Reliability and Validity of a Substance Craving Scale: Testing Invariance across Substances

BY

JESSICA MAZZA
B.A., New College of Florida, 2004
M.S.P.H., University of South Florida, 2008

THESIS

Submitted as partial fulfillment of the requirements for the degree of Doctor of Philosophy in Public Health Sciences in the Graduate College of the University of Illinois at Chicago, 2013

Chicago, Illinois

Defense Committee:

Kendon Conrad, Chair and Advisor
Sage Kim
Ross Mullner
Barth Riley, Nursing
Michael Dennis, Chestnut Health Systems, Inc.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>A. Definition and Construct of Craving</td>
<td>3</td>
</tr>
<tr>
<td>B. Craving as a Variable and Treatment Goal</td>
<td>4</td>
</tr>
<tr>
<td>C. Measurement of Craving</td>
<td>5</td>
</tr>
<tr>
<td>D. Rasch Measurement</td>
<td>8</td>
</tr>
<tr>
<td>E. Research Objectives</td>
<td>10</td>
</tr>
<tr>
<td>F. Significance</td>
<td>11</td>
</tr>
<tr>
<td>II. LITERATURE REVIEW</td>
<td>12</td>
</tr>
<tr>
<td>A. Theoretical Models of Craving</td>
<td>12</td>
</tr>
<tr>
<td>1. Conditioning models</td>
<td>12</td>
</tr>
<tr>
<td>2. Cognitive models</td>
<td>13</td>
</tr>
<tr>
<td>3. Psychobiological models</td>
<td>14</td>
</tr>
<tr>
<td>B. Craving as an Outcome Variable of Treatment</td>
<td>15</td>
</tr>
<tr>
<td>C. Craving and Substance Use History</td>
<td>15</td>
</tr>
<tr>
<td>1. Craving and withdrawal</td>
<td>16</td>
</tr>
<tr>
<td>2. Craving and substance dependence</td>
<td>17</td>
</tr>
<tr>
<td>D. Craving by Drug</td>
<td>18</td>
</tr>
<tr>
<td>1. Alcohol</td>
<td>19</td>
</tr>
<tr>
<td>a. Alcohol craving measurement</td>
<td>19</td>
</tr>
<tr>
<td>2. Cannabis</td>
<td>20</td>
</tr>
<tr>
<td>a. Cannabis craving measurement</td>
<td>21</td>
</tr>
<tr>
<td>3. Cocaine</td>
<td>21</td>
</tr>
<tr>
<td>a. Cocaine craving measurement</td>
<td>22</td>
</tr>
<tr>
<td>4. Amphetamines</td>
<td>23</td>
</tr>
<tr>
<td>a. Amphetamines craving measurement</td>
<td>24</td>
</tr>
<tr>
<td>5. Opioids</td>
<td>24</td>
</tr>
<tr>
<td>a. Opioids craving measurement</td>
<td>24</td>
</tr>
<tr>
<td>E. Predictive Validity of Craving</td>
<td>25</td>
</tr>
<tr>
<td>F. Properties of Craving Measures</td>
<td>26</td>
</tr>
<tr>
<td>1. Single-item measures</td>
<td>26</td>
</tr>
<tr>
<td>2. Multiple-item measures</td>
<td>27</td>
</tr>
<tr>
<td>a. Length of measures</td>
<td>28</td>
</tr>
<tr>
<td>b. Time frame</td>
<td>28</td>
</tr>
<tr>
<td>G. Review of Existing Self-Report Measures of Drug Craving</td>
<td>29</td>
</tr>
<tr>
<td>H. Preliminary Work: Development of the Substance Craving Scale</td>
<td>32</td>
</tr>
<tr>
<td>1. Construct map</td>
<td>32</td>
</tr>
<tr>
<td>2. Rationale for Rasch Measurement</td>
<td>35</td>
</tr>
<tr>
<td>I. Demonstrating Reliability and Validity of the Substance Craving Scale</td>
<td>36</td>
</tr>
<tr>
<td>1. Reliability</td>
<td>37</td>
</tr>
<tr>
<td>2. Dimensionality</td>
<td>37</td>
</tr>
<tr>
<td>3. Validity of the items</td>
<td>38</td>
</tr>
</tbody>
</table>
TABLE OF CONTENTS (continued)

CHAPTER                  PAGE

a. Item fit................................................................. 38
b. Differential Item Functioning .................................. 38
4. Construct and criterion validity of the measure .......... 39
   a. Construct validity................................................. 40
   b. Criterion validity............................................... 40
   c. Multivariate models ............................................. 44
J. Research Objectives of the Current Study ................ 47

III. METHODS ............................................................ 48
A. Rasch Measurement Model........................................ 48
B. Data Collection and Sample ..................................... 50
   1. Data collection .................................................. 50
   2. Sample.................................................................. 50
C. Analysis Plan .......................................................... 50
   1. Objective 1: Reliability ......................................... 51
      a. Cronbach’s alpha .............................................. 51
      b. Rasch reliability .............................................. 51
   2. Objective 2: Unidimensionality ................................. 52
   3. Objective 3: Item quality .................................... 52
   4. Objective 4: Invariance of hierarchy (using DIF) ........ 53
   5. Objective 5: Construct and Criterion Validity .......... 53
      a. Construct validity ............................................. 53
      b. Criterion validity using hypothesized correlations .... 54
      c. Multivariate models ......................................... 54
         1) Linear regression model ................................. 55
         2) Stepwise regression model ............................... 55

IV. RESULTS .................................................................. 57
A. Demographics .......................................................... 57
B. Analysis Sample ....................................................... 58
C. Preliminary Analysis ................................................. 59
D. Rasch Analysis: Objectives ....................................... 60
   1. Objective 1: Reliability .......................................... 60
      a. Reliability by drug .......................................... 61
      b. Item reliability ................................................. 62
   2. Objective 2: Unidimensionality ................................ 62
   3. Objective 3: Item quality ..................................... 63
   4. Objective 4: Invariance of hierarchy (using DIF) ....... 65
   5. Objective 5: Construct and criterion validity .......... 68
      a. Construct validity ............................................. 68
TABLE OF CONTENTS (continued)

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>V. CONCLUSIONS</td>
<td>75</td>
</tr>
<tr>
<td>A. Purpose and Summary of the Study</td>
<td>75</td>
</tr>
<tr>
<td>B. Psychometric Analysis – Reliability and Validity</td>
<td>76</td>
</tr>
<tr>
<td>1. Reliability</td>
<td>76</td>
</tr>
<tr>
<td>2. Dimensionality</td>
<td>77</td>
</tr>
<tr>
<td>3. Validity of the items</td>
<td>77</td>
</tr>
<tr>
<td>4. Invariance of hierarchy by drug</td>
<td>79</td>
</tr>
<tr>
<td>5. Construct and criterion validity of the measure</td>
<td>80</td>
</tr>
<tr>
<td>a. Construct validity</td>
<td>81</td>
</tr>
<tr>
<td>b. Criterion validity</td>
<td>81</td>
</tr>
<tr>
<td>C. Limitations of the Study</td>
<td>83</td>
</tr>
<tr>
<td>1. Homogenity of the Sample</td>
<td>83</td>
</tr>
<tr>
<td>2. Items came from the literature</td>
<td>83</td>
</tr>
<tr>
<td>D. Recommendations for Future Research</td>
<td>84</td>
</tr>
<tr>
<td>1. Improving the scale and its items</td>
<td>84</td>
</tr>
<tr>
<td>2. Establishing long form, short form and single item scales</td>
<td>85</td>
</tr>
<tr>
<td>a. Long form</td>
<td>85</td>
</tr>
<tr>
<td>b. Short form</td>
<td>86</td>
</tr>
<tr>
<td>c. Single item</td>
<td>86</td>
</tr>
<tr>
<td>3. Study potential reactivity to the scale</td>
<td>87</td>
</tr>
<tr>
<td>E. Significance and Conclusion</td>
<td>87</td>
</tr>
</tbody>
</table>

REFERENCES | 99 |

CURRICULUM VITAE | 100 |
### LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. MEASURES USED IN THE DEVELOPMENT OF THE SUBSTANCE CRAVING SCALE</td>
<td>31</td>
</tr>
<tr>
<td>II. VARIABLES USED TO TEST CRITERION VALIDITY</td>
<td>42</td>
</tr>
<tr>
<td>III. DEMOGRAPHIC CHARACTERISTICS OF SAMPLE</td>
<td>58</td>
</tr>
<tr>
<td>IV. ITEM MISFIT, 12-ITEM SCALE</td>
<td>59</td>
</tr>
<tr>
<td>V. CRONBACH’S ALPHA AND RASCH PERSON RELIABILITY, BY DRUG</td>
<td>61</td>
</tr>
<tr>
<td>VI. DIMENSIONALITY ANALYSIS</td>
<td>63</td>
</tr>
<tr>
<td>VII. SUBSTANCE CRAVING SCALE ITEMS, LABELS, MEASURES, FIT STATISTICS, AND CORRELATIONS</td>
<td>64</td>
</tr>
<tr>
<td>VIII. NUMBER OF ITEM ENDORSEMENTS BY DRUG</td>
<td>65</td>
</tr>
<tr>
<td>IX. AVERAGE ITEM MEASURE BY DRUG, WITH DIF CONTRASTS</td>
<td>67</td>
</tr>
<tr>
<td>X. CORRELATIONS OF CRAVING WITH OTHER MEASURES</td>
<td>71</td>
</tr>
<tr>
<td>XI. HIERARCHICAL LINEAR REGRESSION OF CRAVING PREDICTORS</td>
<td>72</td>
</tr>
<tr>
<td>XII. STEPWISE REGRESSION OF CRAVING PREDICTORS</td>
<td>73</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Substance Craving Scale construct map</td>
<td>34</td>
</tr>
<tr>
<td>2. Theoretical model of craving with hypothesized correlations</td>
<td>46</td>
</tr>
<tr>
<td>3. Person DIF measured by item</td>
<td>68</td>
</tr>
<tr>
<td>4. Wright map of the Substance Craving Scale</td>
<td>70</td>
</tr>
</tbody>
</table>
SUMMARY

Introduction

Craving for alcohol or drugs has been studied for many years by scholars and clinicians in an effort to improve the efficiency of substance abuse treatment across individuals and settings. The reduction of craving has been conceptualized as a target of intervention in addiction treatment programs and post treatment relapse prevention programs. The purpose of this research was to develop a substance craving scale that is reliable and valid for use across substance use disorders. The scale was examined via the Rasch model, an item response theory model that provides the analysis of data through use of a single, difficulty parameter. To date, there has been no Rasch analysis conducted of a substance craving scale that assesses cravings for all substances.

Methods

The current Substance Craving Scale was developed based on several of the prior validated, drug-specific scales that are used to measure substance craving. An additional craving item, used on the National Longitudinal Alcohol Epidemiologic Survey and a key basis of the DSM-V decision to add craving as a criterion, was also included in the final scale. Analysis of the data employed both Rasch and traditional test theory techniques. The measure was analyzed with a Rasch dichotomous model with Winsteps statistical software, and other techniques used SPSS statistical software. Analyses were performed to provide summary statistics of the data, to determine item fit, dimensionality of the measure, reliability, validity, and differential item functioning (DIF) for substance of choice.
Results

The results showed that all but one item on the measure were appropriate based on a preliminary item analysis, and the final psychometric analyses were conducted on the revised 11-item scale. The craving scale was found to be unidimensional, reliable for research purposes, and showed evidence of validity of construct. Most importantly, there was no significant differential item functioning by substance of choice. Discussion. Several items were discussed regarding their theoretical importance, and the validity and reliability of the scale were critiqued. Limitations of the study included homogeneity of the sample, hence reliability and validity of the Substance Craving Scale should be replicated in different samples, especially for those individuals entering substance treatment.
I. INTRODUCTION

Craving for alcohol or drugs has been studied for many years by scholars and clinicians in an effort to improve the efficiency of substance abuse treatment across individuals and settings. In fact, the American Psychiatric Association has added craving as a criterion for the definition of substance use disorder (SUD) as part of the new Diagnostic and Statistical Manual Version V (DSM-V) released this year (APA, 2013). The reduction of craving has often been conceptualized as a goal of intervention in addiction treatment programs and post treatment relapse prevention programs. Unfortunately, progress towards this goal is illusive since the addiction field has lacked clear definition of the construct of craving. Consequently, there is a lack of well operationalized and validated measures for treatment or research.

This research addressed this problem by developing a substance craving scale that is reliable and valid for use across the major substance use disorders. It examined the Substance Craving Scale using the Rasch model, including dimensionality, item quality, reliability, validity (including construct and criterion) and invariance of hierarchy, where there are no differences in item response by primary substance.

Historically, craving was first recognized in 1955 by Jellinek and his colleagues as a critical component of alcohol dependence. It is widely accepted that craving is part of the addiction process of other drugs, as well (Heishman, Singleton, & Liguori, 2001). The reduction of craving has been conceptualized as a target of intervention in substance abuse treatment programs and is often cited as a desirable outcome in substance abuse treatment. It is also sometimes considered as a contributing factor to substance use relapse. However, the lack of clarity surrounding the definition of craving has produced significant confusion and controversy in the addiction literature. As an example, one dilemma in the study of craving concerns which
components to include in the research, since there are various domains affected by craving: emotional, cognitive, behavioral, and physiological factors influence an individual’s craving in nuanced ways (Rosenberg, 2009).

Craving can be thought of both as a process and an outcome (Drummond, Litten, Lowman, & Hunt, 2000). As a process, it is how an individual reacts to certain situations or stimuli, which could lead to relapse. As an outcome, it can be used to track treatment progress, or to develop clinical profiles of substance use disorders. The clinical applications of assessing an individual’s self-report of craving for substances are myriad. These include: supporting and informing treatment planning (O’Brien, 2005), predicting readiness for discontinuing or stepping-down treatment (Anton & Drobes, 1998), assessing likelihood of relapse after treatment discharge (Doherty, Kinnunen, Militello, & Garvey, 1995), and evaluating those treatments (pharmacological and otherwise) that intend to lower substance craving (Conklin & Tiffany, 2002). Most recently, the availability and development of anti-craving drugs has increased the need for a reliable and valid self-report of craving. Although researchers disagree both about the validity of craving as a construct, and its association with relapse in the addictive disorders (Lowman, Hunt, Litten, & Drummond, 2000), craving has been a very commonly used variable in measuring outcomes of substance treatment programs.

The issues surrounding craving as a definable and measurable construct expose a significant gap in both the understanding of addiction and the demonstrative utility of craving-specific treatment interventions. As one of the leaders in addiction research recently claimed, “Craving…will not disappear from the researcher’s or clinician’s ledger. People addicted to drugs will demand that scientists acknowledge its intrusive, noxious qualities” (Tiffany & Wray, 2009, p. 1618).
The information in this problem statement will highlight the need for an appropriate measure of craving. First, the problems regarding the definition of craving and its conceptualization as a construct will be presented. Next, craving as a variable, i.e., an independent variable that predicts relapse or reduction of craving as a dependent outcome variable, will be discussed. The issues surrounding the existing measurement scales used to assess alcohol and illicit-drug craving (nicotine craving will not be included) will also be examined.

A valid and reliable measure of craving is needed to help clarify many of the conceptual gaps in the literature; this discussion will describe how the use of Rasch analysis (analysis of data through use of a specific single-parameter model) can aid in the examination of craving as a unidimensional and generalizable measure, across substances, individuals, and treatment settings. Finally, an explicit statement of the problem will be presented as a framework for the literature synthesis required to develop a more thorough understanding of the knowledge base of substance craving prior to the eventual conceptual development, implementation, and analysis of a substance craving scale.

A. Definition and Construct of Craving

Despite the fact that craving is considered an important topic of interest by scholars, clinicians, and addicted persons, definitions of craving vary across studies (Connolly, Coffey, Baschnagel, Drobes, & Saladin, 2009; Heishman, Singleton, & Liguori, 2001; Merikle, 1999; Rosenberg, 2009; Sayette et al., 2000; Tiffany & Wray, 2009). One review of the literature found that many researchers have posited that craving is an individual state that motivates
substance use (Eliason & Amodia, 2007), and it can be most simply defined as the ‘desire to use a drug’ (Sayette et al., 2000).

However, while some researchers define craving explicitly and sufficiently as the desire to use a substance (Sayette et al., 2000), others believe that craving refers to a broader definition or set of factors, including such varying factors such as anticipation of positive effects when using, desire to alleviate withdrawal symptoms through use, and the explicit purposefulness of a user to acquire and use drugs or alcohol (James, Davies, & Willner, 2004).

