patient’s perspective on their own problems in the areas of behavior, functioning, and symptoms instead of relying on personality traits or diagnosis. Clinical social workers and nurses do the bulk of mental health work, and the profession is having a hard time catching up with recent changes happening in the area of patient assessment.

Clinical social workers could play a major role in the development and validation of patient-centered, culture-specific assessment tools. There are several examples of patient
assessment tools outlined by (Maruish, 2004) that focus on measurement of treatment planning and outcome. One of the tests described in Maruish's book provides a model of how mental health assessment can adapt to changing trends. BASIS-24 is one of a handful of instruments that measure both functioning and symptoms. The use of BASIS-24 among psychiatric inpatients is growing significantly (S. V. Eisen et al., 2006). The self-report instrument, SF-36, and the Quality of Life Interview (QOLI) similarly examine symptoms and functioning, as well as related factors, in the broader context of health-related quality of life.

Mental health assessment has become an important part of social work clinical practice and in research and policy. In this context, BASIS-24 can be used with inpatient psychiatric patients to measure their functioning and symptoms. Second, this test confirms that the mean score differences among males and females in BASIS-24 subscales are actual differences and are not due to the measurement error under normal circumstances. Thus BASIS-24 can be considered a reliable instrument for measuring mental health outcomes by comparing the pre (admission) and post (discharge) mean scores. This confirms the utility of the BASIS-24 for both clinical and research purposes. These findings represent a major recognition for the use of the BASIS-24 in mental health service and research. This has particular relevance to social work where there is a strong commitment to gender and culture-specific mental health services. These findings suggest that BASIS-24 can be used with confidence in practice settings in which practitioners require an assessment of symptoms across gender with inpatient population. Likewise, social work researchers can use BASIS-24 in studies that target gender and severe mental disorders.
Limitations

The results of the current study may not apply to residential, partial or outpatient populations with different types of mild and moderate conditions. Further study is recommended concerning the utility of self-rated measures such as BASIS-24 with a clinician’s ratings such as GAF score. Furthermore, though admission and discharge data were used to test the factorial invariance of the BASIS-24 subscale across gender, changes in mental health status that may have occurred over the two time points and that may affect patient’s report were not measured. Finally, this study did not test the invariance of the BASIS-24 across age, race/ethnicity or diagnostic groups.

In summary, the two important findings of the present study are: 1) BASIS-24 can be used as a reliable and valid symptom measure in assessing a psychiatric inpatient population and can be used to compare quantitative differences in the magnitude of patient symptoms and functioning across gender; 2) the current study provides an example of a useful statistical methodology for examining specific questions related to factorial invariance of the BASIS-24 instrument across gender. It is hoped that the BASIS-24 instrument may help social workers determine treatment goals and measure treatment progress in their work with inpatient psychiatric populations.
REFERENCES


http://www.naswpress.org/publications/standards/code.html


Vandenberg, R. J. (2002). Toward a further understanding of and improvement in measurement invariance methods and procedures. *Organizational Research Methods, 5*(2), 139-158.


World Health Organization. (2002). *Gender and mental health*
APPENDIX A: BASIS-24® (BEHAVIOR AND SYMPTOM IDENTIFICATION SCALE)

Instructions to Staff: Please fill in the following information completely.

Client ID: __ __ __ __ __ __ __ __

HCO ID: __ __ __ __

Admission / Intake Date: __ __ / __ __ / __ __

Time Point:
1 Admission/Intake
2 Mid-treatment
3 Discharge termination
4 Post-treatment follow-up

Level of Care:
1 Inpatient
2 Outpatient
3 Partial/day hospital
4 Residential

Program Type or Unit: __ __

Instructions to Respondents:
This survey asks about how you are feeling and doing in different areas of life. Please check the box to the left of your answer that best describes yourself during the PAST WEEK. Please answer every question. If you are unsure about how to answer, please give the best answer you can.

EXAMPLE: During the past week, how much difficulty did you have sleeping?
0 No difficulty
1 A little difficulty
2 Moderate difficulty
3 Quite a bit of difficulty
4 Extreme difficulty

During the PAST WEEK, how much difficulty did you have...

1. Managing your day-to-day life?
0 No difficulty
1 A little difficulty
2 Moderate difficulty
3 Quite a bit of difficulty
4 Extreme difficulty

2. Coping with problems in your life?
0 No difficulty
1 A little difficulty
2 Moderate difficulty
3 Quite a bit of difficulty
4 Extreme difficulty

3. Concentrating?
0 No difficulty
1. A little difficulty
2. Moderate difficulty
3. Quite a bit of difficulty
4. Extreme difficulty

During the PAST WEEK, how much of the time did you...

4. Get along with people in your family?
   0. None of the time
   1. A Little of the time
   2. Half of the time
   3. Most of the time
   4. All of the time

5. Get along with people outside your family?
   0. None of the time
   1. A Little of the time
   2. Half of the time
   3. Most of the time
   4. All of the time

6. Get along well in social situations?
   0. None of the time
   1. A Little of the time
   2. Half of the time
   3. Most of the time
   4. All of the time

During the PAST WEEK, how much of the time did you...