Given the many definitional challenges of craving, is it not surprising that there is no universally accepted conceptual construct of craving. Although there has been 30 years of dedicated research focused on substance craving, there has been limited study on the theory of the phenomenon or construct of craving in alcohol- or drug-addicted persons (Eliason & Amodia, 2007), and debate about the validity of an accepted construct continues (Lowman, Hunt, Litten & Drummond, 2000; Rosenberg, 2009). Again, this may be related to the definitional challenges surrounding craving, since many researchers define and operationalize substance craving in different ways (Heishman, Singleton, & Liguori, 2001; Rosenberg, 2009).

B. Craving as a Variable and Treatment Goal

As previously mentioned, craving is often considered to be an independent variable that is used to predict relapse of alcohol/drug use (Tiffany & Wray, 2009). In 2000, Drummond, Litten, Lowman, and Hunt postulated four possible associations between craving and relapse: (1) craving is not associated with relapse, (2) craving is associated with relapse, but current measures of craving do not suffice, (3) craving can only predict relapse under certain conditions, and (4) craving is not associated with relapse, but certain correlates of craving are associated.
Now, nearly a decade later, the association between craving and relapse has still not been determined (Tiffany & Wray, 2009). As suggested in the second postulation by Drummond and his colleagues, appropriate craving measures are needed to test hypotheses if researchers want to establish the existence of a relationship between craving and relapse.

Craving is also directly targeted during substance abuse treatment, since craving-reduction is often a principal component of program modalities and therapeutic approaches, such as in cognitive-behavioral therapy (O’Brien, 2005; Heishman, Singleton, & Liguori, 2001; Sussner et al., 2006). Since craving-reduction is a desired process goal of substance abuse treatment, it would be useful to demonstrate success of individual craving reduction, through appropriate measurement while a person is entered into substance abuse treatment.

Craving reduction has also been considered as an outcome measure of treatment (Singleton & Gorelick, 1998). It is relevant to the clinical interventions that need to understand, monitor, and use craving as part of the treatment approach to reduce or stop use (Anton, 1999). Again, this outcome measure is problematic given the uncertainty surrounding craving as a measurable construct. To determine association between craving and any other variable (as predictor or an outcome), it is necessary to develop and improve upon measures that can validly test these associations.

C. Measurement of Craving

Existing measures have not shown a clear association of craving and subsequent drug use (Tiffany & Wray, 2009). It is important to note that a psychometrically valid craving measure is necessary to adequately prove or disprove associations and test hypotheses among variables of interest. As noted by Cronbach in 1970, “Taking a better measure improves the sensitivity of an
experiment in the same way that increasing the number of subjects does” (p.152). There is still no agreement of ‘gold standard’ among researchers and clinicians on how to accurately measure craving (Anton, 1999; Harold, 2009; Rosenberg, 2009; Sayette et al., 2000). Measuring substance craving through a self-report measure is decidedly the best approach at understanding what is, essentially, a subjective psychological experience (Rosenberg, 2009).

Drug and alcohol craving scales in the current literature are limited in both their scope and the sophistication of their conceptual development. Many of the existing craving scales lack systematic conceptual work in their creation through the use of appropriate conceptually-driven scale development techniques (e.g., concept mapping, cognitive interviews, focus groups). In the construction of measures, it is important to progress in the formulation of initial concepts, item development, and potential outcome models with an appropriate degree of attention to how the processes and data relate back to the construct of interest (Wilson, 2005).

When researchers first began examining craving in substance abusing individuals, they started first with single-item ‘scales’ (Connor, Feeney, & Young, 2005; Heishman, Singleton & Moolchan, 2003). Although researchers are still using single-item scales, researchers have developed scales with multiple items to more adequately capture the construct of craving (Sussner et al., 2006). Most of the self-report measures of substance craving, including the ones discussed in this section, are multi-item scales that range from 7-item to 43-item scales. At face value, some of the longer scales have items that are seemingly repetitive or redundant, and some of the shorter scales do not adequately capture enough information to be useful to the understanding of craving as a construct.
Because craving is a subjective experience, the existing craving scales rightly include only self-report items, with questions such as: ‘I want to use cocaine so bad right now, I can almost taste it’ (Sussner et al., 2006), and ‘My desire for speed right now is overwhelming’ (James, Davies, & Willner, 2004). Although many of the scales have overlapping items, it is important to note why these redundancies exist. In several of the existing craving scales, the authors have used items from other scales and adapted these items for the specific substance of interest. For example, in James, Davies & Willner (2004), the authors adapted items from the Desires for Alcohol Questionnaire to create items for their Desires for Speed Questionnaire (Love, James, & Willner, 1998). This is just one example; this seems to be the case for several of the self-report scales.

Further, almost all of the existing drug/alcohol scales are limited to measuring craving for a single substance, i.e., alcohol only, cocaine only (Heishman, Singleton, & Liguori, 2001; Heishman, Singleton, & Moolchan, 2003; James, Davies, & Willner, 2004; Love & Willner, 1998; Mackillop, 2006; Sussner et al., 2006; Tiffany, Singleton, Haertzen, & Henninfield, 1993; Singleton, Tiffany, & Henninfield, 1994). One highly important issue to highlight here is that because these measures are substance-specific, clinicians and researchers who want to use craving as an outcome for patients in substance abuse treatment (with different “primary substances”) cannot aggregate measures across clinical subgroups, for example, cocaine addicts and heroin addicts, in the same sample.

It would be useful for professionals to have a single scale of craving that could be used across substances in order to present findings about craving as a treatment outcome. The development of a universal scale for substance craving would help to build the evidence base about the associations between craving and relapse in treatment across substances. This is
important since researchers posit that craving may help to predict readiness for discharge from substance use treatment, and may also help to predict likelihood of relapse past treatment.

Despite limitations of the current measures, it is still useful to examine the existing measures for their psychometric properties in order to gain a broader understanding of the construct of craving. For example, most existing measures have been validated and critiqued through the use of exploratory and confirmatory factor analyses (Sayette et al., 2000), which can be synthesized to identify important conceptual themes and highlight useful or appropriate scale items. However, the use of a Rasch measurement model for the construct of substance craving can provide additional information about the severity, dimensionality, and validity of measures including across substances using Differential Item Functioning (DIF) as discussed in the next section.

D. Rasch Measurement

Rasch measurement models require that there is an underlying latent construct that is revealed or operationalized through individual participant responses. In this case, it is assumed that craving is the latent construct that is being measured when asking participants various questions about substance urges or desires.

As noted by topical experts in the domain of substance abuse craving, it is important that any given scale be able to discriminate between different levels of craving across different settings, individuals, or drugs (Tiffany, Carter, & Singleton, 2000). One limitation of factor analyses is that there is no hierarchy or severity of items, e.g., endorsing the item “I would do almost anything for some cocaine now,” would probably indicate a more severe level of craving than “Using cocaine now would be pleasant.”
Rasch modeling allows for this differentiation of levels of craving, across treatment settings, drugs and individuals. Moreover, a hierarchy of craving symptomatology will allow researchers to understand the nature of craving, and clinicians would ultimately be able to individualize treatment to appropriate ‘severity’ of cravings. For example, a person with a high measure of craving would receive a more intense intervention than a person with a low or moderate measure of craving.

Most importantly, the item hierarchy provided by the Rasch analysis also provides an indication of construct validity (Smith, 2001). A Rasch analysis of the developed substance craving scale could provide validation of craving as a measurable construct. By using Rasch analysis in addition to other forms of scale analysis (like factor analysis), it may be possible to determine how well substance craving is conceptualized.

Differential Item Functioning is a test of the invariance of the item hierarchy across subgroups, such as age and gender groups (Masters, 1988; Smith, 1992). Presence of significant DIF on item is an indicator that the item may be biased against a group, or simply may function differently for the group. As will be discussed later in this study, this component of the Rasch analysis uniquely allows for the examination of differential item response by subgroups controlling for individuals being at the same level on the trait. In the case of craving, DIF will be useful to determine the differences (if any) of craving across individuals regardless of their primary substance. If the item hierarchies differ on several items per drug, the meaning of the measure could be significantly different by drug. The results may provide information on the differences in the craving experience across users of different drugs, and provide useful information about craving on the whole. Most importantly, if there is little to no DIF on a
A craving measure designed for multiple substances, the tool could be reliably used across large, diverse treatment samples.

E. **Research Objectives**

The addiction field has lacked clear definition of the construct of craving. Consequently, there is a lack of well operationalized and validated measures for treatment or research. The purpose of this research is to develop a substance craving scale that is reliable and valid for use across the major substances of abuse. The proposed research will examine the Substance Craving Scale using the Rasch model, in addition to other traditional validation techniques.

The specific objectives are to test the Substance Craving Scale for:

1. **Reliability** (where Cronbach’s alpha and Rasch person reliability are above .80)
2. **Unidimensionality** (where variance explained by the measurement dimension is greater than 40%, and unexplained variance in the first contrast falls under 15%)
3. **Item Quality** (where mean squares on both infit and outfit are between .75 and 1.33)
4. **Invariance of hierarchy** (no differential item functioning by substance)
5. **Construct and Criterion Validity**, including construct validity against a theoretical hierarchy and criterion validity against hypothesized relationships with Substance Problem Scale, withdrawal problems, and frequency of use (high); Recency of use, Recovery Environment Risk Index (moderate) and Perceived Social Support (low).

Further confirmation of criterion validity will be examined by their placement in a multivariate model to predict craving.
F.  **Significance**

In summary, there is still much debate about the definition, construct, and measurement of substance craving. The limitations of the development and utility of the current substance-specific craving measures highlight the need for a conceptually-sound development of a scale that can be used across individuals, drugs and treatment settings. The reduction of craving has been conceptualized as a target of intervention in addiction treatment programs and post treatment relapse prevention programs. At the same time, the addiction field has lack of clarity surrounding the construct of craving and few well operationalized measures for treatment or research. The proposed measure should allow researchers to generate and validate testable hypotheses; such a measure will allow researchers to examine craving as a variable when evaluating its association with other variables (e.g., relapse).
II. LITERATURE REVIEW

A. Theoretical Models of Craving

There is a multitude of theoretical models, with varying etiologies, that attempt to explain substance abuse craving. Though not discussed here, conceptual information about craving is also derived (to a lesser extent) from philosophy, anthropology, and religious/spiritual practices (Eliason & Amodia, 2007). It is precisely this diversity of theoretical approaches that makes craving such a difficult construct to identify, since there has been little evidence that researchers and theorists incorporate perspectives from outside of their own disciplines (Eliason & Amodia, 2007). Despite differences in craving models, they all assume that it is a complex phenomenon that is motivated and influenced by multiple factors (Anton, 1999; Sussner et al., 2006).

1. Conditioning models

First established by Wickler in a seminal paper on morphine addiction (1948), the conditioning models of substance craving are based on the theory of classical conditioning by Ivan Pavlov (Skinner & Aubin, 2010). In his classic experiment, Pavlov used dogs to demonstrate this learning process. He paired a stimulus (food) that results in a natural response (salivating) with another stimulus (a bell ringing). Over time, the dogs would then associate the food and the bell ringing, such that in the absence of actual food, just the bell ringing would elicit the salivation.

Conditioning models are based on this basic theory. These theories of craving focus on the ‘cues’ that trigger substance cravings (Drummond, 2001; Lowman, Hunt, Litten & Drummond, 2000). The notion here is that craving is an automatic response to a stimulus (Skinner & Aubin, 2010). The basic concept is that in a substance-dependent person, craving can become
the conditioned response (salivation) to a set of paired stimuli (food & bell) over time. As an example, over the many instances of use, a crack-cocaine addict will pair certain stimuli (e.g., the sulphur dioxide smell of a lighted match) with the response of alleviating withdrawal symptoms (the reward of smoking crack cocaine for an addicted person). So the psychological state that motivates this type of behavior, like craving, becomes conditioned with that stimulus—such that when the individual smells the match, he craves to use crack cocaine. Similarly, other conditioning models state that the cues are related to craving because they are paired not with withdrawal symptoms but with the seeking of the pleasurable effects of use (again, this craving behavior is associated with the smell of the match). The conditioning theory of craving will be discussed in more depth later in this synthesis.

2. **Cognitive models**

These types of theories highlight the craving experience as a part of cognitive processes (Eliason & Amodia, 2007). They expand on the conditioning models to include the more sophisticated cognitive functions that work to produce craving (Tiffany, 1999). The basic premise of the craving from a cognitive perspective is that craving stems from social learning, decision-making, memory, and information processes (Skinner & Aubin, 2010). This etiology focuses on the psychological states that occur in persons with substance abuse addiction, including compulsions, problem-solving, negative emotional states, cognitive dysfunctions, and pleasure-seeking behavior (Eliason & Amodia, 2007).

To highlight, in the outcome expectancy cognitive model (Marlatt, 1985), substance cues create the expectations about the positive effects of use, which generates the craving. Although this does sound similar to the conditioned models, it is important to note that
cognitive models focus specifically on higher-level processes being fundamental to generating a craving. So, expectancies about the relief from withdrawal (like one of the craving questionnaire items, “I would be less anxious if I could smoke marijuana right now”) are what generate the actual craving.

3. **Psychobiological models**

Psychobiological models of craving purport that cravings stem from common neurological pathways in the brain. The psychobiological model attempts to integrate psychological processes of substance craving into the theoretical framework by offering the caveat that though cravings share neurological processes in the brain, different types of craving, e.g., a desire to reduce tension vs. anticipation of withdrawal relief, have varying psychological motivators (Eliason & Amodia, 2007).

Essentially, psychobiological theories hypothesize that cravings are explained by biological factors, with an emphasis on psychological motivational factors. As an example, the neuroanatomical model of alcohol craving (Anton, 1999) states that the craving results from the interface of brain structures and memories. That is, the structures of the brain mediate a person’s response to memory-related motivational cues such that these cues are ‘filtered’ differently by the each person’s individual brain neuroanatomy.

Findings from neuroimaging studies have indeed demonstrated that neurobiological processes underlie individual, subjective craving for substances (Paulus, Tapert, & Schulteis, 2009). However, to date, there is no concrete model of what regions of the brain are involved in craving (Skinner & Aubin, 2010). Further, although psychobiological models may demonstrate the neurological processes of where craving might occur neurologically, these
models do not offer much hope by way of therapeutic intervention. In other words, no one has yet proposed to be able to change an addict’s neural circuitry.

B. **Craving as an Outcome Variable of Treatment**

Beyond the theoretical significance of understanding substance craving, the development of a measure of substance craving holds practical importance. First, craving is often seen as an important therapeutic target (Love, James, & Willner, 1998), and second, it also has a role as a predictor of substance relapse (Tiffany & Carter, 1998), even if this relationship is not completely understood. From a clinical perspective, knowledge about a client’s level of craving allows for discussions about the therapy, since it offers an understanding of how craving may lead to use. Additionally, the severity and frequency of craving allow clinicians to tailor their treatment approaches, especially for those clients that have difficulty resisting craving episodes. Therefore, the use of craving as a patient reported outcome (PRO) may help both clinicians and researchers monitor treatment success, and also predict relapse. This may help in designing, targeting and timing clinical care in order to maximize effectiveness of treatment.

C. **Craving and Substance Use History**

Researchers have tested the correlation of craving measures with measures of recent past consumption, dependence severity, and number of drug related problems. Indeed, these researchers have found significant associations between past consumption and craving for alcohol (Anton, 2000; Bohn, Barton, & Barron, 1996; Collins, Koutsky, Morsheimer, & MacLean, 2001; Drummond & Phillips, 2002; Klein, Stasiewicz, Koutsky, Bradizza, & Coffey, 2007; Leonard, Harwood, & Blane, 1988; Love, et al., 1998; McEvoy, Stritzke, French, Lang, &
Ketterman, 2004; Mol et al., 2003), cocaine (Fox et al., 2005; Tiffany, Singleton, Haertzen, & Henningfield, 1993), and benzodiazepines (Mol, et al., 2003). These results are thought to be indicative of validity since craving is assumed to result in higher consumption, but it is also likely that higher consumption results in higher craving, or some confounding variable accounts for both higher consumption and higher craving (Ferguson & Shiffman, 2009).

1. **Craving and withdrawal**

   Many theories around craving assume that the desire to use comes from either the withdrawal from a drug, or the positive reinforcement of use of a particular drug of choice (Tiffany, 1997). Although craving is not uniquely associated with withdrawal, measures of craving appear to be sensitive to withdrawal, and measures of craving rise in response to abstinence from a drug (Baker, Japuntich, Hogle, McCarthy, & Curtin, 2006; Baker, Piper, McCarthy, Majeskie, & Fiore, 2004). Further, other research has suggested that craving is at its most intense in acute withdrawal (Carol, Smelson, Losonczy, & Ziedonis, 2001). Indeed, one test of validity for craving measures is the assumption that craving scores should go up as the blood level of the substance goes down (Rosenberg, 2009).

   Presence of withdrawal symptoms has been used to support the construct validity of several substance-specific measures (Davies, Willner, & Morgan, 2000; Greenwald, 2002; Schuster, Greenwald, Johanson, & Heishman, 1995; Tiffany & Drobes, 1991). It remains questionable if craving represents a physiological or a behavioral state (Cherpitel et al., 2010; O'Brien, 2005), but it is likely both. There are mixed results regarding the association of withdrawal symptoms and the craving experience. For example, Mol et al. (2003) reported that there was no significant correlation between self-report craving for benzodiazepines and
withdrawal symptoms, or dependency severity. Similarly, Ollo et al. (1995) found that craving for cocaine was not associated with withdrawal symptoms or DSM symptoms of dependence.

What is interesting about withdrawal is that, pharmacologically, it is brief and cannot account for late-stage relapse (people who relapse after a long period of abstinence) (Baker, et al., 2006). So craving can indeed also occur in the absence of pharmacological withdrawal. There is some evidence to suggest that craving exerts its own motivational influences for substance use (Baker, et al., 2006), above and beyond withdrawal.

2. **Craving and substance dependence**

Craving is often described as a component of drug dependence, and is seen as a contributor to drug use and relapse either during abstinence from substance use or after treatment (American Psychatric Association, 2000; Heishman, et al., 2001; Keyes, Krueger, Grant, & Hasin, 2011). Some investigators have reported a positive significant association between self-reported craving with severity of dependence or alcohol-related problems (Anton, Moak, & Latham, 1996; Collins, et al., 2001; Drummond & Phillips, 2002; Klein, et al., 2007; Leonard, et al., 1988). Further, the same association was found by James, Davies, and Willner (2004) when they evaluated their measure of craving for amphetamines. Furthermore, they found that those people who injected the drug had higher craving measures than those who used amphetamines orally or nasally.

Although the current DSM-IV-TR does not list craving as a symptom of dependence (American Psychatric Association, 2000), it is an optional symptom in the ICD-10 (World Health Organization, 2005). Craving, which is thought to promote and maintain dependence on substances, was just added as a substance use disorder criteria in the DSM-V, and
is supported by evidence that there is a neurological or genetic foundation for the phenomenon (Cherpitel, et al., 2010). Currently in question is whether craving should be included as a criterion for abuse and dependence. Craving as a new criterion should only be added to the DSM if it improves the reliability and validity of diagnoses (Cherpitel, et al., 2010; Keyes, et al., 2011).

Related, the experience of craving may differ based on where someone fits on the continuum of recreational use to substance dependence (Heishman, et al., 2001; Rosenberg, 2009). Indeed, the presence of craving may reveal a phenotype of substance use or dependence that is more severe than for those individuals who do not experience craving (Keyes, et al., 2011). Some researchers also argue that craving can even be used to discriminate substance abuse from dependence, better than the current DSM-IV criteria (de Bruijn, Korzec, Koerselman, & van Den Brink, 2004).

D. **Craving by Drug**

Although craving is currently assessed with specific measures for individual substances, it should be considered that polysubstance abuse or dependence can make the subjective experience of craving even more complicated. Many substance users use more than one substance, and have variable amounts of craving for varying types of drugs. For example, some drug users who use multiple substances may only experience craving for a single drug, or a single class of drugs, e.g., stimulants only. Further, individuals report that a craving for a drug of choice might be satisfied by another drug (Junghanns, Veltrup, & Wetterling, 2000). This highlights the complexities of craving in polysubstance users, and demonstrates the need for a scale that assesses craving for substances, broadly.
For the purposes of this study, the measurement of craving was reviewed for the following drugs: alcohol, cannabis, cocaine, amphetamines and opioids.

1. **Alcohol**

Alcohol is the most frequently reviewed substance in the craving literature, with reason. According to the 2001-2002 National Epidemiologic Survey on Alcohol and Related Conditions (NESARC), approximately 18 million adults in the United States (8.5 percent of the population) meet DSM criteria for an alcohol use disorder. First outlined by Jellinek in 1946, the pattern of alcoholism includes the concept of an urge or craving for a drink. Jellinek’s seminal research, which used a sample of participants from Alcoholics Anonymous (AA), has been foundational in the current understanding of alcoholism, and recognized craving as a central component of alcohol dependence. It is in AA meetings that the word ‘craving’ is used often to describe the urge for a drink, and the avoidance and/or management of craving is considered paramount to sobriety by its members.

a. **Alcohol craving measurement**

Of the current multi-item self-report questionnaires of various substances, measures of alcohol craving are the most prevalent (Rosenberg, 2009). To illustrate, 9 of the 24 measures covered in a recent review article were for alcohol craving (Rosenberg, 2009), and research on the theoretical aspects of alcohol craving are more developed than for any other substance. The most widely used assessments include the Desires for Alcohol Questionnaire (DAQ) (James, Davies, & Willner, 2004; Love, et al., 1998), the Alcohol Craving Questionnaire (ACQ) (Singleton & Gorelick, 1998). Alcohol craving remains relevant as researchers and clinicians attempt to understand, monitor, and use craving as a structured approach to alcohol treatment (Anton & Drobes, 1998). There is certainly a lot of opportunity for craving
measurement research in alcoholism: Anti-craving medications like naltrexone, disulfiram might reduce cravings and have been found to prevent relapse in alcoholics, and as similar types of drugs for alcohol and other substances are developed, it will be important to be able to measure craving as a self-reported outcome (Anton et al., 2003). Also, as alcohol craving progresses, researchers have greatly expanded their understanding of how certain brain mechanisms potentially underlie alcohol dependency and those parts and processes of the brain that may be associated with craving (Anton, et al., 2003; Love, et al., 1998). As methods in studying alcoholism and, specifically, craving for alcohol, progress, the study of how alcohol craving relates to clinical outcomes will become increasingly reliable (Anton, 1999; Anton, 2000; Flannery, Volpicelli, & Pettinati, 1999; Roberts, Anton, Latham, & Moak, 1999).

2. **Cannabis**

Cannabis, commonly referred to as marijuana, is the most commonly used illicit substance throughout the world (Coffey et al., 2002; Degenhardt et al., 2008; Heishman, et al., 2001; Perkonigg et al., 2008). Most people that use cannabis do not develop dependence, and drug craving is strongly associated with dependence. Dependency in cannabis users, however, is not insignificant. The Substance Abuse and Mental Health Services Administration estimates that 11 million individuals are cannabis users. Cannabis withdrawal has been documented in adolescents and adults seeking treatment, those not seeking treatment, and in human laboratory studies (Heishman, et al., 2001). Research on marijuana craving has been limited, because many individuals that use marijuana do not develop dependence on the drug (Heishman, et al., 2001). However, Budney et al. (1974) reported 93% of cannabis users seeking treatment reported craving, and 44% of these users rated their craving as ‘severe’, suggesting that withdrawal and craving do impact the success
of treatment of cannabis dependence. Among marijuana users who reported a recent attempt to quit using, it was found that patients reported craving to be the most noxious symptom after irritability, and further found that more than half of these participants attributed craving to past failed quit attempts (Budney, Vandrey, Hughes, Thostenson, & Bursac, 2008). Further, in both residential (Haney et al., 2003) and non-residential (Vandrey, Budney, Hughes, & Liguori, 2008) treatment settings, daily marijuana users reported significant increases in craving during periods of forced abstinence.

a. **Cannabis craving measurement**

The gold standard of self-report marijuana craving is the Marijuana Craving Questionnaire (MCQ) (Heishman et al., 2009; Heishman, et al., 2001). It was developed by adapting items from the Quitting Smoking Urges Questionnaire (QSU), which is used to assess tobacco craving, and also contains items adapted from the Alcohol Craving Questionnaire (ACQ) (Heishman, et al., 2001). The authors have also developed a short form of the measure (Heishman et al., 2009), and it had been used in various research settings to test marijuana craving (Alexander, 2003; Budney, Moore, Vandrey, & Hughes, 2003; Singleton, Trotman, Zavahir, Taylor, & Heishman, 2002). To date, there are no other widely used cannabis craving measures.

3. **Cocaine**

According to the World Drug Report (2006), the approximate one-year prevalence of cocaine use in the United States is 2.8%. Although cocaine use has decreased in the past few decades, use and dependence are still a concern. According to the National Survey on Drug Use and Health, Sponsored by SAMHSA and DHHS, in 2008 there were approximately 1.9 million past-month users (359,000 were current crack cocaine users).
Young adults ages 18 to 25 are the group with the heaviest use, at 1.5% past-month cocaine use. In general, men report more use of cocaine than women.

The illicit nature of cocaine and its relatively low prevalence present challenges to studying its use and correlates (van den Bree, Johnson, Neale, & Pickens, 1998). Cocaine abusers are thought to be particularly prone to relapse (McKay, Rutherford, Alterman, Cacciola, & Kaplan, 1995), and reasons for relapse have been attributed to several factors, including craving, in a number of prospective studies. Craving for cocaine is often cited as a precursor to drug relapse in cocaine-dependent individuals who are in recovery from their dependence (McKay, et al., 1995; Wexler et al., 2001), and it is considered a prominent symptom of cocaine dependency (American Psychiatric Association, 2000). The central role that craving is thought to play in cocaine relapse is evidenced by how often it is used as a dependent variable in addiction treatment studies (Sussner, et al., 2006; Tiffany, et al., 1993). Furthermore, utilizing MRI technology, cocaine craving has been shown to significantly alter brain activity in cue-elicited research (Wexler, et al., 2001). Comparing these images to other emotional states in individuals in cocaine-dependence is a potential step in the understanding of how craving affects cocaine users. Although supporters of the use of this type of technology have argued that neurobiology is more reliable than self-report measures, it is also considerably more expensive and complicated to collect. Thus both are important methods in furthering the understanding of craving’s contribution to cocaine relapse.

a. **Cocaine craving measurement**

After alcohol, measures of cocaine craving seem to be the most prevalent in the literature. The most widely used craving measures include the Cocaine Craving Questionnaire (CCQ) (Tiffany, et al., 1993), the Questionnaire of Cocaine Use
(QCU) (Ollo, et al., 1995), and Minnesota Cocaine Craving Scale (MCCS) (Halikas, Kuhn, Crosby, Carlson, & Crea, 1991).

4. **Amphetamines**

Amphetamines are a psychostimulant class of drugs that produce increased focus and wakefulness. The physical dependence of amphetamine is less prominent than other substances, including alcohol and heroin (James, et al., 2004), but ending dependence on amphetamines can be just as difficult as for heroin or cocaine (Srisurapanont, Jarusuraisin, & Kittirattanapaiboon, 2001). Indeed, as amphetamines are prescribed for health problems, ongoing use can lead to dependence on prescription amphetamines. Recreationally, amphetamines are referred to as ‘speed’ and have shown a wide surge as a public health problem in the United States with methamphetamine production taking hold of many communities. The current annual prevalence of amphetamine use in the United States (community-based sample) is estimated at about 1.4% (Maxwell & Rutkowski, 2008). Despite the low prevalence of amphetamine use among the general population, the production and use of amphetamines and methamphetamines is cited as a great drug threat, with escalating rates of use in treatment settings (Gonzales, Mooney, & Rawson, 2010). According to the Substance Abuse and Mental Health Services Administration, U.S. admission to substance abuse treatment programs for abuse/dependency went up from about 40,000 in 1995 to nearly 140,000 in 2007 (Substance Abuse and Mental Health Services Administration, 2009). Additionally, approximately 40% of law enforcement agencies in the US cite it as ‘their greatest drug threat’ (National Drug Intelligence Center, 2007).
a. **Amphetamines craving measurement**

Currently, the only validated multiple item self-report measure for amphetamines is the Desire for Speed Questionnaire (DSQ) developed by James, Davies, and Willner (2004), which was adapted from the Desires for Alcohol Questionnaire (DAQ) (Love, et al., 1998). Self-reported craving for amphetamine has been used as an outcome variable to determine the efficacy of recent psychopharmatherapies for abuse and dependence, like naltrexone and fluoxetine (Maxwell & Rutkowski, 2008; National Drug Intelligence Center, 2007; Srisurapanont, et al., 2001). As more of these anti-craving drugs become available, use of self-reported measures of craving for amphetamines should increase (James, et al., 2004).

5. **Opioids**

Opioids, which have long been used in pain management, have the risk of recreational abuse and dependence. There are a number of different classes of opioids, but those most prone to recreational use include heroin and non-prescription use of drugs that are prescribed to manage pain, including methadone, codeine, morphine, oxycodone, hydromorphone, and opium (Substance Abuse and Mental Health Services Administration, 2009). Although overall prevalence of opioid use is lower than or some of the other major drugs, there is still an increasing number of individuals entering treatment for opioid addiction (McMillan & Gilmore-Thomas, 1996). To highlight, between 1993 and 2003, admission to treatment increased from 1.6 million to 1.8 million (an increase of 14%) (Substance Abuse and Mental Health Services Administration, 2009).

a. **Opioids craving measurement**

Multiple item self-report measures to assess heroin craving have been developed, and have been used to study the efficacy of the use of methadone maintenance
and other anti-craving drugs, like buprenorphine, in heroin- and opioid-dependent individuals (Curran, Bolton, Wanigaratne, & Smyth, 1999; Pérez de los Cobos et al., 2000; Resnick et al., 1992). The two most widely used craving measures for heroin include the Heroin Craving Questionnaire (Tiffany, Fields, Singleton, Haertzen, and Henningfield, 1993), which was adapted from the CCQ, and the Obsessive-Compulsive Drug Use Scale (OCDUS) (Franken, Hendriks, & van den Brink, 2002).

E. **Predictive Validity of Craving**

A key reason to measure craving is its potential place in relapse from substance use after a period of abstinence in substance dependent persons (Drummond, et al., 2000). As discussed, the predictive validity of craving has been questioned, but these researchers point out the lack of reliable and valid measures used in the past.

There are several aspects of the measurement of craving that have lead to the non-conclusiveness of past research. Critically, craving fluctuates over time, and the predictive value of craving depends on how and when it is measured. For example, craving measured at a time point that is far away from a relapse event has low predictive value, but craving that occurs right before relapse signifies a stronger association.

Also, the way that an individual’s craving can change widely across time, and can be an important predictive factor, as well. Someone who has a long, stable craving for a drug may have a higher chance of relapse than someone who has an extreme, salient experience of craving that may create a more disruptive experience (Drummond, et al., 2000).
F. **Properties of Craving Measures**

1. **Single-item measures**

   In many cases, studies that measure craving across substances utilize a single-item approach (Tiffany, Carter, & Singleton, 2000). Single-item craving assessments tend to utilize either a Likert-type rating or a visual analogue scale (VAS) (Rosenberg, 2009). These are typically anchored by phrases such as “none” or “not at all” and “extremely” or “strongest ever experienced.”

   The respondent is then asked to select a number or mark on a line (in the VAS case) to report the degree of craving for a target drug. Advantages of such scales is that they are easy to administer, are appropriate for repeated and frequent measurement, and are sensitive to rapid changes in the psychological construct being measured (McCormick, Horne, & Sheather, 1988). A limitation of using a VAS is that responses are treated as if they are on an interval scale, but these responses can be highly subjective (McCormick, et al., 1988). A larger criticism is that these types of scales do not adequately capture the multiple dimensions of craving (Tiffany, 1992).

   Single item scales assume that a respondent will understand craving as it is described in the single item (Sayette, et al., 2000; Sussner, et al., 2006). Also, these single item scales contain more error variance than multiple item measures (Sussner, et al., 2006; Tiffany & Drobes, 1991). Researchers interested in the topic have suggested that the use of single item measures in craving is not sufficient (Heishman, et al., 2001). As mentioned by Tiffany (1992), single items often have low reliability and are unable to capture the complex nature of craving that is experienced by people in different stages of drug dependence under different conditions. They also do not reflect wide range of theoretical perspectives on drug urges and cravings.
An unreliable scale impedes the ability of researchers or clinicians to detect changes in craving throughout a course of treatment or in laboratory studies (Tiffany, 1997). Another criticism of single-item scales is that any one item may not be able to capture the words a respondent might use to describe their ‘craving’, and would thereby limit the breadth of craving report. Measures of craving are more likely than single-item scales to differentiate levels of craving, both across individuals and treatment settings (Heinz et al., 2006). Also, single items do not capture the myriad ways in which respondents at different stages of dependence may experience craving (Heinz, et al., 2006). Recent findings have suggested that craving is manifested in multiple ways, and may manifest itself in ways that are behavioral, cognitive, and emotional (Connolly, Coffey, Baschnagel, Drobes, & Saladin, 2009).