7. Feel close to another person?
   0. None of the time
   1. A Little of the time
   2. Half of the time
   3. Most of the time
   4. All of the time

8. Feel like you had someone to turn to if you needed help?
   0. None of the time
   1. A Little of the time
   2. Half of the time
   3. Most of the time
   4. All of the time

9. Feel confident in yourself?
   0. None of the time
   1. A Little of the time
   2. Half of the time
   3. Most of the time
   4. All of the time
During the PAST WEEK, how much of the time did you...

10. Feel sad or depressed?
   0  None of the time
   1  A Little of the time
   2  Half of the time
   3  Most of the time
   4  All of the time

11. Think about ending your life?
   0  None of the time
   1  A Little of the time
   2  Half of the time
   3  Most of the time
   4  All of the time

12. Feel nervous?
   0  None of the time
   1  A Little of the time
   2  Half of the time
   3  Most of the time
   4  All of the time

During the PAST WEEK, how often did you...

13. Have thoughts racing through your head?
   0  Never
   1  Rarely
   2  Sometimes
   3  Often
   4  Always

14. Think you had special powers?
   0  Never
   1  Rarely
   2  Sometimes
   3  Often
   4  Always

15. Hear voices or see things?
   0  Never
   1  Rarely
   2  Sometimes
   3  Often
   4  Always

16. Think people were watching you?
   0  Never
   1  Rarely
   2  Sometimes
   3  Often
   4  Always
17. Think people were against you?
0 Never  
1 Rarely  
2 Sometimes  
3 Often  
4 Always  

*During the PAST WEEK, how often did you…*

18. Have mood swings?
0 Never  
1 Rarely  
2 Sometimes  
3 Often  
4 Always  

19. Feel short-tempered?
0 Never  
1 Rarely  
2 Sometimes  
3 Often  
4 Always  

20. Think about hurting yourself?
0 Never  
1 Rarely  
2 Sometimes  
3 Often  
4 Always  

*During the PAST WEEK, how often…*

21. Did you have an urge to drink alcohol or take street drugs?
0 Never  
1 Rarely  
2 Sometimes  
3 Often  
4 Always  

22. Did anyone talk to you about your drinking or drug use?
0 Never  
1 Rarely  
2 Sometimes  
3 Often  
4 Always  

23. Did you try to hide your drinking or drug use?
0 Never  
1 Rarely  
2 Sometimes  
3 Often  
4 Always
24. Did you have problems from your drinking or drug use?
0  Never
1  Rarely
2  Sometimes
3  Often
4  Always

ABOUT YOU

25. How old are you? _____

26. What is your sex?
1  Male  2  Female

27. Are you…
1  Hispanic or Latino
2  NOT Hispanic or Latino

28. What is your racial background?
1  American Indian or Alaskan native
2  Asian
3  Black or African-American
4  White/Caucasian
5  Native Hawaiian or other Pacific Islander
6  Multiracial or other (specify)

29. How much school have you completed?
1  8th grade or less
2  Some high school
3  High school graduate/GED
4  Some college
5  4-year college graduate or higher

30. Are you now…
1  Married
2  Separated
3  Divorced
4  Widowed
5  Never married

31. Outside of your treatment providers, what is your main source of social support?
1  wife, husband, or partner
2  Other family (parents, children, relatives)
3  Friends/roommates
4  Community/church
5  Other
6  No one

32. Where did you sleep in the past 30 days?
1  Apartment or house
2. Halfway house/group home/board and care home/residential center/supervised housing
3. School or dormitory
4. Hospital or detox center
5. Nursing home/assisted living
6. Shelter/street
7. Jail/prison
8. Other (fill in)_____________________

33. At any time in the past 30 days, did you work at a paying job?
   1. No
   2. Yes, 1 – 10 hours per week
   3. Yes, 11 – 30 hours per week
   4. Yes, more than 30 hours per week

34. At any time in the past 30 days, did you work at a volunteer job?
   1. No
   2. Yes, 1 – 10 hours per week
   3. Yes, 11 – 30 hours per week
   4. Yes, more than 30 hours per week

35. At any time in the past 30 days, were you a student in a high school, job training, or college degree program?
   1. Yes
   2. No

36. Do you now receive disability benefits; for example, SSI, SSDI, or other disability insurance (Check one or more)
   1. No
   2. Yes, I receive disability for medical reasons
   3. Yes, I receive disability for psychiatric reasons
   4. Yes, I receive disability for substance abuse

37. Today’s Date: __ __ / __ __ / __ __

THANK YOU VERY MUCH!

© McLean hospital  10/2001
APPENDIX B: NASW ABSTRACT

The Behavior and Symptom Identification Scale 24 (BASIS-24) is a psychiatric outcome measure used for inpatient and outpatient populations. The 24-item measure comprises six subscales: depression/functioning, interpersonal relationships, self-harm, emotional lability, psychosis, and substance abuse. The psychometric properties of the measure have never been analyzed for its validity and reliability across gender. The purpose of this study was therefore to assess the construct validity of the BASIS-24 six-factor model and find evidence of configural, metric, strong and strict factorial invariance across gender. The sample consisted of 1398 psychiatric inpatients that completed BASIS-24 at admission and discharge at 11 facilities. Multiple-group confirmatory factor analysis was used to test measurement invariance of the BASIS-24 six-factor model across males and females. The present study supported the configural, metric, as well as the strong and strict factorial invariance of the BASIS-24 across males and females. Implications of social work practice and research are discussed.