Because there are likely multi-dimensional features or presentations of craving, it cannot be adequately measured by a single item. This is not to be confused or in conflict with psychometric unidimensionality. The factors/facets of these scales have been shown to be moderately or strongly inter-correlated, and this implies the presence of a unidimensional, higher-order factor (Tiffany, et al., 2000). In fact, analyses of existing multiple-item measures have been shown to be organized hierarchically (Tiffany, et al., 2000), though no hierarchy has been established or tested.

2. **Multiple-item measures**

Multiple item measures perform better for monitoring craving throughout the course of substance abuse treatment, since they are more sensitive than single-item measures (Sussner, et al., 2006). Indeed, asking patients to rate their experience through specific items regarding their craving provides a more detailed assessment than a single item that asks about general craving or urge for a substance, which lacks specificity (Rosenberg, 2009).
a. **Length of measures**

A craving scale that is brief, reliable, and valid is important for measurement of symptoms for substance using and dependent individuals. Too many questions are burdensome and, in the case of craving, may elicit or trigger the desire to use. This is not simply an ethical concern: the impact of self-report measures on craving is important because it has implications for the validity of craving assessments. If completing a craving assessment is a cue that increases the self-report of craving, the validity of the measure can be compromised (Shadel, Niaura, & Abrams, 2001). This was a topic investigated regarding the assessment of smoking craving (The Questionnaire of Smoking Urges), and the authors concluded that the administration of the QSU did not increase craving more than two control groups (Shadel, et al., 2001). Even though more items on a scale improve validity and internal consistency, shorter measures are more desirable for therapeutic and real-world settings (Rosenberg, 2009).

More research is needed in this domain of craving work, and should be considered for both validity and reasons of professional ethics. Any craving measure can be open to reactivity biases in self-reports of craving. Because of this susceptibility, it is desirable to have an instrument that is short enough to avoid undesirable reactions, but lengthy enough to be sufficiently reliable and valid while capturing the different nuances of self-reported craving. To provide an example, different people may use different words to express ‘crave’ (such as ‘urge’ or ‘desire’), and these may mean different things to different people.

b. **Time frame**

The assessment of craving is complicated because it can be thought of as a persistent inclination to seek substances and also as a short-term urge (Ferguson & Shiffman, 2009). How a craving measure is administered is important, given the specific predictions and
associations one is trying to test. Existing measures have tended to focus on real-time craving, and it is suggested that this is preferred to questionnaires regarding a long-term look back period. For example, measures that ask the respondent to report on past craving may present problems with interpretation because it might be unclear if the respondent was answering items based on a particularly salient experience of craving, an ‘average’ of past craving experiences, or something else (Rosenberg, 2009).

G. **Review of Existing Self-Report Measures of Drug Craving**

There is no ‘gold standard’ of measurement for drug craving (Anton, 1999; Harold, 2009; Rosenberg, 2009; Sayette, et al., 2000). A recent comprehensive review of the literature (Rosenberg, 2009) noted that the past 20 years have seen a dramatic increase in the amount of multi-item, self report questionnaires for craving measures, for both illicit and licit drugs. To date, these instruments are drug specific, and none are considered a ‘gold standard’ tool. In fact, as confirmed by the review by Rosenberg (2009), many of these instruments borrow items from existing questionnaires. However, these measures are explicitly designed to assess craving and have been shown to have practical utility in natural, clinical, and laboratory settings. Since the nature of drug craving can be different across drugs, it there are advantages to having separate measurement scales. However, as these questionnaires have developed over the past twenty years, these items, in addition to being copied from other scales, were selected to reflect a conceptualization of addiction or craving rather than any outcomes that were drug specific (Rosenberg, 2009).
The items of the proposed craving scale build upon the work of several of these validated measures. For the purposes of review, the development of these scales and the procedures used to establish reliability and validity will be presented. The psychometric properties of these measures are summarized in TABLE I.
**TABLE I**

MEASURES USED IN THE DEVELOPMENT OF THE SUBSTANCE CRAVING SCALE

<table>
<thead>
<tr>
<th>Authors (year)</th>
<th>Name of Measure</th>
<th>Drug</th>
<th>Sample used for validation</th>
<th>Number and type of items</th>
<th>Response Format</th>
<th>Reliability, factor/subscale alphas</th>
</tr>
</thead>
<tbody>
<tr>
<td>James, D., Davies, G., Willner, P. (2004).</td>
<td>Desire for Speed Questionnaire (DSQ)</td>
<td>Amphetamines</td>
<td>N=174 amphetamine users in treatment and community (Wales)</td>
<td>40 (adapted from DAQ, added items on impaired control)</td>
<td>7-point scale “strongly disagree to strongly agree”</td>
<td>Subscale alphas= .92 (positive and negative reinforcement), .93 (strong desires and intentions), .89 (mild desires), .84 (impaired control)</td>
</tr>
<tr>
<td>Tiffany, S., Singleton, E., Haertzen, C., &amp; Henningfield, J. (1993).</td>
<td>Cocaine Craving Questionnaire (CCQ)</td>
<td>Cocaine</td>
<td>N=225 cocaine abusers (USA)</td>
<td>45 (many adapted from QSU)</td>
<td>7-point scale “strongly disagree to strongly agree”</td>
<td>Alpha=.93</td>
</tr>
<tr>
<td>Heishman, S., Singleton, E., &amp; Liguori, A. (2001).</td>
<td>Marijuana Craving Questionnaire (MCQ)</td>
<td>Marijuana</td>
<td>N= 217 regular marijuana smokers (USA)</td>
<td>45 (adapted from QSU, ACQ)</td>
<td>7-point scale “strongly disagree to strongly agree”</td>
<td>Subscale alphas= .82 (compulsivity or loss of control), .78 (emotionality or anticipated relief of negative emotions), .55 (positive outcomes), .68 (immediate intention to use)</td>
</tr>
<tr>
<td>Heishman, S., Singleton, E., &amp; Moolchan, E. (2003).</td>
<td>Tobacco Craving Questionnaire (TCQ)</td>
<td>Tobacco</td>
<td>N=213 cigarette smokers (USA)</td>
<td>45 (some adapted from Quitting Smoking Urges by Tiffany and Drobes, 1991)</td>
<td>7-point scale “strongly disagree to strongly agree”</td>
<td>Subscale alphas from .48-.82</td>
</tr>
<tr>
<td>Mackillop, J. (2006).</td>
<td>Alcohol Urge Questionnaire (AUQ)</td>
<td>Alcohol</td>
<td>N=351 abstinent alcoholics (USA)</td>
<td>8 items</td>
<td>7-point scale “strongly disagree to strongly agree”</td>
<td>Alpha = .91</td>
</tr>
<tr>
<td>Singleton, E., Tiffany, S. &amp; Henningfield, J. (1994).</td>
<td>Alcohol Craving Questionnaire (ACQ)</td>
<td>Alcohol</td>
<td>N=217 drinkers, not with alcohol problems</td>
<td>45 (some adapted from Quitting Smoking Urges by Tiffany and Drobes, 1991)</td>
<td>7-point scale “strongly disagree to strongly agree”</td>
<td>Alpha=.96</td>
</tr>
<tr>
<td>Love, A., James, D, &amp; Willner, P. (1998).</td>
<td>Desires for Alcohol Questionnaire (DAQ)</td>
<td>Alcohol</td>
<td>N=380 recreational drinkers, n=131 alcohol-dependent drinkers</td>
<td>36 (based on unpublished article)</td>
<td>7-point scale, no anchors provided</td>
<td>Subscale alphas: .97 (negative states and positive outcomes), .97 (strong desires and intentions). .95 (mild desires and intentions)</td>
</tr>
</tbody>
</table>
H. Preliminary Conceptual Work: Development of the Substance Craving Scale

The results across the previously cited studies found that several of the factor structures were the same across substances (Franken, et al., 2002; James, et al., 2004; Love, et al., 1998). This supports recent conclusions by researchers that dimensions of craving found for drugs may be more general, rather than drug specific. If this is the case, there is no need to have several substance-specific measures when a single one could suffice. Furthermore, with the advancements of quantitative techniques that can control for the unique effects of drugs (e.g., hierarchical models that allow for items nested within drug), it is suggested that more parsimonious models can better describe craving for substances.

The current scale (Mazza, 2011) was developed based on several of the prior validated, drug-specific scales that are used to measure substance craving (Heishman, et al., 2001; James, et al., 2004; Love, et al., 1998; Sussner, et al., 2006; Tiffany, et al., 1993). They were chosen based on their prominence in the literature. The psychometric properties of these were detailed earlier in the literature review section, and were summarized in TABLE 1.

The procedure for selection of the items included a crosswalk that included a list of all of the items for all of the scales that were reviewed. Then, common items that were included across scales were retained. An additional craving item, used on the National Longitudinal Alcohol Epidemiologic Survey, was also included in the final scale.

1. Construct map

A construct map can be thought of the depiction of a unidimensional latent variable. It is useful to consider the construct map to have some anchors e.g., from low to high, rare to frequent, and that there exists a range of space where a respondent or item response can be placed along such a continuum. It is assumed in the construct map that there is a clear
definition of the construct or latent trait being measured, and that there is some ordering to the underlying continuum. The construct map depicts two main ideas: a respondent map, and an item response map.

The important aspects of the construct map are that some continuum or hierarchy exists, that respondents and item responses are related, and that there are some stages or range between the anchors of the construct map. Developing a construct map is ideal in ruling out any closely related constructs. In the case of substance craving symptoms, withdrawal symptoms might be seen as a rival construct. One side of the map may be more conceptually developed first: if one were to start with, say, responses that measured substance craving, theoretical development would need to be carried out to determine how respondents might be classified on the continuum. In the current example, different classes of outcomes were modeled: respondents with a low amount of craving might be classified as not needing intensive treatment, whereas those individuals that are ‘higher’ on the construct map might need more intensive intervention. One will note that the construct map and subsequent ‘building blocks’ that Wilson presents is agreeably related to the Rasch single parameter model and its basic qualitative principles (Bond & Fox, 2007). The following figure provides a proposed theoretical hierarchy of substance craving severity, and presents all of the items contained in the Substance Craving Scale.
**Figure 1.** Substance craving scale concept map.

In addition to confirming the construct validity of the measure, this hierarchy can be tested for its invariance across substance of choice, via a Differential Item Functioning analysis. Testing for DIF also allows for determining that substance craving and its severity is the same for all substances. Examination of (DIF) allows us to test whether items reflect significantly different levels of symptom severity for different groups, i.e., differing item

<table>
<thead>
<tr>
<th>RESPONDENTS</th>
<th>RESPONSES TO ITEMS</th>
</tr>
</thead>
</table>
| Severe craving | - I am going to use alcohol or other drugs as soon as I possibly can.  
                  - I would do almost anything to use alcohol or other drugs.  
                  - All I want to do is use alcohol or other drugs.  
                  - Right now, I want to use alcohol or other drugs so bad I can almost taste it.  
                  - I want to use alcohol or other drugs so badly that I can't think of anything else. |
| Moderate craving | - My desire for alcohol or other drugs seems overwhelming.  
                     - Nothing would be better than alcohol or other drugs right now.  
                     - I would not be able to control how much alcohol or other drugs I used if I had some. |
| Low craving | - I crave alcohol or other drugs right now.  
               - I have an urge for alcohol or other drugs.  
               - Using alcohol or other drugs would make things seem just perfect.  
               - If I were using alcohol or other drugs I would feel less nervous. |
| No craving | -                              |

More craving
calibrations. This analysis examines the validity issue of construct irrelevance for groups (Messick, 1995).

2. **Rationale for Rasch Measurement**

All of the substance-specific measures were validated using classical test theory techniques, and not the Rasch model. In fact, there is only one known dissertation that has applied the Rasch model in substance craving, for a scale of benzodiazepines (Mol, 2007). To date, there has been no Rasch analysis conducted of a substance craving scale that assesses cravings for multiple substances. The use of quantitative psychological models, such as item response theory (IRT) and latent-variable models, has allowed conceptualization of psychopathology in ways that are more continuous, rather than categorical (Conrad et al., 2012; Krueger & Markon, 2006). This approach is especially useful in the context of substance abuse/dependence, where growing evidence exists that abuse and dependence comprise a single latent variable with symptoms that are arrayed across the spectrum of severity (Keyes, et al., 2011; Krueger & Markon, 2006; Martin, Chung, Kirisci, & Langenbucher, 2006; Saha, Chou, & Grant, 2006).

Moreover, although previous factor analyses have revealed the varying components that contribute to the experience craving, there little literature to suggest a hierarchy of severity of craving symptoms (i.e., does negative reinforcement contribute to a more severe experience of craving than positive reinforcement). Modern measurement techniques based in item response theory (Embretson & Reise, 2000), such as Rasch measurement models (Rasch, 1960, 1980; Wright & Stone, 1979), hold promise in addressing these issues, since these models can generate item severity calibrations based on the probability of item endorsement. For an individual, a probabilistic estimate of a person’s severity level is obtained based on his or her
endorsement of items whose severity has been estimated as a unidimensional measure. If items
misfit, this may suggest that they should be revised or dropped. Additionally examination of the
severity hierarchy by demographic characteristics enables us to estimate whether persons in
different groups (e.g., gender, age groups, different drug of choice) respond in systematically
different ways, perhaps indicated bias, even though they are at the same level on the construct.

The aforementioned multi-item scales used to measure substance craving do not
contribute an understanding of the severity of craving, and how such measures could be used to
triage patients for treatment, and to track outcomes. Rasch analysis will provide insight on item
severity, and could contribute to a broader understanding and ability to triage substance-
dependent patients.

I. Demonstrating Reliability and Validity of the Substance Craving Scale

To determine if the new scale is appropriate for use in clinical and research settings, a
thorough examination of reliability and validity is necessary. As explained by Messick (1989),
validity is a matter of degree, where evidence is compiled to justify the use of a scale for its
potential use. As such, the Substance Craving Scale should be explored for reliability and
validity at both the individual item level and for the measure as a whole. The validation of the
items and the measure are testing the hypotheses that the items and the measure are sufficiently
useful in measuring substance craving (Messick, 1989). The validation techniques that will be
used in the examination of the scale are both theory-driven and data-driven, as Messick suggests.
Together, the conclusions will build the argument for or against construct validity of the scale.

Before a thorough examination of construct validity can be conducted, it is important that
the scale is appropriately reliable and unidimensional. These tests will be conducted as a
precursor to an examination of construct validity.
1. **Reliability**

Reliability is an important aspect of construct validity (Messick, 1989). It refers to how accurate and precise a measure is, and informs a researcher that scores generated by a measure are both reproducible and consistent (Thorndike, 1997). To reiterate, reliability typically frames as how accurately a test measures what it intends to measure, and although non-reliability (that is, inconsistency) is present in all measurement procedures, errors in measurement that are broadly systematic invite critique. Strong reliability is necessary to establish construct validity (Messick, 1989).

2. **Dimensionality**

Testing for unidimensionality is an important step in any psychometric analysis, and is most valuable when ideas surrounding a particular construct are in need of clarification. This pertains to substance craving, since definitions of craving vary across previous studies (Connolly, et al., 2009; Harold, 2009; Heishman, et al., 2001; Merikle, 1999; Sayette, Martin, Wertz, Shiffman, & Perrott, 2001; Tiffany & Wray, 2009). Prior to examining the evidence for construct validity, a testing of unidimensionality will be employed.

While some researchers define craving explicitly and sufficiently as the desire to use a substance (Sayette et al., 2000), others believe that craving refers to a broader definition or set of factors, including such varying factors such as anticipation of positive effects when using, desire to alleviate withdrawal symptoms through use, and the explicit purposefulness of a user to acquire and use drugs or alcohol (James, Davies, & Willner, 2004). This suggests that a scale intended to be a unidimensional measure may contain competing, or rival, factors. The most likely rival factor in a craving scale would be items that relate to physical withdrawal, and the desire to alleviate withdrawal through administering the addictive substance (Drummond, 2001).
3. **Validity of the items**

To establish the validity of the items on the Substance Craving Scale, both item quality and differential item functioning should be examined. First, items should be examined for the appropriateness in relation to the construct, by examining their fit statistics. Also, the items should be tested for invariance across substances. The items should function the same across substances for them to be considered valid for the measurement of craving for any drug.

a. **Item Fit**

Rasch analysis provides fit statistics to test assumptions of fundamental measurement for items (Wright & Stone, 1979). The Rasch model provides two indicators of misfit: infit and outfit. The infit is sensitive to unexpected behavior (e.g., random responding), affecting responses to items near the person ability level or item difficulty level, and the outfit is outlier sensitive. Understanding poor fit can lead to the improvement or dropping of items because they fail to perform in a manner consistent with the principal measurement dimension.

For example, the item “I would not be able to control how much alcohol or other drugs I used if I had some”, which is used on some current scales, might be diagnosed as a misfitting item since it may measure a separate construct (such as impulse control). Other problem items might include those that are endorsed by most respondents, and might not be useful in providing information about the construct or the persons.

b. **Differential Item Functioning**

Differential Item Functioning (DIF) is an analysis that can be performed to determine if item calibrations vary significantly among subgroups of the population being measured (Masters, 1988; Smith, 1992). Presence of significant DIF on item is an indicator that the item may be biased against a group, or simply may function differently for the group. DIF
analysis addresses an important construct validity criterion concerning the comparability and fairness of items, their calibrations, and the interpretation of the resulting scores (Messick, 1995). Furthermore, significant DIF can be an indication that specific item calibrations should be set for different subgroups (Conrad, et al., 2007). Common DIF contrasts include male/female, racial/ethnic group comparisons, and age (youth vs. adults). To highlight an example, in their DIF analysis of a substance use measure, Conrad and colleagues (2007) found that significant DIF existed for 8 of 16 items. These conclusions led the researchers to conclude that age-specific calibrations would produce a more reliable measure for these subgroups.

Perhaps the most critical contrasts to be employed for the Substance Craving Scale will be for drug of choice, as the self-report measure for craving intends to be general across types of drugs. However, given the availability of DIF calibration techniques, it would be possible to set different item calibrations for the craving measure by substance of choice. Such a refinement would allow persons to be measured on the same Rasch ruler, regardless of substance of choice.

If the items of Substance Craving Scale are not invariant across subgroups, then changes to the measures could be considered, e.g., dropping items and developing new ones or developing separate measures for certain subgroups. Findings of differential item functioning might also have theoretical implications for the treatment of craving depending on the subgroup.

4. **Construct and criterion validity of the measure**

Several types of validity should be examined to establish the appropriateness of the test. Perhaps the most important test is that of construct validity: the degree to which the items account for the latent construct, \( \theta \). Again, it is useful to think of the many different types of validity as evidence of a strong measure (Trochim & Donnelly, 2008). The analysis of the
Substance Craving Scale should include an examination of both construct and criterion-related validity.

a. **Construct validity**

Construct validity can be tested in different ways. According to Rasch theorists, the item hierarchy provided by the item difficulty estimates support construct validity (Smith, 2001). Rasch analysis provides this item severity hierarchy using the endorsement probabilities of persons and items on a single logit scale, which is displayed on a Wright map (Wilson, 2005). It is argued that the items of a scale should form a hierarchy with low severity items on the bottom, and higher severity items on top. In a craving measure, we expect to see a low severity item (e.g., ‘I would not be able to control how much alcohol or other drugs I used if I had some’) at the bottom of such a hierarchy, and a higher severity item (e.g., ‘I would do almost anything to use alcohol or other drugs.’). Again, this hierarchy should both confirm and inform the construct map (Wilson, 2005) generated at the earlier stage of measure development.

Additionally, other tests of construct validity, such as correlations with scales that are also designed to measure craving were used to support construct validity. Such correlations are related to criterion-related validity, to be discussed next.

b. **Criterion validity**

Criterion-related validity is performed by making predictions of how a measure performs in relation to other measures (Trochim & Donnelly, 2008). Perhaps the best illustration of criterion validity expectation is thoroughly addressed by Campbell & Fiske’s multitrait-multimethod matrix (MTMM) (1959). The MTMM proposed by Campbell and Fiske allows a researcher to validate a measure by testing correlations of a measure with other theoretically related (or, theoretically unrelated) tests. In their seminal paper, the researchers
introduced two important construct validity-related concepts: convergent and discriminant validity (Trochim & Donnelly, 2008). Convergent validity involves testing the hypothesis that the measure of interest is positively correlated with other measures that are, theoretically, related.

To examine the correlates of craving, with other variables as criterion validity criteria (Campbell & Fiske, 1959; Messick, 1989), a pattern of expected correlations will be proposed, where high is more than .5, moderate is more than .3, and low is less than .3. All measures that were available in the Pathways protocol (the battery of measures administered to current participants) were examined for their potential use in the hypothesized correlations. TABLE II shows the measures and variables selected to test criterion validity.
<table>
<thead>
<tr>
<th># of Items</th>
<th>Short Description</th>
<th>Interpretation</th>
<th>Hypothesized Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Substance Problem Scale</td>
<td>16 Dichotomous scale that measures symptoms of substance abuse, dependence, or substance-induced disorders.</td>
<td>Higher scores on this scale represent greater severity of drug problems. The scale includes physiological, psychological and social criteria.</td>
<td>High (&gt; .5)</td>
</tr>
<tr>
<td>Withdrawal</td>
<td>1 Recency of withdrawal problems.</td>
<td>Higher score indicates more recent withdrawal.</td>
<td>High (&gt; .5)</td>
</tr>
<tr>
<td>Substance Frequency</td>
<td>8 Average percent of days (of the past 30) reported of any alcohol or other drug (AOD) use; heavy AOD use; problems from AOD use; and days of alcohol, marijuana, crack/cocaine, heroin/opioid and other drug use.</td>
<td>Higher scores represent increasing frequencies of substance use in terms of days, days staying high most of the day and days actually causing problems.</td>
<td>High (&gt; .5)</td>
</tr>
<tr>
<td>Recovery Environment Index</td>
<td>13 Measure of the past 90 days of alcohol in the home, drug use in the home, fighting, victimization, being homeless, and structured activities that involved substance use and inverse percentage of days going to self-help meetings, and involvement in structured substance-free activities.</td>
<td>Higher scores indicate less involvement in supportive activities and more environmental risk from alcohol/drug use in the home, fighting and/or victimization.</td>
<td>Moderate (&gt; .3)</td>
</tr>
<tr>
<td>Substance Recency</td>
<td>1 Computed variable of recency of use of any type of drug.</td>
<td>Higher score indicates more recent use.</td>
<td>Moderate (&gt; .3)</td>
</tr>
<tr>
<td>Perceived Social Support Scale</td>
<td>20 Dichotomous scale that measures positive types of social support.</td>
<td>Higher scores on this scale represent more perceived positive types of social support.</td>
<td>Low (&lt; .3)</td>
</tr>
</tbody>
</table>
High correlations (>.5) are expected with variables derived from constructs that are theoretically closely related to craving for substances. In this case, those individuals with high measures of craving should also have a high measure of substance problems, as indicated by their measure on the Substance Problem Scale. Also, a high correlation was also expected for a specific item of the Substance Problem Scale, Withdrawal Problems. It was also hypothesized that Substance Frequency would be highly correlated with craving.

In contrast, only a moderate correlation (>.3) should be expected with constructs that are moderately related to craving (i.e., Recovery Environment Risk Index, RERI). The RERI was chosen to be moderately correlated with craving as the evidence shows that that individuals with groups of friends that do not use substances may lower individual craving, and those individuals with friends that do use substance may increase individual craving (Hunter-Reel, McCrady, & Hildebrandt, 2009). Related, continued association with substance-using peers predicts relapse for substance users, whereas association with non-substance using peers predicts abstinence (Brewer, Catalano, Haggerty, Gainey, & Fleming, 1998; Havassy, Hall, & Wasserman, 1991; Havassy, Wasserman, & Hall, 1995; Schroeder et al., 2001). In addition to the RERI, it was hypothesized that recency of substance use (any drug) would also be moderate. Although this variable is a construct related to craving, the fact that this variable accounts for any drug use, it should not be strongly related to craving for a substance of choice.

Similarly, a measure of craving should have a low correlation (<.3) with a measure of positive social support, as evidence suggests that positive social support should be negatively correlated with craving (Schroeder, et al., 2001). The Perceived Social Support Scale
is a 20-item dichotomous scale that measures positive types of social support (Procidano & Heller, 1983).

c. **Multivariate models**

Given the hypothesized correlates of craving reviewed in the literature, it is appropriate to construct a model that would explain the effect of the various predictors on the experience of craving. Confirmation of these relationships will contribute to the validity argument of the Substance Craving Scale.

First, the Substance Problem Scale should have the strongest effect on craving, as researchers have reported very strong, positive associations between self-reported craving with severity of dependence (Anton, Moak, & Latham, 1996; Collins, et al., 2001; Drummond & Phillips, 2002; Klein, et al., 2007; Leonard, et al., 1988). Craving is seen as a very strong contributor to drug (American Psychiatric Association, 2000; Heishman, et al., 2001; Keyes, Krueger, Grant, & Hasin, 2011). Since craving is considered to be strongly present in those with abuse and dependency problems, the Substance Problem Scale likely has the strongest effect on the outcome of craving.

Frequency should also have a similarly strong effect on craving, and will be entered next into the model. It has been found that frequency of use results in higher craving, though the direction causality of this relationship has not been determined (Ferguson & Shiffman, 2009).

Measures of craving appear to be sensitive to withdrawal, above and beyond substance abuse and dependence criteria being met (Baker, Japuntich, Hogle, McCarthy, & Curtin, 2006; Baker, Piper, McCarthy, Majeskie, & Fiore, 2004). The next variable that should
be entered into the model, withdrawal, was selected for its effect on craving, above and beyond the Substance Problem Scale.

Another variable thought to have a positive effect, recency of drug use, was entered in last to see if it had an independent effect on craving. Researchers have found significant associations between past consumption for alcohol (Anton, 2000; Bohn, Barton, & Barron, 1996; Collins, Koutsky, Morsheimer, & MacLean, 2001; Drummond & Phillips, 2002; Klein, Stasiewicz, Koutsky, Bradizza, & Coffey, 2007; Leonard, Harwood, & Blane, 1988; Love, et al., 1998; McEvoy, Stritzke, French, Lang, & Ketterman, 2004; Mol et al., 2003), cocaine (Fox et al., 2005; Tiffany, Singleton, Haertzen, & Henningfield, 1993), and benzodiazepines (Mol, et al., 2003). However, the variable that was used here is a composite variable for the use of any substance, and thus the effect on craving should not be very strong.

There is an assumption that exposure to substance using peers in a contributing factor in eliciting craving, though this relationship would likely be very weak (Brewer, Catalano, Haggerty, Gainey, & Fleming, 1998; Havassy, Hall, & Wasserman, 1991; Havassy, Wasserman, & Hall, 1995; Schroeder et al., 2001). Therefore, the Recovery Environment Risk Index should have a moderate effect on craving.

The last variable selected for entry into the regression model was that variable hypothesized to have a protective effect. The Perceived Social Support Scale was entered to determine if there was an independent protective effect on the craving. It was hypothesized that this variable would have a unique protective effect on craving, since it is suggested that social support should be negatively correlated with craving (Schroeder, et al., 2001).
Figure 2 shows a theoretical model of the hypothesized relationships and the directions and strength of their correlations. The figure shows the strong effect of substance problems on craving, in addition to its association with withdrawal. Substance frequency is also a strong predictor of craving. Substance recency and risk environment have moderate associations with craving, and positive social support indicates a protective effect.

Figure 2. Theoretical model of craving with hypothesized correlations.
J. Research Objectives of the Current Study

In summary, the review of the literature presented the theoretical nature of craving, the drug-specific measures of craving that are currently in use, and the preliminary conceptual work which led to the development of the Substance Craving Scale. The following research objectives will be paramount in establishing the scale as an appropriate measure.

1. Reliability
2. Unidimensionality
3. Item Quality
4. Invariance of hierarchy
5. Construct and Criterion Validity

The conceptual and theoretical support for establishing reliability and validity of the scale was outlined in this chapter, and the testing of these research objectives with decision criteria will follow in the Methods chapter.
III. METHODS

This chapter outlines the Rasch measurement model which will be used to examine the Substance Craving Scale. The research objectives of this study are to determine the reliability and unidimensionality of the Substance Craving Scale, and to establish the validity of the items and the complete measure. In this chapter, the analysis plan for establishing reliability and validity are presented. The criteria for decision-making are described for all of the proposed objectives.

A. Rasch Measurement Model

The Rasch model is a special class of an item response theory (IRT) model, with a single item parameter. In the dichotomous case, the model is a probability model that estimates response to an item by a respondent as:

\[
\ln \frac{P_{ni}}{1 - P_{ni}} = B_n - D_i,
\]

where \(\ln\) is the natural logarithm, \(P_{ni}\) is the probability of respondent \(n\) endorsing craving item \(i\), \(1 - P_{ni}\) is the probability of respondent \(n\) not endorsing craving item \(i\), \(B_n\) is the person measure of craving for respondent \(n\), and \(D_i\) is the difficulty of craving item \(i\) (Bond & Fox, 2007).

The principles of the Rasch model, as presented by Bond & Fox (2007), are 1) each person is characterized by their ‘ability’, 2) each item is characterized by its difficulty, 3) each of these estimates can be expressed by a number, 4) and the difference between these numbers can be used to compute the probability of any particular response.

Person ability and item difficulty are both measured in logits. During the estimation process, person ability is estimated from the total number of correct responses by calculating each person’s success to failure odds. Item difficulty is estimated from the total number of
correct responses by calculating each item’s failure to success odds. Following these maximum likelihood estimates, logarithmic transformations of the item and person scores are converted from ordinal data to yield interval measures, through use of a logarithmic conversion. The item and person probabilities of the data determine the interval sizes. These linear, interval measures are an attractive quality of the Rasch model since they provide meaningful differences in person and item classifications. In contrast, classical test theory relies on raw scores, which are biased in that they are treated as linear measures in analyses when they are, in fact, not truly linear (Wright, 1997).

Further, three assumptions must be satisfied for any IRT model, including the Rasch single-parameter model. First, that there is a unidimensional latent trait being modeled, and ability on this trait is denoted by \( \theta \) (theta). Second, at any specific value of \( \theta \), there is no correlation between responses to items on a scale (this is known as the condition of local independence). Third, the response functions are monotonic. That is, as ability increases, the probability of answering correctly or endorsing a higher category increases. Beyond this class of assumptions, the Rasch model also includes an additional assumption of invariant item ordering, which states that it is possible to assign indices to each item such that item difficulty is ordered in the same fashion for all persons (Sijtsma & Junker, 1996). Invariant item ordering allows for the Rasch hierarchy to hold substantive meaning, and to hold items to their placement on the map.

A self-report measure of craving would be appropriately used at intake upon entering a substance use treatment program; and for the purposes of capturing treatment effects; it could also be administered periodically during treatment, or at treatment completion.
**B. Data Collection and Sample**

1. **Data collection**

   The research site was located in Chicago, IL. All data were collected by Chestnut Health Systems, Inc., a non-profit agency involved in the treatment and research of substance use (Scott, 2011). All data were de-identified by Chestnut Health Systems, and the principal investigator had no means of identifying respondents. All data were held on an encrypted database with no linked health or patient information.

2. **Sample**

   The sample included a cohort of 1,043 adults who had presented to publicly funded substance abuse treatment between 1996 and 1998 operating on the west side of Chicago, who are being followed longitudinally for a variety of health and substance-related outcomes. The racial/ethnic demographics of the sample were 89% (n=933) African American, 4% (n=43) Caucasian, 6% (n=64) Hispanic, and .2% (n=3) other. The gender distributions were 62% female (n=644) and 38% (n=399) male.

**C. Analysis Plan**

   Analysis of the data employed both Rasch and classical test theory techniques, as appropriate. The measure was analyzed with a Rasch dichotomous model (Rasch, 1960; B.D. Wright & Stone, 1979) with Winsteps (Version 3.72.0) statistical software (Linacre, 2011), and SPSS (SPSS Inc., 2009). Analyses were performed to provide summary statistics of the data, to determine fit statistics, dimensionality, item and person hierarchy, and differential item functioning (DIF) for substance of choice and gender. Further tests of construct validity were
conducted, including concurrent validity with theorized correlates of craving. The psychometric analyses and criteria for making psychometric decisions are detailed below.

1. **Objective 1: Reliability**

   Two forms of reliability analyses were performed, including Cronbach’s alpha and Rasch person reliability.

   a. **Cronbach’s alpha**

      Cronbach’s alpha estimates the internal consistency of the measure’s items by using responses obtained from a single administration of the tests (Crocker & Algina, 2006). As such, the criteria for a sufficient Cronbach’s alpha for the measure was set at .80. Cronbach’s alpha was calculated for the entire sample, and was also calculated individually by drug of choice.

   b. **Rasch reliability**

      Both Cronbach’s alpha and Rasch item/person reliability statistics provide estimates of the proportion of variance of the person scores or measures to total variance (Bond & Fox, 2007; Nunnally & Bernstein, 1994). However, Rasch person/item reliability reflects the reliability of the placement of both on the measurement scale, excluding those with measures reflecting extreme scores (zero or all yes). Rasch person reliability, which is analogous to Cronbach’s alpha, was expected to meet a .80 criterion (the same as for the alpha). Rasch person reliability was calculated for the entire sample, and was also be calculated by drug of choice. Item reliability, which is reflective of reproducibility of the order of persons on the scale if given a parallel test, was expected to be above the .80 criterion.
2. **Objective 2: Unidimensionality**

A principal components analysis of Rasch residuals was performed to test for dimensionality. This analysis was used to drive qualitative understanding of the construct, and assure that the unidimensional requirement of measurement prior is met.

In Rasch analysis, the data are first constructed into a linear measure through typical Rasch analysis estimation, and then a principal components analysis of the residuals that remain after the Rasch analysis is conducted. The Rasch principal components analysis of the residuals is used to detect common variances that are left unmodeled after the Rasch measure has been formed (Bond & Fox, 2007). This allows for the detection of a substantial factor in the residuals after a primary measurement dimension (in this case, craving) is estimated.

Criteria that were used for unidimensionality are as follows: first, variance explained by the measurement dimension was to be greater than 40%. Reckase (1979) uses 20% as a criterion for a substantial factor, so 40% is quite conservative. Unexplained variance in the first contrast of the data should also be low, and fall under the criterion of 15% for a rival factor. Moreover, additional criterion for unidimensionality included a 3 to 1 ratio of variance explained by the measurement to variance in the first component of residuals.

3. **Objective 3: Item quality**

Under Wilson’s (2005) criteria of more than 1.33 and less than .75, an item will be regarded as misfitting if its mean squares on both infit and outfit are higher than 1.33 or lower than .75 (overfit). These items much have a significant ZSTD (this is a t-statistic, so acceptable values are those accepted for t, which is -2 to +2). These criteria are appropriate for even large samples (Bond & Fox, 2007).
Another useful diagnostic provided by the person/item maps of the Rasch model is that of how well the items are centered on the population of interest. As a criterion, the mean of the items and the mean of the persons should be within one logit of each other to indicate that the items fit the population of interest. This would be an indication that the items are appropriate for the target population. Additionally, items should have an appropriate spread that ranges across the span of persons measured to capture the wide range of variability of person abilities on the construct (Bond & Fox, 2007).

4. **Objective 4. Invariance of hierarchy (using DIF)**

In the analysis of a craving scale, DIF analyses were performed, with contrasts for substance of choice and gender to determine whether at a given level of the trait difficulty varies significantly for these subgroups. Because the sample was predominantly African-American (89%), DIF contrasts for race/ethnicity were not performed. DIF by substance of choice was examined closely for differences in item response by drug.

Standards for what is considered an important DIF effect size vary from .4 to .6 logits (Draba, 1977; Elder, McNamara, & Congdon, 2003; Scheuneman & Subhiyah, 1998; Wang, 2000). To satisfy criterion for comparability among a diverse population, no DIF contrasts were to exceed .6 logits.

5. **Objective 5: Construct and Criterion Validity**

a. **Construct validity**

The item map (Wright Map) will be examined for its support of construct validity. The items on the Substance Craving Scale should form a hierarchy with low severity items on the bottom, and higher severity items on top. These items and their placement were presented in the construct map in the literature review (see *Figure 1*). Also, the Wright Map
should also present the items with a reasonable spread along the logit scale, and the mean of the items and persons should be within a single logit.

b. **Criterion validity using hypothesized correlations**

A pattern of expected correlations were proposed, where high is more than .5, moderate is more than .3, and low is less than .3. High correlations (> .5) were expected with the Substance Problem Scale, Withdrawal Items, and Substance Frequency. A moderate correlation (> .3) was expected with the Recovery Environment Risk Index. A low correlation (< .3) was expected with the Perceived Social Support Scale.

c. **Multivariate models**

To further confirm the results of the criterion validity tests, regression models were constructed with the hypothesized correlates using the linear regression function in SPSS software.

All regression equations take on the basic algebraic function, which can be expressed as: Predicted response = Average parameters + Random or Variance parameters (Subramanian, Jones, & Duncan, 2003).

Such a model might be represented statistically as:

\[ y_i = \beta_0 x_{0i} + \beta_1 x_{1i} + \ldots + \beta_n x_{ni} + (e_{0X_{0i}}) \]

where \( y_i \) is the dependent variable for an individual, \( \beta_0 \) is the intercept parameter estimate when all other parameters (\( x_{1i} - x_{ni} \)) are held constant, and all other \( \beta \) parameters are slopes for average change per average \( x \). The intercept and slope parameters are fixed portions of the regression model, and the random part of the model (\( e_{0X_{0i}} \)) represents any individual variation from the fixed regression line.
With the appropriate reliable measures, covariates at the individual level can be modeled for their contribution to the dependent outcome. Within the context of substance use craving, multiple linear regression techniques were employed to adequately model the influences of individual factors on individual substance craving. Further, in the interest of parsimony, a linear regression model will adequately explain individual variance without adding unnecessary parameters into the predicted equation.

The variables were entered into the models based on the hypothesized effect size of each of the independent variables on the dependent variable, which was the total score of the craving measure. These were based on the theorized relationships explained in the Literature Review.

1) **Linear regression model**

Conceptually, the final linear regression model to predict craving was:

\[
Y_{\text{craving}} = \beta_0 x_{0i} + \beta_{\text{SPS}} x_{1i} + \beta_{\text{Freq}} x_{2i} + \beta_{\text{Wthd}} x_{3i} + \beta_{\text{RERI}} x_{4i} + \beta_{\text{PSS}} x_{5i} + (\epsilon_{0i} x_{0i});
\]

where \(y_{\text{craving}}\) is craving for an individual and \(\beta_0\) is the intercept parameter estimate when all other parameters are held constant. The other \(\beta\) parameters are slopes for average change per average \(x\). \(\beta_{\text{SPS}}\) represents the slope for the Substance Problem Scale, \(\beta_{\text{Freq}}\) represents the slope for Substance Frequency, \(\beta_{\text{Wthd}}\) represents the slope for Substance Withdrawal, \(\beta_{\text{RERI}}\) represents the slope for the Recovery Environment Index, and \(\beta_{\text{PSS}}\) represents the slope for the Substance Problem Scale.

2) **Stepwise regression model**

As a further confirmation of the entry of independent variables into the regression model, a stepwise regression analysis was performed using SPSS. It was
expected that the stepwise regression model would confirm the sensitivity of the linear regression model based on theoretical relationships of predictors on the craving measure.
IV. RESULTS

This chapter presents the demographics of the sample and the full psychometric analysis of the Substance Craving Scale.

A. Demographics

TABLE III describes the demographic characteristics of the sample. Overall, 38.3% of the total sample of 1,044 were male (n=400), and 61.7% were female (n=644). The sample was predominantly African-American (n=934; 89.5%), with smaller percentages of Hispanic (n=64; 6.1%), Caucasian (n=43; 4.1%), and other (n=3; 0.3%). The average age was 48 years old (SD=7.8 years), with the sample of adults ranging from 32 to 77 years of age. The most reported primary substance based on past year substance use disorder symptom counts was cocaine (n=390; 37.4%), followed by opioids (n=372; 35.6%), alcohol (n=180; 17.2%), cannabis (n=50; 4.8%), amphetamines (n=43; 4.1%), and other drugs (n=9; 0.9%).
B. **Analysis Sample**

The analysis presented here includes data of a subset of the larger sample of 1,044 persons measured. In Rasch analysis, the maximum and minimum extreme scores are not analyzed. The results of the following analysis only include those 308 individuals of the larger sample who endorsed at least one item on the measure. For the analysis presented here, 724
individuals did not endorse any of the items, 11 individuals endorsed all of the items on the measure, and there was 1 person that did not respond to any of the items (i.e., missing data).

C. Preliminary Analysis

The first pass of the data included a Rasch analysis of all 12 items on the Substance Craving Scale. Item fit statistics revealed substantial infit and outfit for two of the items. TABLE IV presents the item analysis for the full 12-item measure.

<table>
<thead>
<tr>
<th>Item Descriptors</th>
<th>Measure</th>
<th>Infit</th>
<th>Outfit</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. If I were using alcohol or other drugs, I would feel less nervous.</td>
<td>-.66</td>
<td>1.55</td>
<td>3.59</td>
</tr>
<tr>
<td>2. I have an urge for alcohol or other drugs.</td>
<td>-.46</td>
<td>.83</td>
<td>.88</td>
</tr>
<tr>
<td>3. I crave alcohol or other drugs right now.</td>
<td>-.05</td>
<td>.72</td>
<td>.46</td>
</tr>
<tr>
<td>4. Using alcohol or other drugs would make things seem just perfect.</td>
<td>-.56</td>
<td>.93</td>
<td>.82</td>
</tr>
<tr>
<td>5. I would not be able to control how much alcohol or other drugs I used if I had some.</td>
<td>-2.58</td>
<td>1.51</td>
<td>4.03</td>
</tr>
<tr>
<td>6. Nothing would be better than alcohol or other drugs right now.</td>
<td>.42</td>
<td>.88</td>
<td>.63</td>
</tr>
<tr>
<td>7. My desire for alcohol or other drugs seems overwhelming.</td>
<td>-.30</td>
<td>.70</td>
<td>.54</td>
</tr>
<tr>
<td>8. I want to use alcohol or other drugs so badly that I can't think of anything else.</td>
<td>1.25</td>
<td>.74</td>
<td>.72</td>
</tr>
<tr>
<td>9. Right now, I want to use alcohol or other drugs so bad I can almost taste it.</td>
<td>.70</td>
<td>.59</td>
<td>.34</td>
</tr>
<tr>
<td>10. All I want to do is use alcohol or other drugs.</td>
<td>1.34</td>
<td>.61</td>
<td>.27</td>
</tr>
<tr>
<td>11. I would do almost anything to use alcohol or other drugs.</td>
<td>2.91</td>
<td>.93</td>
<td>.20</td>
</tr>
<tr>
<td>12. I am going to use alcohol or other drugs as soon as I possibly can.</td>
<td>-.02</td>
<td>1.00</td>
<td>1.0</td>
</tr>
</tbody>
</table>
The item “I would not be able to control how much alcohol or other drugs I used if I had some” (Control) was the most misfitting item. Infit was high at 1.51 (ZSTD=9.9), and outfit was also substantial at 4.03 (ZSTD=7.8). The item “If I were using alcohol or other drugs I would feel less nervous” (Nervous) was also misfitting, with an infit of 1.55 (ZSTD=7.5), and outfit of 3.59 (ZSTD=9.9). Because the item Control does not seem to pertain to the theoretical construct of craving, and was highly misfitting, the decision was made to remove it from the scale and re-run the analysis. Although the item Nervous also had a high infit and outfit, the item was retained for theoretical importance. The item also kept the scale’s reliability at a desirable .70 level (below).

The Rasch analysis was re-run with the item Control removed. The remainder of the results presented here are from the analysis of the 11 remaining items.

D. Rasch Analysis: Objectives

1. Objective 1: Reliability

The Winsteps estimation showed a mediocre person-reliability of .70 for the measure. The Cronbach’s alpha for the craving scale was 0.93. This statistic is higher because it estimates extreme scores as measured perfectly, i.e., with no error. The Cronbach’s alpha was highly inflated because so many of the respondents said ‘no’ to all of the items, which are all coded as zero. The more consistent within-subjects are in their responses, the higher Cronbach’s alpha will be. This indicates the upper bound of reliability that includes the ability to distinguish persons with no craving from those with some craving. This distinction may be especially important in outcome measurement to estimate improvement.
a. **Reliability by drug**

To further support reliability conclusions, Cronbach’s alpha and Rasch person reliability by substance were computed (see TABLE V). Two of the drugs, amphetamines and marijuana, did not have enough positive item endorsements to reliably compute Rasch person reliability. For primary alcohol users, Cronbach’s alpha was 0.929, with a Rasch person reliability of 0.68. For primary cocaine users, the Cronbach’s alpha was 0.916, and person reliability was 0.67. For opioids, the Cronbach’s alpha was 0.920, and Rasch person reliability was the highest for all of the drugs, at 0.73. The Cronbach’s alphas for marijuana and amphetamines were 0.926 and 0.766, respectively.

**TABLE V**

<table>
<thead>
<tr>
<th>Primary Drug, Most Severe</th>
<th>N</th>
<th>Cronbach’s Alpha</th>
<th>Rasch Person Reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alcohol</td>
<td>180</td>
<td>0.929</td>
<td>.68</td>
</tr>
<tr>
<td>Cocaine</td>
<td>390</td>
<td>.916</td>
<td>.67</td>
</tr>
<tr>
<td>Opioids</td>
<td>372</td>
<td>.920</td>
<td>.73</td>
</tr>
<tr>
<td>Marijuana</td>
<td>50</td>
<td>.926</td>
<td>--</td>
</tr>
<tr>
<td>Amphetamines</td>
<td>43</td>
<td>.766</td>
<td>--</td>
</tr>
</tbody>
</table>
b. **Item reliability**

Item reliability, which is reflective of reproducibility of the order of items on the scale if given a parallel test, was .98, well above the .80 criterion.

2. **Objective 2: Unidimensionality**

TABLE VI presents the results of the principal components analysis (PCA). The analysis showed that the variance explained by the measurement dimension was 48.5%, greater than the 40% criterion set in the Methods section. The variance explained in the first component of residuals was 18.3%, which was slightly above the criterion of 15% for a rival (or second) factor. Although this did not meet the criterion established in the Methods section, it should be noted that the recommendation by Reckase for a strong measurement dimension is 20% (1979), so this was below that standard. The standardized residual loadings for the items suggested that the rival factor could consist of the following items: *Nervous, Perfect,* and *Urge.* These are the 3 easiest items, which suggests that severity is splitting the items into a two factors, with the easier items falling into what appears to be a rival second factor. The suggestion that the rival factor is severity mitigates the concern about a second dimension since severity on a single dimension can mimic a second dimension.
3. **Objective 3: Item quality**

Since the Substance Craving Scale was found to be sufficiently reliable and unidimensional, the remainder of the validity analyses were performed. The revised 11-item measure revealed only one substantial infit and outfit problem (see TABLE VII); the item *Nervous* still had high infit and outfit. Additionally, the point measure correlations were also reviewed for the items. Under classical test theory conventions, correlations should be 0.3, 0.4 or better. Notably, point measure correlations for the measure were high, ranging from .63-.82.

To further test the validity of the items, the DIF contrasts for the items, by drug, were performed.
# TABLE VII

**SUBSTANCE CRAVING SCALE ITEMS, LABELS, MEASURES, FIT STATISTICS, AND CORRELATIONS**

<table>
<thead>
<tr>
<th>Item Descriptors</th>
<th>Item Label</th>
<th>Measure</th>
<th>Infit</th>
<th>Outfit</th>
<th>Pt. Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. If I were using alcohol or other drugs, I would feel less nervous.</td>
<td>Nervous</td>
<td>-2.18</td>
<td>1.77</td>
<td>6.58</td>
<td>.63</td>
</tr>
<tr>
<td>2. I have an urge for alcohol or other drugs.</td>
<td>Urge</td>
<td>-1.91</td>
<td>.93</td>
<td>1.40</td>
<td>.80</td>
</tr>
<tr>
<td>3. I crave alcohol or other drugs right now.</td>
<td>CraveNow</td>
<td>-.28</td>
<td>.72</td>
<td>.51</td>
<td>.82</td>
</tr>
<tr>
<td>4. Using alcohol or other drugs would make things seem just perfect.</td>
<td>Perfect</td>
<td>-.86</td>
<td>1.04</td>
<td>1.12</td>
<td>.78</td>
</tr>
<tr>
<td>5. I would not be able to control how much alcohol or other drugs I used if I had some.</td>
<td>Control</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Nothing would be better than alcohol or other drugs right now.</td>
<td>NoBetter</td>
<td>.24</td>
<td>.96</td>
<td>.84</td>
<td>.76</td>
</tr>
<tr>
<td>7. My desire for alcohol or other drugs seems overwhelming.</td>
<td>Overwhelming</td>
<td>-.57</td>
<td>.82</td>
<td>.71</td>
<td>.81</td>
</tr>
<tr>
<td>8. I want to use alcohol or other drugs so badly that I can't think of anything else.</td>
<td>CantThink</td>
<td>1.14</td>
<td>.82</td>
<td>.59</td>
<td>.74</td>
</tr>
<tr>
<td>9. Right now, I want to use alcohol or other drugs so bad I can almost taste it.</td>
<td>TasteIt</td>
<td>.54</td>
<td>.61</td>
<td>.41</td>
<td>.80</td>
</tr>
<tr>
<td>10. All I want to do is use alcohol or other drugs.</td>
<td>AllWant</td>
<td>1.23</td>
<td>.31</td>
<td>.76</td>
<td>.76</td>
</tr>
<tr>
<td>11. I would do almost anything to use alcohol or other drugs.</td>
<td>DoAnything</td>
<td>2.90</td>
<td>.92</td>
<td>.25</td>
<td>.60</td>
</tr>
<tr>
<td>12. I am going to use alcohol or other drugs as soon as I possibly can.</td>
<td>UseSoon</td>
<td>-.25</td>
<td>.93</td>
<td>1.40</td>
<td>.76</td>
</tr>
</tbody>
</table>
4. **Objective 4. Invariance of hierarchy (using DIF)**

A cross tabulation of item responses by primary drug was performed to make sure that the number of item endorsements was large enough for the analysis. A DIF analysis requires at least 10 positive item endorsements (a respondent answering ‘yes’). The results of the cross tabulation revealed that for cannabis, amphetamines, and other drugs, there was not a large enough number of endorsements by item for a reliable DIF analysis (see TABLE VIII). Therefore, DIF could only be reliably analyzed for invariance of the hierarchy for alcohol, cocaine, and opioids.

**TABLE VIII**

**NUMBER OF POSITIVE ITEM ENDORSEMENTS BY DRUG**

<table>
<thead>
<tr>
<th></th>
<th>Alcohol</th>
<th>Amphetamines</th>
<th>Cannabis</th>
<th>Cocaine</th>
<th>Opioids</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Nervous</em></td>
<td>41</td>
<td>1</td>
<td>4</td>
<td>60</td>
<td>88</td>
<td>2</td>
</tr>
<tr>
<td><em>Urge</em></td>
<td>35</td>
<td>1</td>
<td>3</td>
<td>58</td>
<td>83</td>
<td>3</td>
</tr>
<tr>
<td><em>CraveNow</em></td>
<td>19</td>
<td>0</td>
<td>3</td>
<td>27</td>
<td>60</td>
<td>1</td>
</tr>
<tr>
<td><em>Perfect</em></td>
<td>23</td>
<td>2</td>
<td>2</td>
<td>38</td>
<td>67</td>
<td>1</td>
</tr>
<tr>
<td><em>NoBetter</em></td>
<td>17</td>
<td>1</td>
<td>2</td>
<td>24</td>
<td>46</td>
<td>1</td>
</tr>
<tr>
<td><em>Overwhelming</em></td>
<td>22</td>
<td>1</td>
<td>2</td>
<td>35</td>
<td>61</td>
<td>0</td>
</tr>
<tr>
<td><em>CanThink</em></td>
<td>10</td>
<td>0</td>
<td>3</td>
<td>20</td>
<td>34</td>
<td>0</td>
</tr>
<tr>
<td><em>TasteIt</em></td>
<td>15</td>
<td>0</td>
<td>2</td>
<td>23</td>
<td>41</td>
<td>1</td>
</tr>
<tr>
<td><em>AllWant</em></td>
<td>12</td>
<td>0</td>
<td>3</td>
<td>18</td>
<td>31</td>
<td>0</td>
</tr>
<tr>
<td><em>DoAnything</em></td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>17</td>
<td>0</td>
</tr>
<tr>
<td><em>UseSoon</em></td>
<td>18</td>
<td>0</td>
<td>4</td>
<td>23</td>
<td>63</td>
<td>0</td>
</tr>
</tbody>
</table>
TABLE IX shows the average item measure by drug, with DIF contrasts for alcohol, cocaine, and opioids. This is the difficulty of each item by drug, with all else held constant. Using the criterion of half a logit, only a few items had any DIF by substance. The item *Nervous* showed DIF for alcohol and opioids, with the average measure for alcohol of -2.68, contrasted with -1.91 for opioids. In general, this item showed larger variability than the other items. It is also important to note that this item was also highly misfitting. The item *CantThink* showed significant DIF for alcohol versus the other two drugs, with alcohol users finding it much more difficult to endorse the item. The average measure for alcohol was 1.83, much higher than the .91 for cocaine and 0.69 for opioids. The other item with substantial DIF was *UseSoon*, with significant DIF for opioid users versus other drugs, with opioid users finding it easier to endorse the item. The average measure for opioids was -0.67, lower than the 0.04 for alcohol and 0.39 for cocaine. *Figure 3* shows a plot of Person DIF measure by item, and from this graph these contrasts can be seen pictorially.
**TABLE IX**

AVERAGE ITEM MEASURE BY DRUG, WITH DIF CONTRASTS

<table>
<thead>
<tr>
<th>Item</th>
<th>Alcohol</th>
<th>Cocaine</th>
<th>Contrast</th>
<th>Alcohol</th>
<th>Opioids</th>
<th>Contrast</th>
<th>Cocaine</th>
<th>Opioids</th>
<th>Contrast</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nervous</td>
<td>-2.68</td>
<td>-2.29</td>
<td>-0.39</td>
<td>-2.68</td>
<td>-1.91</td>
<td><strong>-0.77</strong></td>
<td>-2.29</td>
<td>-1.91</td>
<td>-0.38</td>
</tr>
<tr>
<td>Urge</td>
<td>-2.11</td>
<td>-2.14</td>
<td>0.03</td>
<td>-2.11</td>
<td>-1.66</td>
<td>-0.45</td>
<td>-2.14</td>
<td>-1.66</td>
<td>-0.48</td>
</tr>
<tr>
<td>CraveNow</td>
<td>-0.12</td>
<td>0.06</td>
<td>-0.18</td>
<td>-0.12</td>
<td>-0.51</td>
<td>0.39</td>
<td>0.06</td>
<td>-0.51</td>
<td>0.57</td>
</tr>
<tr>
<td>Perfect</td>
<td>-0.71</td>
<td>-0.92</td>
<td>0.21</td>
<td>-0.71</td>
<td>-0.86</td>
<td>0.15</td>
<td>-0.92</td>
<td>-0.86</td>
<td>-0.06</td>
</tr>
<tr>
<td>NothingBetter</td>
<td>0.24</td>
<td>0.28</td>
<td>-0.04</td>
<td>0.24</td>
<td>0.28</td>
<td>-0.04</td>
<td>0.28</td>
<td>0.28</td>
<td>0</td>
</tr>
<tr>
<td>Overwhelming</td>
<td>-0.57</td>
<td>-0.67</td>
<td>0.1</td>
<td>-0.57</td>
<td>-0.57</td>
<td>0</td>
<td>-0.67</td>
<td>-0.57</td>
<td>-0.1</td>
</tr>
<tr>
<td>CantThink</td>
<td>1.83</td>
<td>0.92</td>
<td><strong>0.91</strong></td>
<td>1.83</td>
<td>1.14</td>
<td><strong>0.69</strong></td>
<td>0.92</td>
<td>1.14</td>
<td>-0.22</td>
</tr>
<tr>
<td>TasteIt</td>
<td>0.61</td>
<td>0.39</td>
<td>0.22</td>
<td>0.61</td>
<td>0.63</td>
<td>-0.02</td>
<td>0.39</td>
<td>0.63</td>
<td>-0.24</td>
</tr>
<tr>
<td>AllWant</td>
<td>1.27</td>
<td>1.23</td>
<td>0.04</td>
<td>1.27</td>
<td>1.31</td>
<td>-0.04</td>
<td>1.23</td>
<td>1.31</td>
<td>-0.08</td>
</tr>
<tr>
<td>DoAnything</td>
<td>2.94</td>
<td>3.15</td>
<td>-0.21</td>
<td>2.94</td>
<td>2.7</td>
<td>0.24</td>
<td>3.15</td>
<td>2.7</td>
<td>0.45</td>
</tr>
<tr>
<td>UseSoon</td>
<td>0.04</td>
<td>0.39</td>
<td>-0.35</td>
<td>0.04</td>
<td>-0.67</td>
<td><strong>0.71</strong></td>
<td>0.39</td>
<td>-0.67</td>
<td><strong>1.06</strong></td>
</tr>
</tbody>
</table>

* = significant DIF contrast at the .6 logit level
Due to the homogeneity of the sample, DIF analysis by race/ethnicity was not performed. The Rasch model showed no significant differential item functioning by gender (results not shown).

5. **Objective 5: Construct and criterion validity**

   a. **Construct validity**

   The items of the Substance Craving Scale formed a hierarchy with low severity items on the bottom, and higher severity items on the top. With one very notable
exception, the hierarchy of item difficulty essentially confirmed the placement of the items on the construct map, which was generated at the earlier stage of measurement development based on empirical theory. The easiest items to endorse were *Nervous* (‘If I were using alcohol or other drugs I would feel less nervous’) and *Urge* (‘I have an urge for alcohol or other drugs’), which were both predicted during measurement development to be low severity items. Similarly, the most difficult items to endorse were *AllWant* (‘All I want to do is use alcohol or other drugs’) and *DoAnything* (‘I would do almost anything to use alcohol or other drugs’), which were also predicted to be high severity items. The item *Use Soon*, which was predicted to be the most severe item, was actually one of the easier items to endorse.

i) **Examination of the item map**

The Wright map (Figure 4) displayed the item hierarchy of the Substance Craving Scale. The difference between the mean of the items (0.00) and the mean of the persons (-1.27) suggested that although the items generally fit the population of interest, this particular sample had low amounts of craving. Again, a floor effect was evident here as over 724 people did not endorse any of the 11 items.

The items have a decent spread across the map, though some gaps do exist, especially at the higher logits. Another noticeable gap exists between the -1 and -2 logit range. There is little overlap of item difficulties, with the exception of and *UseSoon* and *CraveNow*, which were both around -2 logits (-.28 and -.25, respectively).
Figure 4. Wright map of the Substance Craving Scale.
b. **Criterion validity**

TABLE X presents the observed relationships of the Substance Craving Scale with other items and scales available from the Pathways study. All of the correlations were significant at the 0.01 level.

<table>
<thead>
<tr>
<th>Subscale</th>
<th>Correlation</th>
<th>Predicted Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Substance Problem Scale (Past Month)</td>
<td>.616*</td>
<td>High (&gt; .5)</td>
</tr>
<tr>
<td>Frequency of Use</td>
<td>.554*</td>
<td>High (&gt; .5)</td>
</tr>
<tr>
<td>Withdrawal Problems</td>
<td>.484*</td>
<td>High (&gt; .5)</td>
</tr>
<tr>
<td>Recency of Use (any drugs)</td>
<td>.386*</td>
<td>Moderate (&gt; .3)</td>
</tr>
<tr>
<td>Recovery Environment Risk</td>
<td>.277*</td>
<td>Moderate (&gt; .3)</td>
</tr>
<tr>
<td>Perceived Social Support Scale</td>
<td>-0.23*</td>
<td>Low (&lt; .3)</td>
</tr>
</tbody>
</table>

* p < .01
The correlations of the Substance Problem Scale, past month, \(r=0.62\) and Frequency of Use \(r=0.54\) met their high predicted values. The item correlation with Withdrawal Problems \(r=0.48\) was very near the hypothesized value. The Recency of Use \(r=0.386\) correlation met its predicted value. The Recovery Environment Risk \(r=.277\) was very close to the hypothesized value. The Perceived Social Support \(r=-0.23\) to the Substance Craving Scale corresponded exceptionally well.

Overall, the correlations were supportive of craving criterion validity. To further examine the relationship of these variables to the Substance Craving Scale, a series of hierarchical linear regression models were constructed.

c. **Multivariate models**

TABLE XI describes the results of hierarchical linear regression models for the craving predictors, which were based on the theoretical expectations developed from the review of the literature.

Regression coefficients and standard errors are reported in the table. In each of the models, the Substance Problem Scale remained a significant predictor of craving, as did Substance Frequency, and Withdrawal Problems. Substance Recency failed to achieve significance in any of the models. In Model V, Recovery Environment Risk was included and was significant as a predictor. However, when Perceived Social Support was introduced in Model VI, Recovery Environment Risk was no longer significant. Hence, with the exception of the Recovery Environment Risk, all of the variables conformed to the expectations of their effect as craving predictors.
TABLE XI
HIERARCHICAL LINEAR REGRESSION OF CRAVING PREDICTORS

<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
<th>VI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coef</td>
<td>SE</td>
<td>Coef</td>
<td>SE</td>
<td>Coef</td>
<td>SE</td>
</tr>
<tr>
<td>Substance Problem Scale</td>
<td>.489**</td>
<td>.019</td>
<td>.354**</td>
<td>.026</td>
<td>.265**</td>
<td>.029</td>
</tr>
<tr>
<td>Substance Frequency</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4.043*</td>
<td>.524</td>
<td>4.099*</td>
<td>.515</td>
<td>3.809*</td>
<td>.608</td>
</tr>
<tr>
<td>Withdrawal Problems</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.329**</td>
<td>.052</td>
<td>.33**</td>
<td>.052</td>
<td>.343**</td>
<td>.053</td>
</tr>
<tr>
<td>Substance Recency</td>
<td>.026</td>
<td>.029</td>
<td>.016</td>
<td>.029</td>
<td>.015</td>
<td>.029</td>
</tr>
<tr>
<td>Recovery Environment</td>
<td>1.579*</td>
<td>.748</td>
<td>.912</td>
<td>.765</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Risk Index</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived Social Support</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>.405**</td>
<td>.067</td>
<td>.200**</td>
<td>.070</td>
<td>-.009</td>
<td>.077</td>
</tr>
</tbody>
</table>

* p < .05; ** p < .01

TABLE XII presents the results of the stepwise regression procedure. The results supported the selection of variables in the hierarchical regression models. The first variable entered into the models was the Substance Problem Scale, followed by Substance Frequency. Next, Withdrawal Problems was entered into the model. Perceived Social Support was added in the fourth model. Recovery Environment Risk and Substance Recency were excluded from the models.
### TABLE XII

<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coeff</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Substance Problem Scale</td>
<td>.507**</td>
<td>.373**</td>
<td>.279**</td>
<td>.264**</td>
</tr>
<tr>
<td>Substance Frequency</td>
<td>3.916**</td>
<td>4.014**</td>
<td>3.908**</td>
<td></td>
</tr>
<tr>
<td>Withdrawal Problems</td>
<td>.342**</td>
<td>.344**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived Social Support</td>
<td></td>
<td></td>
<td>-.050**</td>
<td>.190**</td>
</tr>
<tr>
<td>Intercept</td>
<td>.402**</td>
<td>.205</td>
<td>-.014</td>
<td>.732</td>
</tr>
</tbody>
</table>

** p < .01
V. CONCLUSIONS

This chapter begins with the purpose and summary of the study, followed by an examination of the psychometric analysis of the Substance Craving Scale based on the objectives of the study. The limitations of the study are then addressed, and the implications for craving measurement in substance craving research and clinical practice are examined.

A. Purpose and Summary of the Study

The desire to drink or use a drug is affected by a variety of influences which can be biological, psychological, and social. At the heart of many treatment approaches is a desire to reduce craving for a substance in an individual. The extent to which reduction of craving leads to a reduction of use is still relatively unknown. The ability to reliably and validly measure craving at an individual level will allow for the demonstration and testing of different relationships between craving and other aspects of substance use, such as frequency of use, substance use severity, and withdrawal. The absence of a measure that is useful across substances has limited the ability of such conclusions to be drawn.

The ability to measure craving reliably, regardless of substance, fills a significant gap. The existing substance-specific measures were useful in small research studies, but do not consistently provide a measure which allows for conclusions that are generalizable to all users. The development of the Substance Craving Scale was based on a extensive review of the literature, and considered that these previously developed tools had items that had redundancies with each other, and were not based on any drug-specific framework that would prohibit translation across substances. The central hypothesis of this study was that the developed scale
could reliably measure craving across substances, without major differences in item or scale response.

**B. Psychometric Analysis – Reliability and Validity of the Substance Craving Scale**

1. **Reliability**

   The first research objective was to determine if the Substance Craving Scale was reliable in the measurement of craving. The Cronbach’s alpha, an indication of internal consistency of the measure’s items was strong, at 0.92, which did meet the expected criteria. This is important, since in outcome measurement it is important to be able to detect improvement, e.g. presence of craving to absence of craving. However, Rasch person/item reliability reflects the reliability of the placement of both on the measurement scale, excluding those with measures reflecting extreme scores (zero or perfect). Here, the estimation of the data showed a reliability of 0.70, which failed to meet the criteria for a strong measure. A reliability of 0.70 is very useful for the context of research, but the scale needs to improved to distinguish levels of craving in the clinical context.

   Notably, the point measure correlations were high, ranging from .63-.82. This indicates that the items are strong, but the reliability of the scale can be improved by adding additional items, or increasing the number of response categories. The addition of response categories, beyond yes/no, might provide a more reliable measure. Indeed, all of the substance-specific measures these items were taken from used a 7-point response scale. While dichotomous items are less burdensome for a respondent to answer, adding one or two more response options could make a big difference in improving reliability. One option could be a 3-point scale with responses of “Yes, Somewhat, No”, so that respondents could indicate a more nuanced perceived level of craving, with “Yes” being a very pervasive craving experience, and
the “Somewhat” category allowing for a more moderate experience of craving. This and other forms of response categories could be tested to determine which is the most appropriate for the construct and comfortable for respondents.

Another suggestion would be to generate additional items and add these to the scale to increase reliability. Suggestions for this process will be discussed later, since the addition of good items is always useful in improving a measure’s reliability. It should be cautioned, however, that increasing items on the Substance Craving Measure could increase reactivity to the test. That is, answering many items about craving could increase craving in the individual. For the assessment of craving, an optimal balance of reliability and respondent burden is highly desirable.

2. **Dimensionality**

The second research objective was to determine if the Substance Craving Scale was unidimensional. That is, measuring a single construct. Although the measurement dimension was strong, there was an indication of a competing factor. The three items that were found on a potential rival factor were *Nervous, Urge, and Perfect*, the 3 easiest items to endorse on the measure. The data here suggest that item severity is splitting the items on the measure, with the 3 easy items masking as a separate factor. There is not enough theoretical reason to believe that these items are tapping into a construct other than craving. These findings indicate that the items of the Substance Craving Scale comprise a unidimensional measure.

3. **Validity of the items**

The third research objective was to verify that the items of the scale were valid, and met the fit criteria, prior to completing the full Rasch analysis.
The item that was removed prior to the complete Rasch analysis was NoControl: “I would not be able to control how much alcohol or other drugs I used if I had some.” The item was removed due to its substantial infit and outfit. Additionally, the DIF analysis of the full scale (results not shown) revealed significant DIF for cocaine versus the other two drugs, by more than one logit. Qualitatively, the wording for this item suggests that the construct being tapped is control of substance use, and not craving. Indeed, in the original articles from which this item was drawn, this item consistently failed to load on a common factor during factor analysis procedures (Heishman, Singleton, & Liguori, 2001; James, Davies, & Willner, 2004; Love, James, & Willner, 1998). These findings suggest that it does not fit the criteria for an item about craving for substances. The decision was made to remove the item from further analyses, and these findings suggest that this item not be used in the future.

The analysis of the revised 11-item measure presented here revealed only one item with substantial infit and outfit problems. The item Nervous had a high infit and a very high outfit, but it was retained since conceptually, this item seems to be related to the construct of craving. Furthermore, when the item is removed from the analysis, reliability falls to 0.69. As such, this item might prove to be more stable in another research sample. Further inspection of individual person response patterns (data not shown) revealed that the most unexpected person responses for this item were found in persons who were probabilistically expected to endorse this easy item based on their overall Rasch measure, but did not endorse the item. The wording of this item might assume that the respondent is feeling nervous at the time of its administration. Qualitative probing on this item may yield additional information about the item and how it is being interpreted by respondents.
Another interesting point to make regarding the items is that the NLAES item *CantThink* was the second most difficult item to endorse. This was the item that Kreuger and colleagues (2011) examined for its addition in the DSM-V criteria for abuse and dependency criteria. These findings suggest that a less severe item be chosen for such criteria.

Overall, the psychometric analyses presented here give insight to the items chosen for inclusion into the scale. The items functioned well, but there are some that should further examined for their appropriateness, or removed altogether. Clearly, the most problematic item was *Nervous* ”If I were using alcohol or other drugs I would feel less nervous.” This item, in addition to having high infit and outfit, also contained substantial DIF. It could be that the high variability of this item is because of the word ‘nervous’. This low severity item could easily be replaced by another item. The other items that should be re-tested are *UseSoon*, and *CantThink*.

4. **Invariance of hierarchy by drug**

The fourth research objective, notably the most important of this project, was to determine if the items on the scale performed similarly regardless of the substance of choice.

Overall, the measure showed little variance of hierarchy by the 3 drugs that were examined. More positive item responses by different classes of drugs will be needed to determine if the item is useful across additional drugs. The results presented here should still be interpreted with caution. Cross tabulation revealed that positive item endorsements for alcohol were still low (ranging from 41 ‘yes’ for the easiest item *Nervous*, and 7 ‘yes’ for the most difficult item *DoAnything*).

Items with DIF present were *Nervous, CraveNow, CantThink,* and *UseSoon.*

Given the highly misfitting pattern of the item *Nervous,* it is not surprising that this item had significant DIF. The rationale for retaining this item in the final analysis was primarily to retain
an item at the easier end of the scale that seemed substantively related to craving. Again, this item might prove to be more stable when administered in a different sample. In general, the item Nervous showed larger variability than the other items. Again, the wording for this item presumes that the respondent is feeling nervous, and believes that using a substance will alleviate this nervousness.

The results of the DIF contrasts, overall, show that there was some differential item response present. One possibility is that rather than an indication of strong item bias, differences in items may reveal differences in the craving or addiction experience by drug. For research, this information can help classify the craving process by substances. The most notable caveat for interpreting these results is the small number of item endorsements by drug. Positive item endorsements, by drug, ranged from 8-88. If the results of this study were replicated with a much larger sample, specifically in a sample likely to endorse higher levels of craving, like those people entering or receiving substance use treatment, a much higher rate of positive item endorsements by drug would be expected. Increasing the type and size of the sample would have yielded a more reliable calculation of DIF. Again, DIF contrasts were not even possible to calculate for cannabis, amphetamines, and other drugs, due to the low number of positive item endorsements for those individuals. To improve the testing of the invariance hierarchy of the Substance Craving Scale, the measure should be readministered to a larger and more diverse sample.

5. **Construct and criterion validity of the measure**

The fifth objective was to determine if the scale was sufficiently valid. The findings were very supportive of construct validity and criterion validity.
a. **Construct validity**

Overall, the results of the analysis supported construct validity. In a Rasch analysis, the item hierarchy can be used as evidence with low severity items on the bottom, and higher severity items on top. Rasch models provide a severity hierarchy using the proportion/probability of item endorsement as a severity estimator. It is seen in the hierarchy of the Substance Craving Scale, where people may endorse an *Urge* but are much less likely to *DoAnything* to use a substance. Wilson (2005) states that the empirical hierarchy should both confirm and inform the construct map. In this case, the item map has provided information about the construct of craving. Many of the qualitatively predicted placements of the items were confirmed with the quantitative results of the analysis. One item that was surprisingly easy for respondents to endorse was *UseSoon* (‘I am going to use alcohol or other drugs as soon as I possibly can’). This was predicted to be the most severe item. Further probing of this item might reveal why it was so easy for respondents to endorse it, and how it is being interpreted.

Further examination of the item map revealed the mean of the persons (-1.27) was low; this particular sample had a low level of craving. This is not surprising since many of the individuals in this sample were in long-term recovery (abstinent from drugs or alcohol for a long period of time). It would be hypothesized that the level of craving of a population of respondents in substance treatment would be much higher.

b. **Criterion validity**

All correlations of the Substance Craving Scale corresponded with the expectations and were thus highly supportive of criterion validity. The hypothesized correlations were based on empirical evidence, to the extent it was known that these constructs were related.
In addition to being supportive of validity, these correlations also provided some insight on the nature of craving, as to what factors increase or decrease the presence of craving.

Certainly, it is not surprising that the Substance Problem Scale had the highest correlation with craving. The other two hypothesized correlates to be highly associated with craving were Frequency of Use and Withdrawal Problems. The Recovery Environment Risk Scale failed to meet the predicted correlation but closely approached the value. This finding might suggest that certain aspects of a risky recovery environment may increase the chances of cues that trigger craving. Recency of use for any drugs also met the hypothesized moderate value. Interestingly, the Perceived Social Support Scale was indeed lowly and negatively correlated with craving, which was predicted given the evidence which suggests that social support is protective for craving substances (Schroeder, et al., 2001). This may further confirm the protective effects of positive social support for persons in substance recovery.

The regression models further explored the relationships of the hypothesized correlates with the Substance Craving Scale. Primarily, the results show that craving is mostly driven by diagnostic severity, and frequency of use. Above and beyond that, withdrawal has an effect on craving. The results of the regression models are the first known series of models that use independent variables to predict substance craving based on substance-related correlates. Their conformity to the theoretical model, in addition to supporting criterion validity, offers some additional perspective on what variables affect craving, and also indicate the magnitude of these effects. Simple regression models such as the one presented here demonstrate how having a measure of craving can provide insight on substance use severity. Such knowledge should prove to be useful as research on this construct progresses via additional tests of the effects of different variables on substance craving.
C. **Limitations of the Study**

Although the research has reached its aims, there were some unavoidable limitations. These limitations can be reduced or eliminated in future research surrounding craving.

1. **Homogeneity of the sample**

   This sample was highly homogeneous in several ways. Primarily, they were mostly ‘hard’ drug users, with cocaine and opioid users comprising 73% of the sample. Additionally, approximately 87% of the sample was African-American. Finally, this sample consisted of mostly older adults, with mean age of 48. This limits the generalizability of findings to this group. Clearly, replication on more diverse samples, especially those in treatment, is needed to confirm or deny the reliability and validity of the Substance Craving Scale. More specifically, contrasts in item response by all drugs, including marijuana and amphetamines, are needed to see if the items are generalizable across substances.

2. **Items came from the literature**

   The scale was comprised of items from previously validated measures found in the literature, and no original item development was conducted. Several qualitative methods exist to generating items in addition to synthesizing existing literature. These include participant observation, concept mapping, focus groups, cognitive interviews, and exit surveys. None of these were used in the construction of the measure. It would be suggested that an item bank be developed using some of these methods in order to generate other items useful in measuring craving, especially at the more severe end of the item map.

   These items could be developed by testing more items from available scales, or generating additional items through a technique such as concept mapping. In addition to the
qualitative methodologies available to generate additional items, there are still many items in the existing literature that could be tested for inclusion into the Substance Craving Scale. A few example items from the literature to test could be: “I am thinking of ways to get alcohol or other drugs,” or “If I had the chance to use alcohol or other drugs now I think I would.” These items would likely cover some of the gaps at the more severe end of the scale.

To generate more items, some more qualitative techniques should be employed, including concept mapping. A concept mapping session with topical experts, along with active and recovering individuals, might generate some additional items that are not found in the literature. As the drive exists to more accurately define and conceptualize craving, new insight might provide a better understanding of how clients and clinicians deal with the craving experience.

D. Recommendations for Future Research

1. Improving the scale and its items

The Substance Craving Scale has potential clinical utility if the reliability of the measure can be improved. If craving severity can be determined, appropriate treatment protocols aimed at reducing craving can be tailored. With knowledge of the factors that can reduce craving, such as positive social support, treatment plans can include the incorporation of these approaches to help reduce craving triggers.

To improve the measure, the techniques discussed in the last section could be used to generate more items. In addition to increasing the number of items, the number of response categories could be increased to more than a simple yes/no dichotomous response. Having more response categories generally increases reliability, but only if the increase is not
arbitrary (Linacre, 1995; Wright & Linacre, 1992). These response options could be qualitatively explored and quantitatively tested.

Certainly, future research should examine whether or not craving can predict relapse. Craving alone and in combination with other factors (e.g., stress, availability) is thought to be one of the leading risk factors for predicting relapse, so a craving measure would both a useful clinical measure and outcome for recovery management. A Substance Craving Scale with improved reliability would add to its utility in both research and clinical settings.

2. **Establish long form, short form, and single item scales**

   a. **Long form**

   As was discussed in the literature review, there are varying theories as to the appropriate length of craving measures for research and clinical purposes. To determine the nuanced experiences of self-reported craving, a scale with more items is probably most appropriate. The Substance Craving Scale should be repeated with different populations to get a better idea of the craving experience, which will allow for advanced research into predictive utility, establishing severity cut-points, and will help to gain a better understanding of craving overall. Using the full 11 items or a scale with even more than 11 items, would likely provide more information on craving for these research purposes.

   Also, the current measures of craving have been validated and are suitably reliable, but have not been explored in depth for associations with other factors related to abuse and dependency. The availability of a full-length, standardized measure for use across substances should allow for these relationships to be carefully hypothesized and tested.
b. **Short form**

It is clear that the Substance Craving Scale should be analyzed for its predictive utility. That is, what is the association of craving with relapse? That question has been difficult to answer without a validated measure of craving, across substances. However, it may not be necessary to administer the full measure to gain information on craving’s association with correlates. A shorter form with few items, while being less reliable than a longer measure, might be the most useful in making predictions for other things, such as relapse. An important benefit of a shorter scale is reducing potential reactivity to the scale.

It would be useful to examine what items might be chosen and tested for a short form of the scale. It is important to try to select items that provide the most information (unique variance) about the construct. In this case, items that fall on the same place on the scale could be replaced by choosing just one of the items. The following 5 items (ranging from low to high severity) could be tested as a short-form version of the measure: *Urge, Perfect, CraveNow, TasteIt* and *AllWant*. These items cover the spread of the logit scale at the low, medium, and high range. This short-form of the scale could be tested, and it could be determined through appropriate analyses that this short form of the scale can serve to reliably make predictions, like whether a higher score on the measure leads to subsequent relapse.

c. **Single item**

There are instances when a single item is used, often in research settings where it is important to reliably and rapidly assess a construct. In this case, it would be appropriate to select an item that is in the moderate range of the scale, and generally captures the construct. One single item that may perform well is *CraveNow*: “I crave alcohol or other drugs
right now.” This item should function well to assess an individual’s craving, without being too specific.

3. **Study potential reactivity to the scale**

Knowing if individuals are reacting to the Substance Craving Scale, where craving is increased by exposure to the items, is crucial. Reactivity is a concern for any construct in which the asking of questions may increase the level of the trait. In this case, just asking questions about craving for a substance may increase an individual’s level of self-reported craving. If craving is increased by administering these items, it would signal that the scale should be shortened, even at the expense of its reliability. As an example, in the Pathways study where the current data came from, a single-item measure for craving was asked prior to the full scale. An individual who was ‘reacting’ to the items on the craving scale would answer ‘yes’ to items on the scale when their answer to the single preceding craving question was a ‘no’. To establish a low level of reactivity to the scale, it would have to be shown that only a small percentage of individuals, or no individuals, were demonstrating reactivity to the scale. If the scale is indeed causing reactivity, there would be strong evidence to support reducing the amount of items on the Substance Craving Scale.

E. **Significance and Conclusion**

A reliable and valid measure of craving has many applications. As an outcome measure, craving will allow practitioners to track progress of their clients’ treatment goals. From a research perspective, the measure will allow for investigations into craving, its correlates, predictors, and antecedents. One very important relationship to establish will be that of its
predictive utility. If craving can adequately predict relapse, more interventions could target craving reduction in the quest to reduce or eliminate substance use.

The analysis presented here has showed that there is promise in moving forward with a general, rather than substance specific, measure of craving. Although items may differ slightly in response by drug, such difference can be acknowledged and accounted for. Researchers can accomplish this by advanced quantitative models, like IRT, which can account for differences in item response by different classes of individuals. If these individuals respond in ways that are systematically different, this knowledge can be utilized moving forward in addiction research. Having a clear and unique understanding of craving and ability consistently evaluate a level of craving using a standardized approach will allow practitioners to better understand their clients’ challenges in recovery.


VITA

NAME: Jessica Marie Mazza


M.S.P.H., Public Health, University of South Florida, Tampa, Florida, 2008

Ph.D., Public Health, University of Illinois at Chicago, Chicago, Illinois, 2013

PUBLICATIONS:


