Summer 8-7-2021

Psychometric Synthesis of the Drug Abuse Screening Test 20-Item Version (DAST-20)

Erin Johnson

PSYCHOMETRIC SYNTHESIS OF THE DRUG ABUSE SCREENING TEST 20-ITEM VERSION (DAST-20)

A Dissertation

Submitted to the School of Education

Duquesne University

In partial fulfillment of the requirements for

the degree of Doctor of Philosophy

By

Erin K. Johnson

August, 2021

PSYCHOMETRIC SYNTHESIS OF THE DRUG ABUSE SCREENING TEST 20-ITEM VERSION (DAST-20)

By

Erin K. Johnson

Approved June 25, 2021

David Delmonico, Ph.D. Professor of Counselor Education Department of Counseling, Psychology, and Special Education Duquesne University (Committee Chair)	Bradley T. Erford, Ph.D., Professor Human Development Counseling Peabody College at Vanderbilt University (Committee Member)
William J. Casile, Ph.D. Associate Professor of Counselor Education Department of Counseling, Psychology, and Special Education Duquesne University	
Gretchen Givens Generett, Ph.D. Dean, School of Education Professor of Education Duquesne University	Debra Hyatt-Burkhart, Ph.D. Chair, Department of Counseling, Psychology, and Special Education Associate Professor of Counselor Education Duquesne University

ABSTRACT

PSYCHOMETRIC SYNTHESIS OF THE DRUG ABUSE SCREENING TEST 20-ITEM VERSION (DAST-20)

By

Erin K. Johnson

August 2021

Dissertation supervised by Dr. David Delmonico

Numerous research articles have reported differing data on the psychometric properties of the 20-item version of the Drug Abuse Screening Test (DAST-20, Skinner, 1982a). Aggregating this diverse information can lead to a better understanding of how to use and interpret the instrument with clients and research participants. In this psychometric synthesis, evidence of reliability and validity of the DAST-20 scores was aggregated in order to provide a more comprehensive summary of the psychometric properties of the instrument to better inform counseling professionals when using the DAST-20. Overall, the available evidence indicates that the DAST-20 produces reliable and valid scores when screening for drug abuse consequences. However, data was limited and future research is needed to further assess the psychometric properties of the instrument, including internal consistency, test-retest reliability, convergent validity,

diagnostic validity, and structural validity. Implications for professional counselors and counselor research are discussed.

DEDICATION

This dissertation is dedicated to my husband, Chris, who is my greatest champion. Thank you for being nothing, but supportive, encouraging, and loving.

ACKNOWLEDGEMENT

I cannot truly put into words how much I appreciate all the support and encouragement I have received during this process, but this is my attempt.

Thank you to my committee members, Dr. Delmonico, Dr. Erford, and Dr. Casile, for your guidance, patience, support, and knowledge through the dissertation process. Dr. Delmonico, thank you especially for the many conversations as I worked through my anxiety. Dr. Erford, thank you not only for the support during this process, but for being a mentor over the last 10 years. Dr. Casile, thank you for your guidance not only as a committee member, but during my experience as a research assistant.

To my family, you have always supported me and encouraged me to follow my dreams. Thank you for always being there and for always believing in me.

To the most amazing friends, Sally, Nellie, Katie, and Kourtney, thank you for being there to support me through this process. Thanks for the long phone calls, the many coffee runs, and for knowing when I needed a work accountability partner or some time away from my dissertation. I do not know where I would be without each of you!

To my husband, Chris, you have been my rock through this journey. Thank you for always believing in me, even when I did not. I will always be grateful for your endless support, love, and encouragement. Thank you for being you.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

CHAPTER I: INTRODUCTION

Overview

According to the National Institute on Drug Abuse (NIDA, 2021b), a rise in the use of illicit drugs has been observed in the United States. Based on data collected in 2018, the NIDA reported that 11.7% of Americans over the age of 12 years engaged in illicit drug use within the previous 30 days. In the past two decades, an opioid crisis has been reported, with a significant rise in the number of deaths related to opioid overdose (U.S. Department of Health and Human Services [HHS], 2018). In 2017, there were 70,237 fatal drug overdoses in the United States (Scholl et al., 2019). Opioids contributed to 47,600 of those deaths. Further, the HHS reported heroin overdoses rose by 400 percent in the previous ten years and synthetic opioid overdose contributed to a 300 percent increase in deaths in the previous seven years. In a study examining the prevalence of fatal drug overdoses over a two-year period (2017-2018), a decrease in overall fatal opioid overdoses was reported across the United States, which coincides with a decrease in the prescription of opioids (Wilson et al., 2020). Although there were fewer overall fatal opioid overdoses, synthetic opioid-related deaths increased by almost 10 percent over that two-year span.

Costs associated with drug use are assessed in three categories: crime, health care, and productivity (NDIC, 2011). Illicit drug use costs more than 600 billion dollars annually in the United States (NIDA, 2021a), which included a projected 120 billion dollars in lost economic productivity. In 2013, prescription opioid abuse alone cost the United States a projected $78.5 billion, including a projected 26 billion dollars in health care.

Given the prevalence of drug abuse, it is increasingly important to determine effective screening practices, treatment procedures, and best practices among various populations to identify drug abuse related consequences (Kiluk et al., 2018; USDHHS, 2018). Screening instruments are among the first steps in assessing for drug abuse and are vital tools used by professionals to determine whether individuals abuse drugs (Rasmussen, 2000). Since screening instruments are relied upon by evaluators to help determine if problems related to an individual's drug use are present, it is essential the instruments yield valid and reliable scores. NIDA (2021b) indicated that out of the 22.7 million individuals in the United States who required treatment for drug abuse in 2013, only 2.5 million received any specialized treatment. While many factors contributed to a lack of treatment, use of robust screening instruments could be an important initial step in identifying individuals for treatment. In addition to identifying individuals in need of treatment, screening instruments are crucial in gathering information and identifying problems related to drug use to ensure the most effective treatment is provided (Rasmussen, 2000; Skinner, 1982a). One such screening instrument that identifies problems related to drug abuse is the Drug Abuse Screening Test (DAST; Skinner, 1982b).

The DAST (Skinner, 1982b) was developed to screen for the presence of consequences associated with drug abuse; currently there are four forms of the DAST with varying numbers of items (i.e., DAST-28, DAST-20, DAST-10, DAST-A). The DAST items ask about an individual's drug use and any psychosocial problems that resulted from drug abuse. Common psychosocial issues resulting from drug abuse include psychological well-being, relationship strains, professional consequences, financial

hardships, and legal trouble (Maly, 1993). Identifying the presence of these issues and examining the extent to which they affect an individual's life is essential in the identification and treatment of drug abuse problems (Maly, 1993; Rasmussen, 2000). The data gathered about the number and severity of problems stemming from an individual's drug abuse can assist clinicians and researchers in developing tailored treatment plans, as well as in assessing how these issues will potentially affect an individual's recovery process (Rasmussen). Information should be gathered to help determine whether the psychosocial issue being assessed was present prior to an individual's problem drug use and how the issue may impact use of drugs (Kiluk et al., 2018; Macleod, 2010; Rasmussen).

The DAST was created in response to limited instruments available to screen for consequences of drug use and abuse (Skinner, 1982b). Much like the Michigan Alcoholism Screening Test (MAST; Selzer, 1968, 1971), upon which the DAST was based, the DAST was established due to a lack of psychometrically sound instruments to screen for addictions, as many of the resources available in the early 1980's were nonstandardized surveys with little psychometric support.

Statement of the Problem

This study examined psychometric data published by numerous researchers on the DAST 20-item version (DAST-20; Skinner, 1982a). The various populations and settings used in these studies have contributed to varying reports of psychometric values of the DAST-20 items. These differences in findings can lead to a misunderstanding of how to interpret results of this instrument, which can impact the effectiveness of the DAST-20 in clinical and research settings (Erford, 2021). Even with similar samples, factors such as

researcher, settings, and procedures can lead to varying results. This can be confusing for clinicians and researchers looking at multiple studies with samples that closely match their client. Lipsey and Wilson (2001) described aggregating the data from multiple similar studies as "smoothing the resulting picture into a composite, much as a magazine picture looks more crisp and coherent at arms length than when the pixels are examined through a magnifying glass" (p. 167). By combining the results from all the relevant studies, it provides a more robust examination of the results and allows for clinicians and researchers to have a clearer picture of the data.

Researchers and clinicians have the responsibility to provide services that are supported by empirical evidence and ethical guidelines (American Counseling Association [ACA], 2014). Numerous ethical concerns need to be considered by researchers and practitioners when choosing and using screening instruments with individuals. ACA outlined ethical guidelines pertaining to the psychometric properties of an instrument, the population for which the instrument is being used, and necessity of providing empirical data that support the use of an instrument. The specific relevant clauses follow along with commentary.

> *E.6. Instrument Selection E.6.a. Appropriateness of Instruments*
>
> *Counselors carefully consider the validity, reliability, psychometric limitations, and appropriateness of instruments when selecting assessments and, when possible, use multiple forms of assessment, data, and/or instruments in forming conclusions, diagnoses, or recommendations* (ACA, 2014, p. 11).

Clause E.6.a (ACA, 2014) specifically considers the importance of examining the psychometric properties of an instrument before using it with clients. Empirical data on

an instrument should also be examined to assess the appropriate use with various populations and settings and their implications for interpreting results. This can be difficult when multiple sources are reporting differing reliability and validity values for scores from a particular instrument (Erford et al., 2015). Since differences in study conditions and samples can lead to differing outcomes, a psychometric synthesis can provide an overall summary of the psychometric properties on an instrument, allowing professionals in the field to have a better understanding of appropriate use of the instrument (Cook et al., 1997; Erford et al., 2015; Lipsey & Wilson, 2001).

E.8. Multicultural Issues/ Diversity in Assessment

> *Counselors select and use with caution assessment techniques normed on populations other than that of the client. Counselors recognize the effects of age, color, culture, disability, ethnic group, gender, race, language preference, religion, spirituality, sexual orientation, and socioeconomic status on test administration and interpretation, and they place test results in proper perspective with other relevant factors* (ACA, 2014, p. 11).

Code E.8 (ACA, 2014) emphasizes the importance of considering many multicultural facets when using a particular instrument and interpreting results. Varying study samples can influence findings on psychometric properties of an instrument and can make it difficult for professionals to know how to apply a client's results (Cook et al., 1997; Erford et al., 2015; Lipsey & Wilson, 2001).

E.9.b. Instruments With Insufficient Empirical Data

> *Counselors exercise caution when interpreting the results of instruments not having sufficient empirical data to support respondent results. The specific*

> *purposes for the use of such instruments are stated explicitly to the examinee. Counselors qualify any conclusions, diagnoses, or recommendations made that are based on assessments or instruments with questionable validity or reliability* (ACA, 2014, p. 12).

A psychometric synthesis of the reported data can help summarize differences among research findings due to sampling differences, allowing professionals to be more accurate in determining when to use an instrument and how to interpret results (Cook et al., 1997; Erford et al., 2015; Lipsey & Wilson, 2001).

This study will use psychometric synthesis procedures to extract all relevant data and analyze the aggregated data from all included studies to provide a more comprehensive interpretation of the psychometric properties of the DAST-20 (Erford et al., 2015). This process allows a more inclusive and robust examination of the data to inform clinicians and researchers using the DAST-20.

Purpose of the Study

Due to the prevalence of drug abuse and the importance of using screening instruments that yield valid and reliable scores to assess individuals' drug abuse and to identify the need for treatment, further examination of instruments is essential. The purpose of this study is to provide a comprehensive analysis of the psychometric properties of DAST-20 by conducting a psychometric synthesis using the findings from the identified relevant studies. The results will provide insight into the validity and reliability of the DAST-20 scores and inform clinicians and researchers who use it for screening and research purposes by providing a more comprehensive summary of DAST-

20 psychometric properties (e.g., internal consistency, test-retest reliability, convergent validity, structural validity, diagnostic validity).

A psychometric synthesis shares many characteristics with a meta-analysis, however, a meta-analysis is a specific order of procedures and statistical methods that are used to examine the outcome of a treatment program or intervention, whereas a psychometric synthesis is used to examine the psychometric properties (e.g., reliability estimates, validity estimates, descriptive statistics) of an instrument (B. T. Erford, personal communication, June 25, 2021). Both provide an empirical process to examine the data from various studies on a shared topic of interest to produce an aggregated analysis of outcomes (Erford et al., 2015). As such, the results can provide a more comprehensive summary of the data than looking at each of the studies individually across various populations and study characteristics to allow professionals to make more informed decisions about using and interpreting the DAST-20. By examining the analysis of the aggregated data, researchers can identify limitations of particular studies and differences in conditions that may lead to varying outcomes (Cook et al., 1997; Lipsey & Wilson, 2001). Therefore, psychometric synthesis studies can provide clarity on why these differences might occur and what their implications might mean in counseling practice, which allow for a more unbiased view of outcomes and a more in-depth examination of various conditions that affect results. Further, the procedure of examining the aggregated data from the included studies results in a reduction in sampling error, which allows a more robust report of findings (Lipsey & Wilson, 2001). In this dissertation, an aggregated analysis of the psychometric properties of the DAST-20 will be reported and implications discussed.

There are currently four versions of the Drug Abuse Screening Test occurring in the literature, including the original DAST (28 items), the DAST-20 (20 items), the DAST-10 (10 items), and the DAST-A (a version created for use with adolescents). The 20-item and the 10-item versions of the DAST were created after conducting an item-analysis of the original 28-item DAST (Skinner, 1982b, 1984). A very high correlation was found between the DAST-20 and the original 28-item version of the DAST (r = .99) (Skinner, 1982b). This study specifically examined the psychometric properties of the DAST-20.

In the current study, examination of the DAST and DAST-A versions of the instrument were eliminated because the abbreviated versions were highly correlated with the original 28-item version and the DAST-A was specific to only the adolescent population. Examination of the DAST-20 was chosen over the DAST-10 based on Skinner's (1984) recommendations for the use of each version. The DAST-20 produces a more comprehensive evaluation of the individual's consequences that result from drug abuse and has been recommended for use at "specialized assessment centres and for use as a research evaluation tool," whereas the DAST-10 has been recommended for "screening and case finding purposes" (Skinner, 1984, p. 30). Therefore, the findings of this study may be more useful to individuals using the DAST in both clinical and research settings. Further studies should be completed on the other forms of the DAST instrument in the future.

Research Questions

Three research questions form the basis of this study:

1. What are the aggregated psychometric properties of the DAST-20 across published studies (i.e., internal consistency, test-retest reliability, diagnostic validity, internal structural validity, external convergent validity)?
2. What are the mean scores and standard deviations of the DAST-20 in nonclinical samples across published studies?
3. Are there significant differences in the psychometric properties of the DAST-20 among various sample characteristics (e.g., gender)?

Theoretical Orientation to the Study

As previously stated, the DAST (Skinner, 1982b) measures the presence of consequences related to an individual's drug use. Impaired functioning theory posits that drug abuse has a negative impact on an individual's functioning, which leads to issues in various aspects of life (Newcomb & Bentler, 1988). Further, impaired functioning theory provides a framework for categorizing consequences of drug abuse into physical, emotional, and psychological domains. Impairment in any of these domains can lead to problems coping with life's stressors and completing a number of daily tasks. This theory supports the use of the DAST to measure consequences related to an individual's drug use and the items that are included on the DAST. The 20 items on the DAST-20 inquire about one's drug use and related consequences concerning physical health, relationships, and functioning at work (Skinner, 1982b). Impaired functioning theory will be discussed in further detail in Chapter 2.

Significance of the Study

This study provides an integrated analysis of the psychometric properties of the 20-item version of the DAST, including internal consistency and test-retest reliability,

diagnostic validity, internal structural validity, external convergent validity, and nonclinical sample means and standard deviations. The analysis of these psychometric properties will provide professionals with a more comprehensive understanding of the instrument, which can be used to guide their work with clients (Erford et al., 2015). More specifically, professionals will have better insight when using this instrument with their clients, of implications when interpreting results, and applying this information to create treatment plans with clients. At a broader level, integrative syntheses can have an impact on policy and best practices due to a more robust examination of results (Lipsey & Wilson, 2001).

Summary of Methodology

In the current study, a psychometric synthesis process will be conducted to examine psychometric findings on the 20-item version of the Drug Abuse Screening Test (DAST-20; Skinner, 1982a) since 1982. A psychometric synthesis was chosen as the methodology for this study because it allows for the examination of all available data on the psychometric properties of an instrument, creating an aggregated summary of the reliability and validity (Erford et al., 2015). The process of a psychometric synthesis involves collecting all possible data from studies examining a common research question that meet the established criteria included in the study and using particular statistical procedures to aggregate and summarize these findings. The procedure used in this study was outlined by Erford et al. (2015) in a synthesis that examined the psychometric properties of the Beck Depression Inventory – Second Edition (BDI-II). This procedure includes establishing criteria for article inclusion, searching for relevant articles, examining articles for criteria and completing the selection process, data extraction, and

data analysis. A more comprehensive examination of the psychometric synthesis process and the methodology used in this study will be provided in Chapter 3.

Delimitations of the Study

One criterion required for prospective studies to be accepted in the psychometric synthesis was the linguistic version of the DAST-20 used in the study and the language in which the study was published. This criterion helped to eliminate issues with translation and adaptation, but it is also important to note that there are other studies published with different linguistic versions of the DAST-20 and potentially relevant studies published in other languages that could not be included in the analysis (Lipsey & Wilson, 2001). Further studies are needed to examine these potential studies and the instruments' psychometric properties across cultures.

This study also only examines the DAST-20. Other versions of the DAST (i.e., DAST, DAST-10, DAST-A) were not included in the study. Therefore, the results of this study can only be applied to guide one's use of the English version of the DAST-20. Similar synthesis studies are needed to examine the psychometric properties of the other versions of the instrument.

Further, the data used in this psychometric synthesis are limited by how detailed the researchers of the included studies were in reporting their findings. The extraction of data was completed by reviewing the full text version of each study and, therefore, only the data that was provided in the articles could be used in this analysis of the psychometric properties of the DAST-20.

It is important to note the general limitations of psychometric syntheses and the criticisms that have been made about the process and reporting of results (Erford et al.,

2015; Lipsey & Wilson, 2001), which have certain potential limitations, regardless of the topic being studied. For example, the specific factors of the studies, such as methodology, population, and analysis of data may have implications for the scope of the analysis. Significant factors, such as study limitations, cultural implications, theoretical framework, and the power and quality of the study may be overlooked in the overall analysis.

Key Terms

Drug abuse – Many terms were used in the literature to refer to drug abuse (i.e., drug use, drug use disorder, substance abuse, substance use, etc.). For the purposes of this study, the term "drug abuse" will be used to refer to Skinner's (1982a) definition of drug abuse as used in the DAST and its various forms: "…'drug abuse' refers to (1) the use of prescribed or over the counter drugs in excess of the directions and (2) any nonmedical use of drugs" (p. 4).

Clinical sample – A sample that was selected using a specific diagnosis (e.g., substance use disorder) as a criterion to be included in the study.

Nonclinical sample – A community sample in which no criterion was used for participant inclusion in the study.

Summary of Chapter 1

Due to an increase in illicit drug use and the disparity between the proportions of individuals requiring treatment for drug abuse and those individuals who actually receive treatment (NIDA, 2018), it is essential for medical and mental health professionals to use psychometrically sound instruments to screen for drug abuse and related consequences and the information needed to accurately interpret results. The purpose of this study is to

analyze the psychometric properties of the DAST-20 and provide an overall summary of the instrument in order to give clinicians and researchers a more comprehensive understanding of the usefulness of the instrument. The implications of using the DAST-20 with particular populations and for interpretation of results will be discussed. In this chapter, the study was introduced, and the problem, purpose, and significance of the current study discussed. In Chapter 2, the current literature on drug abuse and the DAST-20 will be reviewed.

CHAPTER 2: REVIEW OF THE LITERATURE

Due to the prevalence and severity of drug abuse, it is imperative that professional services are available to aid individuals dealing with substance abuse problems. In order to provide the most effective treatment, professionals in the field need to have screening and diagnostic tools available that yield reliable and valid scores to correctly identify individuals with drug abuse. This study aims to examine one of these instruments, the 20-item version of the Drug Abuse Screening Test (DAST-20), which examines the presence of consequences related to drug abuse. An overall analysis of the psychometric properties of the instrument will be conducted. In this chapter, relevant literature on substance abuse, consequences related to substance abuse, issues pertaining to causality, ethical considerations, theoretical orientation, and a background on the formation and initial examination of the psychometric properties of the DAST-20 will be reviewed.

Substance Abuse

Rasmussen (2000, p. 8) defined substance abuse as "a maladaptive pattern of substance use that leads to clinically significant impairment or distress." The use of substances can lead to various consequences that affect an individual's social, physical, psychological, occupational, and legal domains (APA, 2013; Newcomb & Locke, 2005; Rasmussen, 2000). The American Psychiatric Association (APA, 2013) provided diagnostic guidelines for assessing substance use disorders. Although there are slight variances among classes of substances and diagnostic criteria, the main tenet that defines all diagnoses of a substance use disorder is the ongoing use of a substance, regardless of the presence of symptoms, and problems that occur associated with the individual's use of the substance. According to the *Diagnostic and Statistical Manual of Mental*

Disorders, 5th edition (DSM-5; APA, 2013), there are 11 criteria used to diagnose substance use disorder, regardless of the illicit substance being used. These 11 criteria are organized in four categories: impaired control, social impairment, risky use, and pharmacological.

The first category, impaired control, includes criteria one through four, which dictate that an individual may increase dosage or use the substance for an extended amount of time, display an inability to reduce dosage or frequency of use, devote a significant portion of their time to either procuring, using, or recuperating from the substance, and exhibit cravings for the substance. The second category, social impairment, consists of criteria five through seven, which indicate that the individual is experiencing decreased functioning in required tasks at work, school, or home, relational issues, and a decrease in participation of social activities.

Risky use is categorized by criteria eight and nine, which assess an individual's substance use regardless of potential physical harm. Criterion eight indicates that an individual is repeatedly using substances in dangerous settings, whereas criterion nine indicates continual use of the substance regardless of physical or psychological consequences. Lastly, the pharmacological category includes criteria 10 and 11, which signify problems of tolerance and withdrawal to the substance.

Substance use disorder exists along a continuum ranging from mild to severe. These disorder levels are determined by the number of criteria evident in relation to an individual's substance use. The DSM-5 standards for determining which severity category (mild, moderate, or severe) an individual's substance use disorder falls under is

defined as two to three criteria present, four to five criteria present, and six or more criteria present, respectively (APA, 2013).

Consequences Related to Drug Use

The DAST-20 is focused specifically on measuring the presence of consequences as an outcome of drug abuse. This is one method for determining the degree of severity the drug addiction has on an individual's life. Research has provided evidence of significant relationships between drug abuse and the consequences an individual experiences. Consequences related to drug abuse can occur in various areas of an individual's life, including physical health, psychological, and social issues, all of which can be classified as short-term or long-term effects (Newcomb & Locke, 2005). A more thorough understanding of the potential consequences of an individual's drug abuse can allow professionals to provide a more effective treatment plan when working with a client (Skinner, 1982a).

Physical Health-related Consequences

Drug abuse was linked to various physical health issues, which ranged in severity depending on the drug being abused and the frequency and quantity of use (Newcomb, 1997; Rasmussen, 2000). Rasmussen categorized physical health problems resulting from drug abuse as gastrointestinal issues, dermatological side effects, genitourinary signs, neuromuscular issues, cardiovascular issues, respiratory issues, and physical trauma related to one's drug use. Gastrointestinal issues can range from mild side effects of a particular drug (e.g., nausea, constipation, diarrhea) to developing more serious complications such as hepatitis and cirrhosis (Lange et al., 1992; Rasmussen, 2000).

Depending on the drugs used and the method of ingesting drugs, individuals may experience dermatological issues such as increased perspiration, complexion fluctuations, scarring due to burns or needle marks, as well as rashes, bruises, and skin infections (Rasmussen, 2000). Individuals who abuse drugs are also at a higher risk of developing sexually transmitted diseases. Infections and a weakened immune system were also associated with drug abuse (Lange et al., 1992). Neuromuscular issues were associated with drug abuse and included problems such as "slurred speech, tremors, lack of coordination, poor muscle tone, extremity weakness, hyperactive reflexes, seizures, and coma" (Rasmussen, 2000, p. 3). Further, individuals who abuse drugs can experience side effects such as "chest pain, palpitations, and various cardiac dysrhythmias" and blood pressure, heart rate, and breathing fluctuations (Rasmussen, 2000, p. 4). It is not uncommon for physical trauma to occur as a result of accidents and/or altercations related to one's drug abuse (Lange et al., 1992; Rasmussen, 2000).

The DAST-20 (Skinner, 1982a) has two items that ask about physical symptoms and consequences that resulted from an individual's drug abuse. Item 17 asks "Have you ever experienced withdrawal symptoms (felt sick) when you stopped taking drugs?" and item 18 asks individuals "Have you had medical problems as a result of your drug use (e.g., memory loss, hepatitis, convulsions, bleeding, etc.)?"

Psychological Consequences

Several studies indicated a significant relationship between drug use and psychological issues (Gove et al., 1979; Newcomb & Locke, 2005; Newcomb et al., 1999). In a study examining individuals' drug use and psychological well-being, a significant positive correlation was found between the number of types of drugs an

individual used and their overall psychological well-being and symptoms of mental health issues (Gove et. al, 1979). These researchers cautioned about the inability to determine directionality of this correlation, however, and posited a theory that an individual's psychological issues contributed to the decision to use drugs as an attempt to increase overall psychological well-being.

Newcomb et al. (1999) found evidence of a reciprocal relationship between drug abuse and psychological well-being when conducting a four-year longitudinal study of a community sample of 470 adults, measuring participants' overall psychological well-being and drug use status. These researchers found significant positive correlations between drug abuse and "dysphoria, agitation, psychoticism, and disorganized thinking" (p. 421). However, there was evidence of both drug abuse following psychological problems, as well as preceding psychological problems. Dysphoria reported at the initial survey was correlated with an increase in alcohol consumption four years later. Thoughts of suicide and lower self-esteem present initially was positively correlated with drug use four years later. Further, drug abuse at the initial survey was correlated with increased anxiety and hostility and decreased feelings of purpose four years later. Marijuana use at the initial survey was associated with an increase in suicidal ideation and psychoticism four years later and cocaine use initially was associated with an increase in hostility.

Social Consequences

Interpersonal issues were associated with drug abuse, including relationship strains with family members and friends (Rasmussen, 2000). It is not uncommon for drug abuse to cause issues within a family, including parenting issues (Newcomb & Locke, 2005; Rasmussen, 2000; Visser, 1991), separation or divorce (Rasmussen, 2000; Visser,

1991), as well as physical and emotional neglect of family members (Visser, 1991). Visser stated that a parent with drug abuse problems can have difficulty providing what a child needs physically and emotionally, which can lead to neglect and abuse. In a study examining the correlation between parental drug abuse and the occurrence of child abuse, researchers found the rate of child abuse more than doubled among families where at least one parent had a drug abuse problem (Walsh et al., 2003). Substance abuse also was associated with a decrease in the overall functioning and satisfaction of a marriage/relationship (Newcomb, 1997). Friendships can also deteriorate (Rasmussen, 2000). The impact of substance abuse on relationships can result in a loss of a social support network. According to Visser (1991), consequences of drug abuse within a family, including dysfunction among family members and feelings of guilt and shame about the drug abuse, can lead to isolation from not only those close to the family, but from society in general.

The DAST-20 (Skinner, 1982a) includes items that inquire about the social consequences of drug abuse. Items 8-11 address the effects of drug use on relationships with family and friends. Item 8 states, "does your spouse (or parents) ever complain about your involvement with drugs?;" item 9 asks "has your drug abuse created problems between you and your spouse or your parents?;" and item 11 asks "have you neglected your family because of your use of drugs?"

Drug abuse can also have a negative impact on an individual's career (Newcomb, 1997; Newcomb & Locke, 2005; Rasmussen, 2000). Researchers found that drug abuse is associated with impaired functioning at work, an inability to focus on tasks, a rise in absenteeism, an increase in errors, lapses in judgement, unnecessary risk-taking, and

negligent actions (Rasmussen, 2000). Due to these issues, it is not uncommon for individuals to experience job loss or repeated changes in employment (Galaif et al., 2001; Rasmussen, 2000). In a study examining the relationship between drug abuse and job adjustment among a community sample of adults, Galaif et al. found a significant correlation between occupational consequences and drug abuse. More specifically, individuals who reported issues with drug abuse had an increased chance of experiencing issues such as unpredictable employment and lowered overall work satisfaction.

There are two items on the DAST-20 that inquire about job/work consequences resulting from drug abuse. Item 12 states, "have you been in trouble at work (or school) because of drug abuse?" and item 13 asks "have you lost your job because of drug abuse?" (Skinner, 1982a).

Legal issues are another commonly reported consequence of drug abuse. Thus, legal issues were included in the DSM-5 as one criterion used to assess the presence of drug abuse (APA, 2013; Newcomb & Locke, 2005; Rasmussen, 2000). At the end of 2016, approximately 15% of inmates in state-level correctional facilities were charged with a drug-related crime (Bronson & Carson, 2019). Items 14 through 16 on the DAST-20 assess the presence or absence of legal consequences experienced due to drug abuse: "have you gotten into fights when under the influence of drugs?;" "have you engaged in illegal activities in order to obtain drugs?;" and "have you been arrested for possession of illegal drugs?;" respectively (Skinner, 1982a). The remaining seven items on the DAST-20 ask the individual about drug use activity and frequency (items 1 through 5), as well as drug treatment history (items 19 and 20; Skinner, 1982a).

It is important to note that instruments which examine the presence and severity of consequences due to drug abuse do not necessarily identify which consequences are a direct result of the drug abuse (Kiluk et al., 2018; Macleod, 2010). In a study that examined drug abuse and work adjustment, Galaif et al. (2001) found drug abuse not only had a significant impact on overall work adjustment, but individuals who had experienced unpredictable employment status earlier in life were more likely to report drug abuse four years later. Macleod (2010) argued that when considering environmental factors and adversity, professionals in the field need to

> distinguish between (a) situations where drug use mediates the association between disadvantage and psychosocial harm, (b) situations where drug use causes harm irrespective of any association with disadvantage, and (c) situations where drug use is mainly a marker of a toxic environment that causes harm through other pathways." (p. 25)

Similarly, when discussing the Addiction Severity Index (ASI), another assessment that measures an individual's addiction severity within different psychosocial categories, Kiluk et al. (2018) indicated that although the severity of issues may be measured, it does not necessarily indicate issues and severity due to one's addiction. For example, the issues present may be a result of substance abuse and present prior to substance use. In this case, the problem contributed to one's substance abuse, and the substance abuse may or may not continue to impact problem severity. Newcomb (1997) examined many facets of the consequences associated with drug abuse and argued that while drug abuse likely led to the deterioration of relationship quality, poor relationship quality may also impact drug abuse.

Although consequences may be associated with the presence of drug use, it is important not to assume that the relationship between the factors are causal or unidirectional. Therefore, ethical obligations and best practices require further data collection to examine the relationship between drug use and problems present. Skinner (1982a) included the criterion of drug use in each item of the DAST-20 instrument. Each item asks about the presence of a consequence in relation to an individual's use of drugs. For example, item 10 asks the individual to answer in yes or no format the question, "Have you lost friends because of your use of drugs?" This criterion may aid in the identification of issues that are a consequence of substance abuse based on the individual's perception of their issues.

Impaired Functioning Theory

Impaired functioning theory is based on the premise that drug abuse causes physical, psychological, and emotional consequences in an individual's life (Newcomb & Bentler, 1988). The negative consequences "can result from the psychoactive effects of the drug on cognitive, affective, and behavioral processes…[or]…arise from the deleterious effects due to the mode of ingestion and/or the metabolizing of the substance by the body on the physical level" (p. 26). The impaired functioning theory postulated that the psychoactive effects on an individual's various processes can lead to alteration of one's insight, awareness, ability to focus and process information, and/or emotional expression. Due to the differences in functioning among the various domains, an individual may have difficulty performing daily activities and coping with life's stressors. The tenets of the impaired functioning theory provide a supportive framework for the

consequences of drug abuse across three domains of an individual's life: physical health, psychological well-being, and social impairment.

Assessment and Screening Instruments

Assessment is a critical process in clinical and research practices. Assessment is necessary to ensure clients are accurately diagnosed and receive the care and resources needed (Erford, 2021). Screening, typically one of the first steps in assessment, is used with the goal of identifying the need for further assessment. Screening instruments are not intended for diagnostic purposes, but instead identify the potential need for a more comprehensive, diagnostic assessment (Erford, 2021; U.S. Preventative Services Task Force [USPSTF], 2020). Although screening instruments are not robust enough to be considered appropriate for diagnostic purposes, it is important that they yield valid and reliable scores to ensure that the proper decision is being made as to whether or not an individual is referred for further assessment. Screening instruments that yield valid and reliable scores help ensure that false negative results do not occur, reducing the risk that an individual who should be referred for further diagnostic assessment does not get overlooked (Erford, 2021).

Screening instruments that assess drug abuse are classified as either logically derived instruments or empirically derived instruments (Piazza et al., 2000). Logically derived instruments are direct in what is being measured and tend to demonstrate high content validity, but are at higher risk for response bias and, therefore, false negatives. Empirically derived instruments that assess drug abuse include items that are not necessarily directly asking about drug abuse, however, have high predictive validity of the presence of drug abuse. This type of drug abuse screening instrument may produce

less risk of response bias; however, it is limited in what information can be gleaned about an individual's drug abuse.

Multiple screening instruments exist to assess individuals for the potential presence of drug abuse, related consequences, or severity of use (USPSTF, 2020). Given the availability of screening instruments assessing drug abuse, clinicians and researchers have multiple considerations when choosing to use a specific assessment, such as client age, length of the assessment, assessment delivery format, pertinent characteristics of a client, and associated financial costs (NIDA, 2018; USPSTF, 2020).

Drug Abuse Screening Test (DAST)

The Drug Abuse Screening Test (DAST; Skinner, 1982a) is a 28-item self-report instrument developed to determine the severity of the impact of drug abuse on an individual's life. The DAST is a logically derived instrument intended to provide a means for screening for drug abuse as well as researching and evaluating treatment. Skinner indicated that the DAST can be used in conjunction with other data collection methods to gain more insight into an individual's issues related to their drug abuse and to measure the effects of treatment.

The DAST is given by means of self-administration or interview (Skinner, 1982a, 1984). Self-administration allows delivery to multiple individuals in an efficient manner. Skinner (1982a) suggested the self-administration method may help lessen bias and under-reporting of issues to allow more valid responses. However, an interview format is recommended for individuals who may have difficulty comprehending the material.

The 28 items require the responder to answer in a yes or no format for each question (Skinner, 1982a). The individual taking the assessment is instructed to answer

the questions based on drug use over the course of the past 12 months with the exception of alcohol use. Skinner (1982a) defined drug abuse in the questionnaire instructions as

> (1) the use of prescribed or over the counter drugs in excess of the directions and (2) any non-medical use of drugs. The various classes of drugs may include: cannabis (e.g. marijuana, hash), solvents, tranquillizers (e.g. Valium), barbiturates, cocaine, stimulants (e.g. speed), hallucinogens (e.g. LSD) or narcotics (e.g. heroin) (p. 1).

Since the development of the initial 28-item version of the DAST, three additional formats have been developed, including DAST-20, DAST-10, and DAST-A. The DAST-20 and the DAST-10 are both shorter versions of the DAST comprised of 20 and 10 items, respectively. The DAST-A was created for use with adolescents.

DAST-20

The DAST-20 version contains 20 yes or no questions about an individual's drug abuse and the number of consequences associated with drug abuse (Skinner, 1982a). The questions assess presence of consequences in various aspects of an individual's life, including issues with friends, family, work, incarceration, and physical and medical implications. The DAST-20 is recommended for use in clinical settings and for research evaluation as it provides more information about an individual's consequences related to drug abuse than the DAST-10 (Skinner, 1984).

The DAST-20 is scored by applying values of either 1 or 0 based on whether the individual responded with yes or no for a particular question (Skinner, 1982a). For all but two questions, a response of "yes" results in a score of 1, indicating an issue associated with substance abuse is present and a response of "no" results in a score of 0. Items 4

("Can you get through the week without using drugs?") and 5 ("Are you always able to stop using drugs when you want to?") have reverse scoring procedures, as a "yes" response indicates less negative impact of drug abuse. The score can range from 0 to 20, which represents a range of severity of consequences associated with an individual's substance abuse; a lower score indicates less severity of consequences associated with substance abuse with a score of 0 signifying that there were no consequences associated with substance abuse reported by an individual.

Additional Forms of the DAST

The DAST was also reduced to 10 yes or no items (DAST-10). Skinner (1984) recommended the DAST-10 be used to screen for drug abuse consequences since it does not provide as much information about an individual's consequences. Still, scores on the DAST-10 were found to have psychometrically sound characteristics.

The DAST-A is an adolescent version that includes items similar to the DAST-20, but the 20 questions focus on relationships with parents when assessing relationship consequences and on school when assessing an adolescent's ability to function. In contrast, the DAST-20 includes relationships with spouses/partners and occupational consequences (Skinner, 1982b). The instrument is scored in the same manner as the DAST-20. Values of either 0 or 1 are added together based on whether the individual responded with "yes" or "no" for a particular question. For all but two questions, a response of "yes" results in a score of 1, indicating a consequence associated with substance abuse is present and a response of "no" results in a score of 0. As with the DAST-20, items 4 ("Can you get through the week without using drugs?") and 5 ("Are you always able to stop using drugs when you want to?") have reverse scoring

procedures, as a "yes" response indicates less negative impact of drug abuse. The score can range from 0 to 20, which represents a range of severity of consequences associated with an individual's substance abuse; a lower score indicates less severity of consequences associated with substance abuse with a score of 0 signifying that there were no consequences associated with substance abuse reported by an individual.

The DAST-20 was chosen for the current study due to the various uses and settings recommended by Skinner (1982a) when compared to the DAST-10 and for the broader demographics with which it is commonly used, compared to the adolescent version of the instrument.

DAST Formation

The DAST (Skinner, 1982b) was developed based on the Michigan Alcoholism Screening Test (MAST; Selzer, 1971). The MAST was developed to address the absence of a standardized instrument to identify the presence of alcoholism. Selzer found that the instruments available were based on varying definitions of alcoholism and alcohol abuse and relied on poorly defined categorizations of individuals' alcohol consumption as opposed to the consequences an individual is experiencing due to their alcohol use. For example, some definitions of alcoholism included invalid observations of one's appearance and hygiene (Selzer, 1968). Discordant and invalid definitions of alcoholism made it difficult to identify individuals with an actual alcohol use problem. A standardized instrument that could produce valid and reliable scores was needed to help identify individuals with alcoholism (Selzer, 1968, 1971).

The MAST (Selzer, 1971) is comprised of 25 yes or no questions that inquire about an individual's consequences associated with use of alcohol related to relationships

with family and friends, work issues, health problems, legal issues, and psychological consequences. Scoring of the MAST is based on the magnitude of each item in addressing alcoholism (Selzer, 1968). A cumulative value of four or more points is considered a positive screening for alcoholism. Studies have indicated that the MAST is a psychometrically sound instrument for screening purposes (Skinner, 1984). A synthesis of the psychometric properties of the MAST scores revealed an aggregated internal consistency coefficient of .84, indicating the MAST to yield reliable scores for screening for alcohol-related issues; however, this value fell short of the recommended coefficient of .90 for diagnostic purposes (Minnich et al., 2018).

Consequences related to alcohol abuse can be compared to consequences an individual can experience from drug abuse (Skinner, 1984). Therefore, Skinner (1982b) created the DAST based on the MAST to provide an instrument that could screen for drug abuse consequences.

Initial Evaluation of the DAST-20

The initial evaluation of the score validity and reliability of the DAST yielded positive results for all forms of the instrument (Skinner, 1982b). Skinner examined the psychometric properties of the original 28-item version of the DAST in a study that sampled 223 individuals in an addiction treatment program and DAST scores had high internal consistency (*KR-20* = .92).

Scores from the DAST-20 were highly correlated with the 28-item DAST (r = .99) and DAST-10 (r = .98) among the total sample (Skinner, 1984). The DAST-28 and DAST-10 scores correlated r = .97. High internal consistency was reported for the DAST-20 scores (*KR-20* = .95) for the entire sample (including substance abuse of drugs

and/or alcohol) and for a subgroup that did not include participants who reported only abuse of alcohol (i.e., drug use only; *KR-20* = .86; Skinner, 1982b).

An evaluation of the DAST-A was conducted by Martino et al. (2000) with adolescents in an inpatient facility. The DAST-A scores were significantly positively correlated with five related instruments. The researchers also found that the DAST-A scores had adequate internal consistency (*KR-20* = .91) and test-retest reliability (*r* = .89).

Advantages and Limitations of the DAST

Skinner (1982a) identified numerous advantages and limitations of using the DAST to assess an individuals' level of issues associated with substance abuse. The DAST is a cost-efficient (i.e., free) means of collecting a measure of drug abuse impact that can be administered individually or in a group format (Skinner, 1982a, 1984). The DAST not only provides an overall score indicating the level of consequences associated with drug abuse, but also can be used as a foundation to gain further data and information on the identified areas most affected by drug abuse. The DAST can be used in conjunction with other instruments to validate findings. However, the DAST is susceptible to participants underreporting their substance abuse. Also, the DAST is not a comprehensive assessment of all aspects of an individual's substance abuse and should be used in conjunction with other means of collecting data.

Psychometric Properties of the DAST-20 Scores

Since the initial study on the psychometric properties of the DAST-20 scores (Skinner, 1982b), numerous studies reported quantitative data on the reliability and validity of DAST-20 scores among various populations (Cassidy et al., 2008; Cocco & Carey, 1998; Grekin et al., 2010; Salehi et al., 2012). Due to varying research conditions,

limitations, and populations, differences in psychometric properties of the DAST-20 have been reported, which can have a significant effect on a professional's ability to interpret and apply DAST-20 scores in treatment and research (Erford et al., 2010; Lipsey & Wilson, 2001). A preliminary review of these findings is summarized below.

Reliability

Numerous studies reported internal consistency values for the DAST-20 scores (Aubry et al., 2011; Bliss et al., 2008; Boyd et al., 2009; Brocato & Wagner, 2008; Burnett et al., 2013; Cassidy et al., 2008; Fineran et al., 2010; Forbey & Ben-Porath, 2008; Forbey et al. 2011; Gerlock, 2004; Kaslow et al., 2006; Kimbrel et al., 2011; Murphy et al., 2007; Nordfjaern, 2013; Perepletchikova et al., 2012; Reviere et al., 2007; Rosenkranz et al., 2012; Sabato et al., 2013; Sullivan et al., 2009; Ting et al., 2009). Among these studies, the reported reliability value (*KR-20*) ranged from 0.71 (Rosenkranz et al., 2012) to 0.998 (Cassidy et al., 2008). Erford (2021) recommended internal consistency values of at least .80 to consider scores from an instrument acceptable for screening purposes and at least .90 for use as a diagnostic tool. Such large variances in reported internal consistency values can affect best practice across diverse populations.

Validity

Data on the validity of the DAST-20 scores were reported in numerous studies over the last few decades. Convergent validity was demonstrated with scores from the DAST-20 and similar instruments, including the MMPI-2 Restructured Clinical Scales (Forbey & Ben-Porath, 2008), the Externalizing Inventory (100-item; Hall et al., 2007), the MAST (Marshall & Marshall, 2006), the Marijuana Screening Inventory (MSI-X;

Alexander & Leung, 2006), the Research Institute on Addictions Self-Inventory (Shuggi et al., 2006), the NEO-PI-R (Conner et al., 2004), the CAGE and the SASSI-2 (Teslak, 2000), and the ADS and MAST (Gavin et al., 1989). Further analysis and discussion of findings in regard to convergent validity will be provided in the Results and Discussion sections of this study.

Multiple studies reported data on the diagnostic validity (decision reliability) of the DAST-20 scores with varying findings on the sensitivity, specificity, positive predictive power, and negative predictive power. Multiple cut-off scores have been recommended across these studies. The data reported from these studies on the criterion validity of the DAST-20 will be summarized in the Results chapter of this study.

Ethical Considerations When Using Screening Instruments

Counselors and researchers must take numerous ethical considerations into account when choosing instruments to use with clients and research participants. Specifically, the American Counseling Association (ACA, 2014) outlined ethical standards pertaining to the psychometric properties of an instrument, the population in which instruments are used, and empirical data that support the use of an instrument. These specific clauses include E.6.a, E.8, and E.9.b.

ACA's *Code of Ethics* (2014) supports the need for various methods of data collection when working with clients and research participants in order to gather a more complex understanding of an individual's drug use and psychosocial factors. When using instruments, professionals also need to be aware of any multicultural considerations that may affect score reliability and validity of an instrument and how results are interpreted. For example, researchers reported differences between the consequences experienced by

men and women due to substance abuse (Newcomb, 1997; Plant et al., 2002; Robbins, 1989). Robbins found women were more likely to report consequences of depression, trust issues, and feeling distressed, whereas men were more likely to report issues related to school, work, finances, legal problems, driving under the influence, and urgent health issues. The reported differences between men and women were attributed to men's higher occurrence of substance abuse than women. Women who reported abusing marijuana and cocaine were also more likely to report relational consequences such as engaging in altercations and arguments. In contrast, a study conducted using a cross-sectional survey of individuals in the United Kingdom found that women were more likely than men to report negative consequences due to substance abuse (Plant et al.). Although various studies reported on differences in experiences among men and women, these differences are not always considered by clinicians and researchers (Newcomb).

Another example of multicultural factors that could have significant impact on outcomes is age (Newcomb, 1997). Newcomb made the argument that different consequences of drug abuse may be present depending on an individual's age and the tasks that occur during that particular life stage.

Issues with Application of Individual Study Results

It is not uncommon for researchers conducting similar studies to report contradicting results leading to differing conclusions (Erford et al., 2015; Lipsey & Wilson, 2001). This can make it difficult for professionals to know how to implement the conclusions from the various studies when working with clients. The differences in findings among the various studies can be attributed to the specific samples, researchers, methods, and conditions of each individual study. Many of the samples in the studies that

reported data on the psychometric properties of the DAST-20 are very specific and not as generalizable to the population of clients of mental health professionals. Some examples of these samples include patients with burn injuries (Salehi et al., 2012), individuals with a mental disorder (Cocco & Carey, 1998), women who recently gave birth (Grekin et al., 2010), and individuals with first-episode psychosis (Cassidy et al., 2008). Further, some studies used clinical samples while others used nonclinical samples, and even different diagnostic criteria (e.g., DSM, ICD, author created) were used to assess the psychometric properties of the DAST-20. Multiple studies reported issues of generalizability of results in the discussion of study limitations (e.g., Alexander & Leung, 2006; Grekin et al., 2010).

Much like a meta-analysis, a psychometric synthesis can allow for a more comprehensive examination of a topic, which better informs professionals making decisions based on the available literature (Erford et al., 2015; Lipsey & Wilson, 2001). Lipsey and Wilson explained the benefits of meta-analysis by comparing it to looking at individual pixels versus looking at the overall picture from a distance. Aggregating the data from all the studies of varying samples, methods, and conditions allows for "smoothing the resulting picture into a composite, much as a magazine picture looks more crisp and coherent at arms length than when the pixels are examined through a magnifying glass" (Lipsey & Wilson, 2001, p. 167).

Chapter 2 Summary

Most studies of the psychometric properties of the DAST-20 concluded the instrument yielded reliable and valid scores among various populations. However, multiple optimal cutoff scores were reported, and findings of various factor analyses and

diagnostic validity studies (sensitivity, specificity), and internal consistency values differed. Many of the studies used homogeneous samples in a specific setting, which limit the generalizability of the findings. The purpose of the current study is to examine the findings from all of the identified studies and analyze the aggregated data to provide insight on the overall psychometric properties of the DAST-20 scores and then discuss implications for researchers and clinicians using the DAST-20 with various populations.

CHAPTER 3: METHODOLOGY

Research Design

Psychometric synthesis procedures outlined by Erford et al. (2015) were used to quantitatively examine the psychometric properties of the DAST-20 (Skinner, 1982a). Although the research design is modeled on the study named above (Erford et al., 2015), the correct classification of this research design is called a psychometric synthesis. A psychometric synthesis has many procedural characteristics in common with a meta-analysis. However, in recent years, very specific procedures define a meta-analysis (Lipsey & Wilson, 2001). Therefore, the article this study models is more correctly classified as a psychometric synthesis. The procedures in a psychometric synthesis are included in a meta-analysis, however not all procedures of a meta-analysis are included in a psychometric synthesis.

A synthesis is conducted by following specific steps to examine the findings of empirical studies that have produced results on a common subject of research (Erford et al., 2015). This research design involves collecting all possible data from studies examining a common research question that meet the established criteria of the study and using particular statistical procedures to summarize these findings. This includes establishing criteria for article inclusion, searching for relevant articles, examining articles for criteria and completing the selection process, data extraction, and data analysis. This psychometric synthesis will examine the reliability and validity of the DAST-20 scores, combining the results of numerous articles that report data for the instrument. After selecting the studies used in the psychometric synthesis, analyses were

conducted on the reported results to produce a more comprehensive understanding of outcomes.

Procedure

The procedure utilized in this study was outlined by Erford et al. (2015) in a psychometric synthesis study that examined the psychometric properties of the Beck Depression Inventory – Second Edition (BDI-II; Beck et al., 1996). The procedure includes specific methods for searching for and identifying potential articles for inclusion in the study, accepting or rejecting these studies, extracting data, and analyzing the data. These procedures are described in detail below.

Article Selection and Criteria

A search for sources that contained psychometric data on the DAST-20 was conducted. The search for relevant sources included published journal articles as well as unpublished documents, such as dissertations, theses, and other unpublished sources that were available through research databases, including all possible sources of psychometric data on the DAST-20 into the study, reduced potential publication bias (Erford et al., 2015).

Lipsey and Wilson (2001) identified several areas of eligibility criteria that apply to most meta-analyses, including "distinguishing features, research respondents, key variables, research methods, cultural and linguistic range, time frame, and publication type" (pp. 16-17). Candidate articles selected for inclusion in the current study were published or available between 1982 and 2014 (the year this study was started). For the purposes of this study, more current studies will not be added, however, any articles published between 2015 and 2021 will be included before this manuscript is submitted

for publication. Selection for inclusion also required the use of the English version of the shortened 20-item form of the DAST (DAST-20) and at least one psychometric property of the DAST-20 reported, including internal consistency, test-retest reliability, convergent validity, structural validity, diagnostic validity, and descriptive statistics. Studies that reported means or standard deviations of the DAST scores were only included if data were collected from nonclinical samples.

Search Procedures

The initial search for included sources was conducted through research databases, including PsychINFO, ERIC, Academic Search Premier, MEDLINE, and Cochrane Database of Systematic Reviews using the search text "Drug Abuse Screening Test" from dates 1982 (the year the DAST was published) to 2014. To ensure all relevant data was included in the analysis, a second search was conducted of the reference list of any source that reported psychometric data. This redundant procedure was undertaken to obtain potential sources that were not included in the original electronic search. Also, to make sure data from a particular study was only used once in the data analysis, the list of articles was searched for repeated occurrences. For example, dissertations and articles resulting from a dissertation were analyzed for relevant psychometric data, but the sample result was only used once to maintain independence of results. Once all the articles were obtained through the multi-step search process, the selection process was started.

Selection Process

The full text version of all articles identified by the search process were obtained and numbered to organize selection procedures. The selection process was independently

conducted by the author and two additional research assistants in order to ensure accurate inclusion and rejection decisions of each article. The inclusion criteria listed above were used to evaluate each article to determine whether or not the study provided usable data for analysis. A further description of the psychometric variables of interest is provided below. Each researcher inserted an independent decision for inclusion of each article in a chart. If the decision was to accept an article for inclusion in the study, a short description of the data of interest was included in the chart. A final compiled chart was completed, listing each researcher's decision for inclusion of each article. An inter-rater agreement percentage was calculated at 95%. Ultimately, the decision to include or exclude articles that were not agreed upon by the selectors was deferred to Dr. Erford, who has completed many psychometric syntheses and has published on the research design.

After the selection process was finalized, data was extracted from all articles accepted for inclusion. To ensure that aggregated data from a particular study was only represented in the data analysis once, duplicative articles were rejected and dissertations and articles using the same sample were used only once so the resulting data would maintain independence.

Psychometric Variables Analyzed

The psychometric variables of interest for the analysis of this study included data on internal consistency, test-retest reliability, convergent validity, structural validity, diagnostic validity, and the means and standard deviations of nonclinical samples. Reliability statistics of interest included internal consistency coefficients and test-retest reliability data. Internal consistency refers to how well a set of items on an assessment inter-correlate or hang together (Erford, 2021). Because responses to the DAST-20 are

dichotomous (yes-no), Kuder-Richardson formula 20 (*KR-20*) was the most common statistic used to examine the internal consistency of the DAST-20. Test-retest reliability statistics represent the degree of consistency of an individual's responses on an instrument when administered twice. Pearson *r* is the statistic most commonly used to examine test-retest reliability of the DAST-20. The time lapsed between the first and second administrations of the DAST-20 was noted for each eligible study.

Multiple validity statistics were examined to determine the ability of the DAST-20 in accurately measuring the construct of interest (Erford, 2021). Validity measures of interest in the current psychometric synthesis included measures of convergent and diagnostic validity (i.e., overall accuracy, sensitivity, specificity, positive predictive power, negative predictive power), and structural validity (i.e., confirmatory factor analysis, exploratory factor analysis). Construct validity refers to the ability of an instrument to accurately measure the construct of interest, such as by providing statistical evidence on convergent and discriminant validity. In the current study, articles were screened for convergent validity statistics on the DAST-20. Convergent validity helps to determine whether an instrument demonstrates adequate construct validity by statistically comparing it with scores on other instruments measuring the same or a very similar construct. If a high positive correlation is found between the two instruments, it is an indication that the instrument under question demonstrates construct validity. In the current psychometric synthesis study, Pearson *r* was normally the statistic used to examine convergent validity.

Diagnostic validity refers to the ability of an instrument to accurately determine the presence or absence of a specific diagnosis or condition. In order to determine the

diagnostic validity of an instrument, the scores of the instrument are compared to the findings of a clinical evaluation or diagnosis of an individual completed by a mental health professional (Erford, 2021). Although the DAST-20 is a screening instrument and is not intended as a diagnostic assessment, it is still important that the instrument accurately identifies individuals who most likely have an issue with drug abuse to allow appropriate and timely referral for further evaluation, as needed. In the initial examination of the DAST-20, Skinner (1982b) recommended a cutoff score of six for screening purposes. Data of interest on the diagnostic validity of the DAST-20 included the percent of individuals correctly classified overall by the instrument, specificity values, sensitivity values, positive predictive values, and negative predictive values.

The test cutoff score and the criterion cutoff determined for the clinical evaluation are used to examine these various indices of diagnostic validity (Erford, 2021). These indices include sensitivity, specificity, positive predictive power, and negative predictive power. Sensitivity refers to the ratio of true positives identified by the instrument cutoff (individuals correctly identified to have a problem with drug abuse) out of all the individuals identified to have a problem with drug abuse by the criterion cutoff determined by the clinical evaluator (true positives plus false negatives of the instrument). Specificity refers to the ratio of true negatives identified by the instrument cutoff (individuals correctly identified to not have a problem with drug abuse) out of all the individuals who were determined not to have a problem with drug abuse by the criterion cutoff determined by the clinical evaluator (true negatives and false positives of the instrument). Positive predictive power refers to the instrument's ability to predict the presence of a drug abuse problem. This statistic is determined by calculating the ratio of

true positives identified by the test out of all the individuals identified by the test to have a problem with drug abuse (true positives plus false positives). Finally, negative predictive power refers to the instrument's ability to predict the absence of a drug abuse problem. This statistic is determined by calculating the ratio of true negatives (individuals correctly identified to not have a problem with drug abuse) out of all the individuals the instrument determined not to have a problem with drug abuse (true negatives plus false negatives).

Structural validity is determined by using a statistical process called factor analysis to examine the format of an instrument and to group related items into various aspects of a construct by identifying which items correlate with each other (Erford, 2021). Each group of related items represents a factor. Studies of both exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) were of interest. An EFA is the process of examining structural validity when evidence of the number of factors and how items are grouped among factors is not available (Erford, 2021). In this process, the factors that represent various constructs and the items that comprise the various factors are determined by mathematical procedures. A CFA is the process of examining structural validity when theoretical evidence is used to identify the number of factors thought to be present and how the items load on each factor prior to the analysis.

Descriptive statistics, including means and standard deviation values, were also extracted for analysis. Only means and standard deviations reported on nonclinical samples were analyzed as data from clinical samples are not likely to represent the general population parameters.

The data extraction process was completed by creating tables for each psychometric property identified in the articles. For example, one table was created to record all articles that reported internal consistency values of the DAST-20, including the article number, sample size (*n*), type of sample (nonclinical or clinical), and the *KR-20* coefficient. If a value was also reported specifically for men or women, these values were also stratified. Once the relevant data were extracted, analytic procedures were used to examine the psychometric properties of the DAST-20.

Data Analysis

Research Questions

The following research questions were explored:

1. What are the aggregated psychometric properties of the DAST-20 scores (i.e., internal consistency, test-retest reliability, diagnostic validity, structural validity, convergent validity)?
2. What are the aggregated mean scores and standard deviations of the DAST-20 in nonclinical samples?
3. Are there significant differences in the psychometric properties of the DAST-20 scores among various sample characteristics?

Statistical Analysis

Statistical analyses were conducted on the reliability and validity data described above. All of the data included in the study were independent, which means that all data were only represented in the study once (Erford et al., 2010). When completing the article selection and data extraction processes, duplicate articles were deleted, and data sets

published in more than one accepted article were used only once in the data aggregation and analysis.

Before aggregation, all data were weighted by the corresponding sample size before being analyzed in order to reduce sample size bias (Erford et al., 2015). According to classical test theory, an observed score is comprised of the true score and an error of measurement (Erford, 2021). The error of measurement is assumed to be random. As a result, the average of the observed scores converges on the true score because they are likely normally distributed. Therefore, internal consistency coefficients were weighted by sample size and then averaged together to get aggregated reliability data. This process was repeated for test-retest reliability aggregations. When interpreting the reliability coefficient, coefficients .80 and higher are considered to be acceptable for screening purposes and coefficients of .90 and above are acceptable for diagnostic purposes.

Unlike analysis of the reliability statistics, validity coefficients cannot be analyzed by operating under the assumption of true score and error variance because systematic error must be considered (Erford, 2021; Erford et al., 2015). Therefore, Pearson *r* coefficients were converted to *z*-values before being weighted by sample size and aggregated using the equation proposed by Hedges and Olkin (1985), $z(r) = \frac{1}{2}\log[\frac{1+r}{1-r}]$. For each convergent validity comparison, the Pearson *r* statistic was converted to a *z*-score, then weighted by sample size and averaged. Finally, the averaged *z*-scores were converted back to Pearson *r* coefficients. To interpret the Pearson *r* coefficients, the following effect sizes were used, as recommended by Lipsey and Wilson (2001): 0.1 signifies a small effect size, 0.3 signifies a medium effect size, and 0.5 signifies a large effect size.

Evidence of diagnostic validity, as described above, was aggregated as possible across the indices of percent of correct classifications, specificity values, sensitivity values, positive predictive values, and negative predictive values. For each cut off score on a similar criterion, index values were weighted by sample size, aggregated, and then averaged. Evidence of structural validity was extracted from the accepted articles and reported in table format. Finally, means and standard deviations of nonclinical samples were weighted by sample size and averaged across studies.

Chapter 3 Summary

In chapter 3, the methodology for this study was reviewed. The research design for this study is a psychometric synthesis, which uses specific procedures to collect and analyze psychometric data from available research on a given topic of interest. The procedures include setting criteria, completing an article search, screening articles for inclusion in the study, data extraction, and data analysis. Data analysis was conducted on reliability data (internal consistency, test-retest reliability), validity data (convergent validity, structural validity, diagnostic validity), and means and standard deviations of nonclinical samples. The data analysis and findings will be provided in chapter 4.

CHAPTER 4: RESULTS

Introduction

Psychometric synthesis procedures (Erford et al., 2015) were conducted and the available data on the psychometric properties (i.e., internal consistency, test-retest reliability, convergent validity, structural validity, diagnostic validity, descriptive statistics) of the DAST-20 (Skinner, 1982a) were analyzed as outlined in Chapter 3. In this chapter, the results of the article search and selection are reported and a detailed report of the data and analysis procedures are provided.

Sample

A total of 839 articles were produced by the electronic database search and 14 additional articles were found by hand searching candidate article reference lists, resulting in a total of 853 potential articles screened for the study. In order to be selected for inclusion into the study each article had to be in English, published between 1982 and 2014, use the English version of the 20-item DAST, and report at least one type of data on the psychometric properties of the instrument. The selection process produced a total of 56 articles accepted into the psychometric synthesis. Out of the 56 articles, 34 articles provided data on internal consistency, three articles provided data on test-retest reliability, 17 articles provided data on convergent validity, three articles provided data on factor analysis, seven articles provided data on diagnostic validity, and 12 studies provided mean and standard deviation descriptive statistics on the DAST-20 for nonclinical samples.

The author and a committee member screened each of the 853 articles for relevant data and the required criteria for inclusion into the study. The interrater agreement for

article selection and inclusion was 94.8%, an inter-rater correlation of $r = .851$ ($p < .001$), both indicators of strong inter-rater concordance (Erford, 2021). Any disagreement in inclusion decisions were discussed until consensus was achieved.

DAST-20 Reliability

In the present study, data on the reliability of the DAST-20 scores was collected, including evidence of internal consistency (i.e., *KR-20*) and test-retest reliability (i.e., Pearson *r*). The results are summarized below.

Internal Consistency

A total of 33 articles ($j = 33$) accepted into the current study provided internal consistency data (*KR-20*). One of these articles included coefficient alpha values for three independent sample groups, resulting in a total of 35 internal consistency scores included in the aggregation ($k = 35$) and a total sample size of 15,546 (see Table 1). Each *KR-20* was weighted by the corresponding sample size. The aggregated total DAST-20 internal consistency for this study embedded within a 95% confidence interval was *KR-20* = .819[.819, .820].

Table 1: Total Aggregated Internal Consistency

Article	***N***	***KR-20***	***N* x *KR-20***
Aubry et al. (2012)	329	.93	305.97
Bliss et al. (2008)	178	.91	161.98
Boyd et al. (2009)	142	.96	136.32
Brocato & Wagner (2008)	141	.75	105.75
Burnett et al. (2013)	1,874	.78	1461.72
Cassidy et al. (2008)	84	.998	83.832
Cocco & Carey (1998)	97	.92	89.24
Fineran et al. (2010)	200	.91	182.00
Fleury et al. (2012)	2,443	.74	1807.82
Forbey et al. (2008)	1,038	.81	840.78
Forbey et al. (2011)	213	.95	202.35

Forbey et al. (2013) sample a	1,065	.81	862.65
Forbey et al, (2013) sample b	613	.93	570.09
Forbey et al. (2013) sample c	164	.94	154.16
Irving & Schweiger (1991)	400	.77	308.00
Kaslow et al. (2006)	274	.94	257.56
Kimbrel et al. (2011)	162	.77	124.74
Looman & Abracen (2013)	348	.90	313.20
Mowbray et al. (2006)	379	.94	356.26
Murphy et al. (2007)	139	.90	125.10
Nelson et al. (2011)	1,637	.77	1260.49
Nochajski et al. (2013)	520	.82	426.40
Nordfjærn (2013)	203	.795	161.385
Perepletchikova et al. (2012)	99	.96	95.04
Reviere et al. (2007)	200	.92	184.00
Rosenkranz, Muller et al. (2012)	216	.71	153.36
Rosenkranz, Henderson et al. (2012)	188	.72	135.36
Saltstone et al. (1994)	318	.88	279.84
Skinner (1984)	223	.94	209.62
Skinner & Goldberg (1986)	105	.74	77.70
Stuart & Holtzworth-Munroe (2005)	86	.91	78.26
Sullivan et al. (2009)	412	.86	354.32
Teslak (2000)	142	.90	127.80
Weber (2008)	824	.81	667.44
Weinstein (1999)	90	.88	79.20
Total	15,546		12,739.737

A total of seven articles ($j = 7$) accepted into the current study provided internal consistency data (*KR-20*) for female participants. The articles resulted in a total sample size of 1,903 (see Table 2). Each coefficient α was weighted by the corresponding sample size. The aggregated total DAST-20 internal consistency for females was *KR-20* = .909[.846, .954].

Table 2: Internal Consistency - Females

Article	*N*	*KR-20*	*N x KR-20*
Bliss et al. (2008)	178	.91	161.98
Boyd et al. (2009)	142	.96	136.32
Kaslow et al. (2006)	274	.94	257.56
Mowbray et al. (2006)	379	.94	356.26
Reviere et al. (2007)	200	.92	184.00
Saltstone et al. (1994)	318	.88	279.84
Sullivan et al. (2009)	412	.86	354.32
Total	1,903		1,730.28

A total of seven articles ($j = 7$) accepted into the current study provided internal consistency data (*KR-20*) for male participants. One article provided two samples, resulting in a total of eight internal consistency values that were aggregated ($k = 8$). The articles resulted in a total sample size of 1,846 (see Table 3). Each coefficient was weighted by the corresponding sample size. The aggregated total DAST-20 internal consistency for males was *KR-20* = .908[.862, .954).

Table 3 – Internal Consistency - Males

Article	*N*	*KR-20*	*N* x *KR-20*
Brocato & Wagner (2008)	141	.75	105.75
Forbey et al. (2011)	213	.95	202.35
Forbey et al. (2013) sample b	613	.93	570.09
Forbey et al. (2013) sample c	164	.94	154.16
Looman & Abracen (2013)	348	.90	313.20
Murphy et al. (2007)	139	.90	125.10
Stuart & Holtzworth-Munroe (2005)	86	.91	78.26
Ting (2009)	142	.90	127.80
Total	1,846		1,676.71

A total of six articles ($j = 6$) accepted into the current study provided internal consistency data (*KR-20*) for nonclinical samples. One article provided two samples,

resulting in a total of seven aggregated internal consistency values ($k = 7$). The articles resulted in a total sample size of $n = 7{,}319$ (see Table 4). Each *KR-20* was weighted by the corresponding sample size. The aggregated total DAST-20 internal consistency for the nonclinical samples was *KR-20* = .793[.770, .816].

Table 4 – Internal Consistency – Nonclinical Samples

Article	*N*	*KR-20*	*N* x *KR-20*
Burnett et al. (2013)	1,874	.78	1,461.72
Fineran et al. (2010)	200	.91	182.00
Fleury et al. (2012)	2,443	.74	1,807.82
Forbey & Ben-Porath (2008)	1,038	.81	840.78
Forbey et al. (2013) sample a	1,065	.81	862.65
Forbey et al. (2013) sample b	613	.93	570.09
Stuart & Holtzworth-Munroe (2005)	86	.91	78.26
Total	7,319		5,803.32

Clinical sample results were also aggregated and included a total of 28 articles and studies ($j = 28$; $k = 28$) accepted into the current study that provided internal consistency data (*KR-20*) for clinical samples. The articles resulted in a total sample size of 8,227 (see Table 5). Each coefficient was weighted by the corresponding sample size. The aggregated total DAST-20 internal consistency for clinical samples was *KR-20* = .843[.821, .865].

Table 5 – Internal Consistency – Clinical Samples

Article	*N*	*KR-20*	*N* x *KR-20*
Aubry et al. (2012)	329	.93	305.97
Bliss et al. (2008)	178	.91	161.98
Boyd et al. (2009)	142	.96	136.32
Brocato & Wagner (2008)	141	.75	105.75
Cassidy et al. (2008)	84	.998	83.832
Cocco & Carey (1998)	97	.92	89.24

Forbey et al. (2011)	213	.95	202.35
Forbey et al. (2013) sample c	164	.94	154.16
Looman & Abracen (2013)	348	.90	313.20
Mowbray et al. (2006)	379	.94	356.26
Murphy et al. (2007)	139	.90	125.10
Irving & Schweiger (1991)	400	.77	308.00
Kaslow et al. (2006)	274	.94	257.56
Kimbrel et al. (2011)	162	.77	124.74
Nelson et al. (2011)	1,637	.77	1,260.49
Nordfjærn (2013)	203	.795	161.385
Nochajski et al. (2013)	520	.82	426.40
Perepletchikova et al. (2012)	99	.96	95.04
Reviere et al. (2007)	200	.92	184.00
Rosenkranz, Muller et al. (2012)	188	.72	135.36
Rosenkranz, Henderson et al. (2012)	216	.71	153.36
Saltstone et al. (1994)	318	.88	279.84
Skinner (1984)	223	.94	209.62
Skinner & Goldberg (1986)	105	.74	77.70
Sullivan et al. (2009)	412	.86	354.32
Ting et al. (2009)	142	.90	127.80
Weber (2008)	824	.81	667.44
Weinstein (1999)	90	.88	79.20
Total	8,227		6,936.417

Test-retest Reliability and Other Measures of Consistency

Only one article included in the current study reported data on test-retest reliability. Peters et al. (2000) reported a test-retest reliability value (r_{tt}) of .95[.70, 1.00; n = 60] with a 72-hour period between test administrations.

Somewhat relatedly, Cocco and Carey (1998) reported intraclass correlations with an ICC value of .78[.58, .98; n = 97]. The period of time between test administrations was between 7-43 days depending on the participant.

In addition, Conner et al. (2004) reported an interrater reliability of .62[.34, .90; *n* = 48] between a participant's self-report and their partner's report of the participant's drug abuse.

DAST-20 Validity

Data on the score validity of the DAST-20 was collected, including evidence of convergent validity (i.e., Pearson *r*), structural validity (i.e., confirmatory factor analysis, exploratory factor analysis), and diagnostic validity (i.e., overall accuracy, sensitivity, specificity, positive predictive power, negative predictive power). The results are summarized below.

Convergent Validity

Evidence of convergent validity was most commonly reported as Pearson *r* coefficients. A total of 17 articles reported convergent validity data on the DAST-20. Only two instruments (not including correlations between the DAST-20 and other forms of the DAST) were used in multiple studies and aggregated. The first instrument reported in multiple studies is the Michigan Alcoholism Screening Test (MAST; 25-items) (see Table 6). A total for four studies ($j = 4$; $k = 4$), with a combined sample size of 766, included Pearson *r* values between the DAST-20 and the MAST (25-item version). Three of the four studies reported a positive correlation between the MAST and the DAST-20 and one study (Gavin et. al, 1989) reported a negative correlation between the two instruments. After converting each Pearson *r* value to *z*-values, weighting the value by sample size ($z \times n$), averaging the scores, and converting the aggregated *z*-score back to a Pearson *r* value, the aggregated correlation was $r = .11$[.04, .18], a no to small effect size (Cohen, 1988).

Table 6 – Correlation with the Michigan Alcohol Screening Test (MAST; 25 items)

Article	*N*	*r*	*z*-value	*z* x *n*
Cocco & Carey (1998)	97	.52	0.576	55.872
Gavin et al. (1989)	501	-.19	-0.192	-96.192
Marshall & Marshall (2006)	80	.56	0.633	50.64
Weinstein (1999)	88	.659	0.793	69.784
Total	766			80.104

The second instrument that had multiple correlation coefficients reported ($j = 2$; $k = 2$; $n = 6{,}504$) with the DAST-20 is the Alcohol Dependence Scale (ADS) (see Table 7). Similar to the MAST (25-item) correlation, Gavin et al. (1989) found a negative correlation between the DAST-20 and the ADS, while Shuggi et al. (2006) reported a positive correlation between the two instruments. The aggregated convergent validity value for the DAST-20 and the ADS is $r = 0.27[.25, .29]$, a small to medium effect size.

Table 7 – Correlation with the Alcohol Dependence Scale

Article	N	*r*	*z*-value	*z* x n
Gavin et al. (1989)	501	-.13	-0.131	-65.13
Shuggi et al. (2006)	6,003	.30	0.310	1,860.93
Total	6,504			80.104

Although the ADS and the MAST (25-item) were the only instruments with multiple correlation data with the DAST-20, a total of 27 additional instruments or subscales had a single report of convergent validity data with the DAST-20. One study (Møller & Linaker, 2010) reported phi coefficients, whereas all other articles reported *r* values. Moller and Linaker (2010) reported a phi value of .41 ($n = 37$) for the ICD-10 and the DAST-20 and a phi value of .34 ($n = 37$) for the DUS and DAST-20. The convergent validity (r) data of the other instruments are reported in Table 8.

Table 8 – Additional Convergent Validity Correlations with DAST-20.

Article	Instrument	*n*	*r*
Alexander & Leung (2006)	Marijuana Screening Inventory (MSI-X)	107	.531

Alterman et al. (2007)	Addiction Severity Index – 5th Edition (ASI-5) Recent Drug Problems	2,142	.65
	ASI-5 Lifetime Drug Problems	2,142	.63
Brocato & Wagner (2008)	DSM-IV TR Checklist	141	.66
Cocco & Carey (1998)	Addiction Severity Index (ASI) – Drug Composite Score	97	.42
	Clinician Rating Scales for Drug Use (CRS-Drug)	97	.40
	ASI – Alcohol Composite Score	97	.33
	Global Assessment of Functioning (GAF)	97	-.14
Forbey et al. (2008)	MMPI-2-RC RC4 subscale – Antisocial Behavior	Males 407	.53
		Females 631	.52
Forbey et al. (2011)	Minnesota Multiphasic Personality Inventory 2 (MMPI-2)	159	.52
	MMPI-2 MAC-R (MacAndrew Alcoholism Scale-Revised)	159	.35
	MMPI-2 AAS (Addiction Acknowledgement Scale)	159	.46
	MMPI-2 APS (Addiction Potential Scale)	159	.45
Gavin et al. (1989)	DSM-III Drug Addiction Diagnosis (Current)	501	.75
	DSM-III Drug Addiction Diagnosis (Lifetime)	501	.74
	DSM-III Alcohol Addiction Diagnosis (Current)	501	-.31
	DSM-III Alcohol Addiction Diagnosis (Lifetime)	501	-.25
Hall et al. (2007)	Externalizing Inventory and Criterion Measure (100-item)	90	.61
Hormes et al. (2012)	Obsessive Compulsive Cocaine Use Scale (OCCUS) Total	107	.42
	OCCUS - Obsessive	107	.27
	OCCUS – Compulsive	107	.47
Reviere et al. (2007)	Michigan Alcoholism Screening Test (10-item)	200	.54
Teslak (2000)	CAGE Drug Test (CAGE D)	70	.52
	SASSI-2 Face Valid Other Drugs (FVOD)	70	.49
Weber (2008)	AUDIT (10-item)	824	.39

Three articles provided convergent validity data on the DAST-20 and two other forms of the DAST. Aggregated correlations between the DAST-20 and DAST-10 are summarized in Table 9. Skinner (1984) produced a correlation with the DAST (28-items) with a sample size of 223 of $r = .99[.86, 1.00]$. Two articles (see Table 9) produced correlations with the DAST-10 ($j = 2$; $k = 2$) with a total sample size of 320. After converting r values to z-values, weighting by corresponding sample sizes, and converting the aggregated z-score back to an r value, the averaged correlation coefficient for the DAST-20 and DAST-10 was $r = .98[.87, 1.00]$.

Table 9 – Convergent Validity between DAST-20 and DAST-10

Article	N	r	*z*-value	*z* x n
Cocco & Carey (1998)	97	.97	2.092	202.924
Skinner (1984)	223	.98	2.298	512.454
Total	320			715.378

Structural Validity

Three articles provided evidence of structural validity, including one confirmatory factor analysis (Cocco & Carey, 1998) and three exploratory factor analyses (Cocco & Carey, 1998; Saltstone et al., 1994; Skinner & Goldberg, 1986). Varying findings are reported below and implications discussed in Chapter 5.

Confirmatory Factor Analysis

Cocco and Carey (1998) performed a confirmatory factor analysis on the DAST-20 administered to a sample of only 100 participants, resulting in a poor fit for the unidimensional model ($X^2 = 473.23$, $p < .001$; $GFI = .72$; $AGFI = .65$). With only 100 participants, this CFA was significantly underpowered.

Exploratory Factor Analysis

Cocco and Carey (1998) reported a possible six-factor solution for the DAST-20. However, after reviewing the factor loadings, in which the first factor accounted for 41% of the variance and factors 2 through 6 accounted for an additional combined total of 30% of the variance, the researchers concluded that a two-factor solution best represented the scale. Cocco and Carey (1998) reported factor 1 was composed of items that inquire about "external consequences experienced by drug use," which included all items except for items 4 and 5 (factor loadings of .47 to .77) (p. 411). Factor 2 was comprised of items 4 (factor loading = .73) and 5 (factor loading = .55). These two items "had the lowest item-scale correlations" (p. 411). However, with only 97 participants, this EFA was significantly underpowered.

Saltstone et al. (1994) performed an exploratory factor analysis of DAST-20 scores from 615 female participants. This analysis resulted in a five-factor solution; however, due to a small variance (.14) between the last two factors, Saltstone et al. determined the four-factor solution was most parsimonious. A second factor analysis was performed for four factors and accounted for 56% of the total variance. The first factor alone accounted for 32.5% of the variance. Factor 1 was composed of items (4, 5, 7-11, 15, 16, 19, & 20) that inquire about drug abuse habits, social consequences, medical consequences, and previous treatment. Factor 2 was composed of items (12 & 13) related to work consequences. Similar to factor one, factor three comprised of items (2, 3, 6, 14, 17 & 18) that inquire about drug abuse habits, medical consequences, and "aggression," however different items loaded on factor 3 than on factor 1. Factor 4 included a single item (1) that inquired about an individual's drug dependence; single items do not

constitute a factor. In addition, items 4, 5, 17, & 18 displayed significant cross loading between two factors.

The third article (Skinner & Goldberg, 1986) that explored structural validity reported data from an exploratory factor analysis of the DAST-20 administered to 105 participants. The analysis resulted in a five-factor solution. Factor 1 was composed of three items (items 4, 5, and 17) that inquire about an individual's drug dependence. Factor 2 was composed of six items (items 8 through 13) that inquire about social consequences. Two items loaded on factor 3 (items 6 and 18), both of which inquire about medical consequences related to drug abuse. Factor 4 included five items (items 1 through 3, 14, and 16) inquiring about one's drug abuse habits and use, as well as legal consequences. The two items (14 and 16) that addressed legal consequences did not load as highly (.54 and .42, respectively). Finally, factor 5 had two items (items 19 and 20) load that inquire about previous treatment. With only 105 participants, this EFA was significantly underpowered.

Diagnostic Validity

A total of six articles reported diagnostic validity data for the DAST-20. Only two of these articles could be aggregated because of diverse criterion measures and cutoff scores. Therefore, a summary of the remaining articles is provided. Five of the articles reported data for similar cutoff scores (see Table 10).

Table 10 – Diagnostic Validity Data with Comparative Cut-off Scores

Studies	n	Cutoff	% CC	Sens	Spec	PPV	NPV	Criterion
Alexander & Leung (2011)	174	≥6	79.1	.818	.608	__	__	DSM-G-CS 21
Cassidy et al. (2008)	84	6	71	.55	.86	.79	.68	SCID
Cocco & Carey (1998)	97	5/6	81	.74	.83	__	__	DSM-HI-R
Gavin et al. (1989)	501	5/6	85	.96	.79	.73	.97	DSM-III

Møller & Linaker (2010)	48	≥5	__	.86	.67	__	__	ICD-10
Peters et al. (2000)	306	>6	82.7	.882	.805	.636	.947	SCID-IV

Alexander and Leung (2011; *n* = 174) reported a sensitivity of .818, a specificity of .608, and 79.1% correctly classified for a cutoff score of 6 or greater. The criterion used was the DSM-G-CS 21. Møller and Linaker (2010; *n* = 48) reported a sensitivity value of .86 and a specificity of .67 for a cutoff score of 5 or greater. Cassidy et al. (2007; *n* = 84) had optimal results at a cutoff score of 3 using the SCID as the criterion measure. At this cutoff score, the sensitivity value was .85, the specificity value was .73, the PPV was .74, and the NPV was .84, with 79% of participants correctly identified. However, with the traditionally recommended cutoff score of 6, Cassidy et al. (2007) reported a sensitivity of .55, a specificity of .86, a PPV of .79, a NPV of .68 and 71% of participants were correctly classified. Coco and Carey (1998) reported diagnostic validity data on a sample of 97 individuals using the DSM-HI-R criterion for drug use disorder. Optimal sensitivity and specificity results were found with a cutoff score of 2/3 (sensitivity = .89, specificity = .68, %CC = 72) and 5/6 (sensitivity = .74, specificity = .83, %CC = 81). Using the DSM-III as the criterion measure, Gavin, et al. (1989; *n* = 501) found that the percent of individuals correctly identified was 85 between the cutoff scores of 5/6 and 9/10. Sensitivity and specificity values were closest (.88 and .84, respectively) at a cutoff score of 6/7. Finally, Peters et al. (2000; *n* = 306) reported diagnostic validity for the DAST-20 for a cutoff score of 6 or greater using the SCID-IV as the criterion measure. The sensitivity was .882, specificity .805, PPV .636, NPV .947, and 82.7% of individuals were correctly classified.

An aggregation of Coco and Carey (1998) and Gavin et al.'s (1989) diagnostic validity data was completed (see Table 11) as the two studies reported cutoff scores in the same format and used a version of the DSM as the criterion.

Table 11 – Aggregated Diagnostic Validity Data for Cocco and Carey (1998) and Gavin et al. (1989) (*n* = 598, *j* =2)

Cutoff Score	Sensitivity	Specificity	%CC
0/1	1.00	.385	59.1
1/2	.992	.542	69.9
2/3	.982	.663	77.9
3/4	.974	.723	81.2
4/5	.957	.765	83.3
5/6	.924	.796	84.4
6/7	.881	.820	84.7
7/8	.831	.850	84.7

Descriptive Statistics

Means and standard deviations reported for nonclinical samples were extracted and aggregated. A total of 12 articles (*j* = 12) included in the current study provided data for nonclinical samples, with six articles providing data on multiple nonclinical groups (*k* = 19). The combined total sample size was 2,617. Each mean and standard deviation was weighted by sample size and aggregated. The total average mean is 1.083 and the total average standard deviation is 1.620. These results are summarized in Table 12.

Table 12 - DAST-20 Descriptive Statistics of Nonclinical Samples

Article	N	M	N x M	SD	N x SD
Aubry et al. (2012)	89	0.0	0.0	0.0	0.0
Burnett et al. (2013) sample a	552	2.1	1,159.2	2.8	1,545.6
Burnett et al. (2013) sample b	1,322	1.1	1,454.2	1.7	2,247.4
Ersche et al. (2010) sample a	30	0.0	0.0	0.0	0.0
Ersche et al. (2010) sample b	30	0.4	12.0	1.0	30.0
Ersche et al. (2012) sample a	50	0.0	0.0	0.0	0.0
Ersche et al. (2012) sample b	50	0.5	25.0	1.1	55.0
Ersche et al. (2013) sample a	52	0.0	0.0	0.0	0.0

Ersche et al. (2013) sample b	50	0.5	25.0	1.1	55.0
Fineran et al. (2010) sample a	79	0.23	18.17	0.97	76.63
Fineran et al. (2010) sample b	79	0.34	26.86	0.85	67.15
Fineran et al. (2010) sample c	7	0.14	0.98	0.38	2.66
Gizewski et al. (2013)	12	0.8	9.6	0.8	9.6
Levy (2013)	22	1.5	33.0	1.4	30.8
Morein-Zamir et al. (2013) sample a	39	.36	14.04	.74	28.86
Morein-Zamir et al. (2013) sample b	41	0.0	0.0	0.0	0.0
Schiffer et al. (2010)	14	0.9	12.6	0.9	12.6
St. Germain & Hooley (2012)	68	0.4	27.2	0.7	47.6
Weinborn et al. (2011)	31	0.5	15.5	1.0	31.0
Total	2,617		2833.35		4,239.9

Chapter 4 Summary

In this chapter, the results of the article search and selection were reported. Further, the psychometric properties (i.e., internal consistency, test-retest reliability, convergent validity, structural validity, diagnostic validity, descriptive statistics) of the DAST-20 (Skinner, 1982a) were summarized. In Chapter 5, the results will be interpreted and I will discuss the significance of the findings, limitations of the study, and implications for future research.

CHAPTER 5: DISCUSSION

Introduction

The current study examined the psychometric properties of the DAST-20 (Skinner, 1982a) following psychometric synthesis procedures outlined in Chapter 3. Various factors in research studies can contribute to variations in reported psychometric properties of the instrument, which can have an impact on how clinicians and researchers use and interpret the results of the DAST-20 (Erford et al., 2010; Lipsey & Wilson, 2001). This is significant because researchers and clinicians are required to provide care that is supported by research and ethical guidelines, including examining the psychometric properties and empirical data of the instrument and considering multicultural facets when using the instrument (ACA, 2014). This study aimed to provide an overview and aggregated results of all relevant data on the DAST-20 in order to provide a more comprehensive and clear picture of the psychometric properties of the instrument.

This psychometric synthesis of the DAST-20 included data from 56 articles and provided aggregated results of internal consistency, test-retest reliability, convergent validity, structural validity, diagnostic validity, and descriptive statistics of the instrument. Overall, the aggregated data produced adequate to excellent score reliability coefficient values for a screening instrument. The validity data was more diverse, with varying structural analysis reports and reported cutoff scores when examining diagnostic validity. In this chapter, interpretation of the results from Chapter 4 and implications will be discussed in detail. Then, limitations of the study, recommendations for clinicians, and recommendations for future research will be provided.

Research Questions

Three research questions formed the basis of this study:

1. What are the aggregated psychometric properties of the DAST-20 scores across published studies (i.e., internal consistency, test-retest reliability, diagnostic validity, internal structural validity, external convergent validity)?
2. What are the mean scores and standard deviations of the DAST-20 in nonclinical samples across published studies?
3. Are there significant differences in the psychometric properties of the DAST-20 scores among various sample characteristics (e.g., gender)?

Discussion of Results

The first research question inquired about the aggregated psychometric properties of the DAST-20 scores, including internal consistency, test-retest reliability, diagnostic validity, structural validity, and convergent validity. When determining the efficacy of internal consistency estimates, Erford (2021) recommended values of at least .80 to consider scores from an instrument acceptable for screening purposes and at least .90 for use as a diagnostic tool. From the initial psychometric examination of the DAST-20 scores, Skinner (1984) reported an excellent internal consistency for the total sample and a good internal consistency for a subsample that omitted participants reporting only alcohol abuse. Both of these estimates indicate that the DAST-20 is adequate for screening purposes, although, the subsample size is small. The internal consistency estimate of the subsample, which included individuals who reported only drug abuse or drug abuse and alcohol abuse and omitted those who reported only alcohol abuse is likely to be more similar to the current sample because the DAST-20 instructions specify that

the items do not inquire about alcohol use. The aggregated internal consistency value from the current study was slightly lower than the initial subsample, although still indicative of an adequate reliability for a screening instrument.

Of the studies included in this study that reported internal consistency values, there was a disparity between sample size and *KR-20* values. Samples that were over 1,000 participants had lower internal consistency estimates than the studies with a smaller sample size. All of these larger studies, except for one consisted of a nonclinical sample. Therefore, the lower internal consistency scores among these studies can be attributed to homogeneity of scores, as nonclinical samples tend to have a very large percentage of low scores with many individuals scoring 0 on the instrument.

Out of the 33 articles reporting internal consistency data, seven articles provided data for female participants (see Table 2) and seven articles provided data for male participants (see Table 3). The aggregated data was consistent among males and females and provided evidence of excellent internal consistency for females and males scores. The internal consistency estimates for gender groups were significantly higher than for the total sample, but they included a much smaller sample size.

The sample also was divided into clinical and nonclinical subsamples in order to assess internal consistency estimate differences between the two groups. The aggregated internal consistency score estimate for the clinical subsample was slightly higher than the estimate for the overall sample (see Table 5). The internal consistency estimate calculated for the nonclinical subsample was significantly lower (see Table 4). This estimate falls just below the cutoff considered adequate for screening tests.

Only two articles reported evidence of test-retest reliability for the DAST-20 scores and each of the studies reported different types of evidence and varying timeframes of test administration, which did not allow for aggregation of the data. Peters et al. (2000) reported excellent test-retest score reliability over a period of 72 hours. Although the sample was small, the results indicate more than adequate test-retest reliability of the DAST-20 scores for screening purposes over a 72-hour timeframe. Cocco and Carey (1998) reported evidence of lower test-retest reliability. Again, the sample size was under 100 participants and the timeframe between test administrations was inconsistent in that study (readministering the test anywhere between 7 to 43 days for each participant). The inconsistency of test administration could have had an effect on the reported reliability. Further, the longer gap between test administrations could contribute to the lower reliability scores (Bressler et al., 2018). Finally, Conner et al. (2004) reported interrater reliability between the participant's self-report of drug abuse and their significant other's report of the participant's drug abuse using the DAST-20. The sample size of this study is smaller than what is considered ideal.

Evidence of convergent validity among scores from included studies ranged from -.13 to .75 depending on the instrument being correlated with the DAST-20. The MAST (25-items; see Table 6) and the ADS (see Table 7) were the only two instruments with multiple convergent validity data reported in multiple studies. The aggregated convergent validity estimate for both the MAST and the ADS indicated small correlations. Both convergent instruments measure problems related to alcohol use, whereas the DAST-20 excludes alcohol when assessing drug abuse. Among the other 27 instruments (see Table 8) with one source of data of convergent validity with the DAST-20 scores, six

instruments or subscales measured alcohol use, including the Alcohol Composite Score, the MMPI-2 MAC-R, the DSM-III Alcohol Addiction Diagnosis current and lifetime, the MAST (10-item), and the AUDIT (10-item). All of the convergent validity estimates indicated medium to large correlations (effect sizes) between the DAST-20 and each instrument scores.

These estimates may support the ability of the DAST-20 scores to discriminate between alcohol abuse and drug abuse during the screening process. Medium to strong correlation values were reported for the DAST-20 and the Marijuana Screening Inventory (MSI-X), the Addiction Severity Index - Recent Drug Problems (ASI-5) and Lifetime Drug Problems, Addiction Severity Index – Drug Composite Score, Clinician Rating Scales for Drug Use, DSM-III Drug Addiction Diagnosis - Current and Lifetime, Obsessive Compulsive Cocaine Use Scale – Total, the CAGE Drug Test, and the SASSI-2 Face Valid Other Drugs. These estimates support convergent validity between the DAST-20 scores and other instruments measuring drug abuse, yielding medium to large effect sizes (Cohen, 1988).

Cocco and Carey (1998) reported a negative, small convergent validity estimate between the DAST-20 and the Global Assessment of Functioning (GAF). As discussed in Chapter 2, drug abuse has been shown to result in negative consequences in various aspects of an individual's life. Therefore, it seems logical that as scores increase on the DAST-20 (indicating more consequences of drug abuse) then scores on the GAF would decrease (indicating lower levels of functioning in various aspects of life).

Finally, multiple studies reported evidence of convergent validity between the DAST-20 and other forms of the instrument. Skinner (1984) reported a strong correlation

between the DAST (28-items) and the DAST-20. Aggregated convergent validity data for the DAST-20 and the DAST-10 indicated a strong, positive correlation. These findings suggest that clinicians and researchers should chose the form of the instrument based on the recommended uses. For example, The DAST-20 is recommended for use in clinical settings and for research evaluation (Skinner, 1984).

Varying findings were reported for the structural validity of the DAST-20. Cocco and Carey (1998) was the only study that performed a confirmatory factor analysis based on the original finding of a one-factor solution, with an undersized sample. EFA and CFA typically requires 10 participants per item to yield reliable results (Tabachnik & Fidell, 2019). At 100 participants for 20 items, the analyses were significantly underpowered. Cocco and Carey decided to run an exploratory factor analysis after the confirmatory factor analysis did not support a one-factor solution.

Three studies performed an exploratory factor analysis on the DAST-20 and all three studies came to different conclusions. Cocco and Carey (1998) reported a two-factor solution, Saltstone et al. (1994) reported a four-factor solution, and Skinner and Goldberg (1986) reported a five-factor solution. Differences in sample characteristics and the primarily small sample sizes could have affected the factor analysis results of the instrument. Two of these three studies were underpowered with sample sizes of 97 and 105 and each consisted of about 75% males and 25% females. The study conducted by Saltstone et al. (1994) was not underpowered, however the sample consisted of only females. It also appears that some of the items had significant cross-loading on multiple factors with less than .1 difference. This could affect the interpretation of the loadings.

Unfortunately, from the available data, a preliminary conclusion of an EFA factor solution of the DAST-20 cannot be made and further research is needed.

Diagnostic validity of the DAST-20 was analyzed, including examining percent correctly classified, sensitivity, specificity, positive predictive values, and negative predictive values. Adequate data was not reported in most of the articles reporting on diagnostic validity of the instrument. Skinner (1982a) originally recommended a cutoff score of 6 or greater to determine an individual's need for further evaluation of drug abuse, but also specified that more research needed to be completed. Multiple articles provided data for a cutoff of 5/6 or greater, with a range of 71 to 85 percent correctly classified. This range was used following the recommendation of Skinner (1982a) and did not provide further statistics for other cutoff scores. Therefore, it is not possible to see if there were more optimal cutoff scores for these studies. The six studies also used different criterion measures (i.e., various DSM versions, ICD). Due to the lack of data and differences in criteria, only two studies (Cocco & Carey, 1998; Gavin et al., 1989) provided enough data using similar criterion measures to be aggregated. The results of this aggregation is in agreement with Skinner's (1982a) original cutoff recommendation, with a maximum percent correctly classified of 84.7 at a cutoff score of 5/6 and higher. However, Cocco and Carey (1998) independently reported optimal sensitivity and specificity results at a cutoff of 2/3. The sample size of Cocco and Carey's study was significantly smaller than the second study accounting for only 97 of the 598 individuals in the study and, therefore had less impact on the overall findings. Cassidy et al. (2007) reported an optimal cut-off of 3. These lower cut-off scores are in agreement with the data collected for the second research question of this study.

The second research question inquired about the mean scores and standard deviations of the DAST-20 among nonclinical samples. The total average mean of nonclinical samples is 1.083 with an average standard deviation of 1.620 (see Table 10). In 2015, the NIH reported that four percent of adults in the United States met the criteria for a diagnosis of drug abuse disorder within the last year. Using this statistic, the DAST-20 cut-off score at the 96th percentile is about 3.9. This lower cut-off score is more congruent with the studies mentioned above. More research is needed to examine optimal cut-off scores, but the tentative conclusion is that a cutoff score of 6, as previously recommended by Skinner (1982a) is too high to effectively and accurately identify all individuals who need further assessment for drug abuse. A cutoff score of four appears to be a more acceptable cut-off score, leading to fewer false negatives when screening for substance abuse.

Finally, the third research question inquired about the presence of significant differences in psychometric properties of the DAST-20 among various sample characteristics. Due to a lack of available data reported in the accepted publications, gender was the only sample characteristic that could be examined. The aggregated data for internal consistency was consistent among males and females and provided evidence of excellent internal consistency for females and males. Only a fraction of the included articles that reported on internal consistency provided scores for males and females, which contributes to the higher scores compared to the overall internal consistency estimates. Unfortunately, none of the studies that included nonclinical samples provided mean and standard deviations values for only males or females.

Limitations

As mentioned in Chapter 1, there were several delimitations prior to conducting the study that should be noted. One category of delimitations concerns data available during the time of the study. Only articles that were published in English and used the English version on the DAST-20 were considered for inclusion in this study. Therefore, potential relevant data from studies published in other languages could not be included (Lipsey & Wilson, 2001). Similarly, this psychometric synthesis only included articles that reported data from the 20-item version of the DAST, meaning that the results of this study should only be used to guide one's use of the English version of the DAST-20. Further, this study is limited by the available data provided by included studies. Although the full text version of each potential study was examined for relevant psychometric data on the DAST-20, we were limited by what was included in the published articles. Therefore, without access to more detailed records, relevant data that could affect the overall aggregated results of the current study are potentially missing. The current study only includes articles that were published between 1982 (the year the DAST was published) and 2014 (the year that the current study was originally started). Seven years of recent studies need to be searched and included before publication of the results.

The second category of delimitations include the general limitations and criticisms of meta-analyses and similar types of studies such as this study (Erford et al., 2015; Lipsey & Wilson, 2001). When completing a synthesis or meta-analysis, specific factors of the individual studies may become lost or unnoticed in the overall analysis. Critics of synthesized methodologies argue that these factors, such as limitations, cultural

implications, theoretical frameworks, and study quality, can have important implications that are not always able to be highlighted and examined in a large aggregation of studies.

A further limitation of this study was the lack of psychometric data reported for specific groups, such as race, ethnicity, different age groups, and so on. Therefore, data aggregation was not possible for diverse groups. Gender (male/female) was the only group that had sufficient reported data to complete analyses, but this was still somewhat limited. Mean and standard deviations of nonclinical samples were not available across gender, which impeded the ability to assess gender score differences.

Validity data in general was very limited for the DAST-20. There were only three available studies that reported evidence of EFA and one study that reported evidence of CFA. Evidence of diagnostic validity was only provided in six studies and there were only two instruments that had multiple convergent validity values with the DAST-20 that could be aggregated. All of the other instruments that had convergent validity scores with the DAST were only reported in one study. There were also no measurement invariance studies available to examine item response differences among various groups. The lack of data made it difficult to examine, interpret, and draw conclusions about many aspects of the reliability and validity estimates of the DAST-20.

Implications for Professional Counselors

After examination of the results of this study, several implications are apparent for clinical practice, test administration and interpretation for clinicians and researchers, as well as recommended future research to shed further light on the use of the DAST-20. Implications of instrument use and recommendations are provided in this section.

Overall, the DAST-20 seems to yield reliable and valid scores to screen for the presence of consequences related to an individual's drug abuse, however, there is need for more research on the psychometric properties of the DAST-20. Recommendations for professional counselors are provided below.

As previously mentioned, there was limited data available to compare the DAST-20 psychometric properties among various cultural groups (i.e., race, ethnicity, age, gender). Therefore, counselors should consider multicultural factors that may impact the results of the instrument until more research can be conducted to examine the use of the DAST-20 with various groups.

Clinicians should only use the DAST-20 as a screening instrument and not as a diagnostic tool. Further, Skinner (1982a) indicated that the instrument is not intended to provide a comprehensive assessment of all aspects of drug abuse. Therefore, counselors should use this instrument in conjunction with other sources of data collection in relation to a client's drug abuse. One area of drug abuse consequences not adequately assessed by the DAST-20 are psychological consequences. In the literature, many psychological consequences have been cited related to an individual's use of drugs, however, the DAST-20 does not directly assess the presence of these consequences. Other screening instruments should be used to get a more complete picture of how drug abuse is affecting an individual's psychological health and vice versa.

Instruments that rely solely on self-report can be at risk for underreporting of issues by a client (Erford, 2021). Skinner (1982a) indicated the risk of underreporting on the DAST. In the current study, no correlations were reported between social desirability scales and the DAST-20. Underreporting of drug abuse can lead to misidentification for

further assessment and treatment, as well as inadequate treatment plans. Therefore, it is recommended that professional counselors use a social desirability instrument in conjunction with the DAST-20 to provide evidence of whether a client may be underreporting issues related to drug abuse.

Significant differences are apparent between the aggregated psychometric estimates found in this study and the estimates reported in the original study of the DAST-20. The internal consistency estimate from this study was much lower than the original internal consistency estimate provided by Skinner (1982a). This is important for professional counselors to consider, however, the aggregated internal consistency still falls in the range considered appropriate for use as a screening instrument. More concerning, the analysis of the aggregated data in this study indicated that the original cutoff score recommended by Skinner (1982a) might be too high, leading to potential false negatives. This means that a significant number of individuals who should be identified for further diagnostic evaluation could potentially be missed. Further research is needed to come to a more definitive conclusion about the most effective cutoff score. Until this data is available, professionals should consider scores in the range of three to six. Other instruments and forms of data collection could be used in conjunction with the DAST-20 when individuals score in this range to help determine whether further diagnostic assessment is recommended.

Implications for Counseling Research

To further examine evidence of reliability and validity of the DAST-20 scores, more research is needed. In regard to limitations of the current study, a search for all relevant published articles from 2015 to 2021 should be acquired and examined for data that can

be included in the psychometric synthesis. A future study that examines reliability and validity of the DAST-20 scores in studies published in languages other than English could provide important multicultural facets of using and interpreting the instrument. It would also provide a more comprehensive examination of the psychometric properties of the DAST-20 scores than what was accomplished with the current study. Also, other psychometric synthesis studies conducted on the other forms of the DAST, including the original 28-item DAST, the DAST-10, and the DAST-A are recommended. All of the instruments have been reported to be highly correlated, which would provide further evidence of validity, as well as implications for the appropriate use and settings for each instrument.

After completing the data analysis, it was apparent that further research was needed to fill gaps in the data and examine conflicting data between some of the studies. As previously mentioned, data for individual groups were not readily available in the studies included in this psychometric synthesis. Due to the lack of data, multicultural implications cannot be identified with confidence. Although the DAST-20, overall, yields valid and reliable scores as a screening instrument for drug abuse, it is important that cultural factors are studied to make sure that it is sufficient to use among various populations. The ACA (2014) *Code of Ethics* specifies the responsibility of clinicians and researchers to consider possible implications of using specific instruments without considering cultural factors (i.e., "age, color, culture, disability, ethnic group, gender, race, language preference, religion, spirituality, sexual orientation, and socioeconomic status") (p. 11). It is also imperative that authors publish data for individual groups in future research articles in order for more psychometric syntheses such as the current

study to have the data necessary to investigate potential differences and how they affect outcomes. Further, professional journal editors should require more detailed data be reported in accepted research articles related to multicultural implications and differences among groups.

Measurement invariance was not assessed in any of the articles. Measurement invariance examines how individuals across diverse backgrounds respond to items on an instrument and how any differences can impact the way the items are interpreted. For example, do men and women interpret the items in the same way and therefore, do the results have the same meaning across gender? Not only does more validity and reliability data need to be collected across various groups, measurement invariance of the DAST-20 items also should be studied.

Conflicting evidence of structural validity was reported among studies. Some of these studies were underpowered. Future psychometric studies of the DAST-20 should include larger sample sizes of at least 500 participants. A future confirmatory factor analysis study using a community sample of at least 500 participants is recommended. Except for the original study of the DAST-20 completed by Skinner (1982) when creating the screening instrument, only one other study found provided evidence from a confirmatory factor analysis. There were only three studies included in this psychometric synthesis that conducted exploratory factor analyses on the DAST-20, resulting in three different conclusions. Also, updated procedures for interpretation of EFA have been established (Tabachnik & Fidell, 2019). Therefore, more EFA studies need to be conducted using the updated procedures to produce a more accurate picture of factor loadings.

One of the issues in comparing findings from various studies was the lack of consistent criterion measures. The DSM-III, DSM-IV, and the ICD were all used as criterion measures for diagnostic validity, making it difficult to aggregate data and make more definitive conclusions about cutoff scores, as well as other psychometric properties. In regard to the DSM, future research should examine the validity of DAST-20 scores when using the DSM-V as the criterion measure, as it is the most recent version of the manual and is widely used for substance abuse evaluation. It must be noted that these studies might already exist, as the current study only accepted articles up to 2014. This data should be added to the current analysis for an updated examination of the DAST-20. Only six articles included in the psychometric synthesis reported data on the diagnostic validity of the DAST-20 and recommended multiple different cutoff scores to use when screening for drug abuse. More studies are needed on the diagnostic validity of the DAST-20 in order to gain greater agreement on cutoff scores among varying groups and to assess the instrument's use as part of a diagnostic protocol.

Finally, as mentioned in the recommendation section for professional counselors, none of the studies included in the current study reported correlation estimates for the DAST-20 and a social desirability scale. Skinner (1982a) acknowledged that the DAST does not prevent or detect underreporting of substance abuse. Therefore, studies examining the use of the DAST-20 and social desirability scales could shed light on how underreporting affects DAST-20 scores, especially diagnostic validity and how these instruments should be used together during evaluation.

Conclusion

This chapter discussed the interpretation of the results of the psychometric synthesis of the DAST-20 and the significance of the findings. Findings conclude that counselors should use the instrument with clients with confidence, as the overall reliability and validity estimates are adequate. However, some of the data was limited and further research is recommended to continue to examine the DAST-20, such as multicultural implications. Limitations of the study were discussed and recommendations for professional counselors and researchers were provided.

REFERENCES

* = Article accepted into the synthesis.

*Alexander, D., & Leung, P. (2006). The Marijuana Screening Inventory (MSI-X): Concurrent, convergent and discriminant validity with multiple measures. *The American Journal of Drug and Alcohol Abuse, 32*(3), 351-378. https://doi.org/10.1080/00952990600753594

*Alexander, D., & Leung, P. (2011). The DSM Guided Cannabis Screen (DSM-G-CS): Description, reliability, factor structure and empirical scoring with a clinical sample. *Addictive Behaviors, 36*(11), 1095-1100. https://doi.org/10.1016/j.addbeh.2011.06.008

*Alterman, A. I., Cacciola, J. S., Habing, B., & Lynch, K. G. (2007). Addiction Severity Index recent and lifetime summary indexes based on nonparametric item response theory methods. *Psychological Assessment, 19*(1), 119-132. https://doi.org/10.1037/1040-3590.19.1.119

American Counseling Association. (ACA). (2014). *ACA code of ethics.* https://www.counseling.org/resources/aca-code-of-ethics.pdf

American Psychiatric Association. (APA). (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.; DSM-5). American Psychiatric Press.

*Aubry, T., Klodawsky, F., & Coulombe, D. (2011). Comparing the housing trajectories of different classes within a diverse homeless population. *American Journal of Community Psychology, 49*(1-2), 142-155. https://doi.org/10.1007/s10464-011-9444-z

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the Beck Depression*

Inventory (2nd ed.). Psychological Corporation.

*Bliss, M. J., Ogley-Oliver, E., Jackson, E., Harp, S., & Kaslow, N. J. (2008). African American women's readiness to change abusive relationships. *Journal of Family Violence, 23*(3), 161-171. https://doi.org/10.1007/s10896-007-9138-3

*Boyd, M. R., Berger, K. C., Baliko, B., & Tavakoli, A. (2009). Predictors of alcohol and drug problems in rural African American women. *Issues in Mental Health Nursing, 30*(6), 383-391. https://doi.org/10.1080/01612840802509478

Bressler, R., Erford, B. T., & Dean, S. (2018). Systematic review of the Posttraumatic Stress Disorder Checklist (PCL). *Journal of Counseling & Development, 96*(2), 167-187. https://doi.org/10.1002/jcad.12190

*Brocato, J., & Wagner, E. F. (2008). Predictors of retention in an alternative-to-prison substance abuse treatment program. *Criminal Justice and Behavior, 35*(1), 99-119. https://doi.org/10.1177/0093854807309429

Bronson, J., & Carson, E. A. (2019). *Prisoners in 2017*. https://www.bjs.gov/content/pub/pdf/p17.pdf

*Burnett, A. J., Sabato, T. M., Walter, K. O., Kerr, D. L., Wagner, L., & Smith, A. (2013). The influence of attributional style on substance use and risky sexual behavior among college students. *College Student Journal, 47*(2), 122-136.

*Cassidy, C. M., Schmitz, N., & Malla, A. (2008). Validation of the Alcohol Use Disorders Identification Test and Drug Abuse Screening Test in first episode psychosis. *Canadian Journal of Psychiatry, 53*(1), 26-33. https://doi.org/10.1177/070674370805300105

*Cocco, K. M., & Carey, K. B. (1998). Psychometric properties of the Drug Abuse

Screening Test in psychiatric outpatients. *Psychological Assessment, 10*(4), 408-414. https://doi.org/10.1037/1040-3590.10.4.408

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.) Hillsdale, NJ: Erlbaum.

*Conner, K. R., Yueying, Z., & Duberstein, P. R. (2004). NEO-PI-R neuroticism scores in substance-dependent outpatients: Internal consistency and self-partner agreement. *Journal of Personality Assessment, 83*(1), 75-77. https://doi.org/10.1207/s15327752jpa8301_07

Cook, D. J., Mulrow, C. D., & Haynes, R. B. (1997). Systematic reviews: Synthesis of best evidence for clinical decisions. *Annals of Internal Medicine, 126*(5), 376-380. https://doi.org/10.7326/0003-4819-126-5-199703010-00006

Erford, B. T. (2021). *Assessment for counselors*. Pearson.

Erford, B. T., Johnson, E., & Bardoshi, G. (2015). Meta-analysis of the English version of the Beck Depression Inventory-second edition. *Measurement and Evaluation in Counseling and Development*, *49*(1), 1-31. https://doi.org/10.1177/0748175615596783

Erford, B. T., Savin-Murphy, J. A., & Butler, C. (2010). Conducting a meta-analysis of counseling outcome research: Twelve steps and practical procedures. *Counseling Outcome Research and Evaluation, 1*(1), 19-43. https://doi.org/10.1177/2150137809356682

*Ersche, K. D., Jones, P. S., Williams, G. B., Smith, D. G., Bullmore, E. T., & Robbins, T. W. (2013). Distinctive personality traits and neural correlates associated with stimulant drug use versus familial risk of stimulant dependence. *Biological*

Psychiatry, 74(2), 137-144. http://doi.org/10.1016/j.biopsych.2012.11.016

*Ersche, K. D., Turton, A. J., Chamberlain, S. R., Müller, U., Bullmore, E. T., & Robbins, T. W. (2012). Cognitive dysfunction and anxious-impulsive personality traits are endophenotypes for drug dependence. *American Journal of Psychiatry, 169*(9), 926-936.

*Ersche, K. D., Turton, A. J., Pradhan, S., Bullmore, E. T., & Robbins, T. W. (2010). Drug addiction endophenotypes: Impulsive versus sensation-seeking personality traits. *Biological Psychiatry, 68*(8), 770-773. https://doi.org/10.1016/j.biopsych.2010.06.015

*Fineran, K., Laux, J. M., & Seymour, J. (2010). The Barnum effect and chaos theory: Exploring college student ACOA traits. *Journal of College Student Psychotherapy, 24*(1), 17-31. https://doi.org/10.1080/87568220903400096

*Fleury, M., Grenier, G., Bamvita, J., Perreault, M., & Caron, J. (2012). Determinants associated with the utilization of primary and specialized mental health services. *Psychiatric Quarterly, 83*(1), 41-51. https://doi.org/10.1007/s11126-011-9181-3

*Forbey, J. D., & Ben-Porath, Y. S. (2008). Empirical correlates of the MMPI-2 Restructured Clinical (RC) Scales in a nonclinical setting. *Journal of Personality Assessment, 90*(2), 136-141. https://doi.org/10.1080/00223890701845161

*Forbey, J. D., Ben-Porath, Y. S., & Arbisi, P. A. (2011). The MMPI-2 computerized adaptative version (MMPI-2-CA) in a veterans administration medical outpatient facility. *American Psychological Association, 24*(3), 628-639. https://doi.org/10.1037/a0026509

*Forbey, J. D., Lee, T. T. C., Ben-Porath, Y. S., Arbisi, P. A., & Gartland, D. (2013).

Associations between MMPI-2-RF validity scale scores and extra-test measures of personality and psychopathology. *Assessment, 20*(4), 448-461. https://doi.org/10.1177/1073191113478154

Galaif, E. R., Newcomb, M. D., & Carmona, J. V. (2001). Prospective relationships between drug problems and work adjustment in a community sample of adults. *Journal of Applied Psychology, 86*(2), 337-350. https://doi.org/10.1037//0021-9010.86.2.337

*Gavin, D. R., Ross, H. E., & Skinner, H. A. (1989). Diagnostic validity of the Drug Abuse Screening Test in the assessment of DSM-III drug disorders. *British Journal of Addiction, 84*(3), 301-307. https://doi.org/10.1111/j.1360-0443.1989.tb03463.x

*Gerlock, A. A. (2004). Domestic violence and post-traumatic stress disorder severity for participants of a domestic violence rehabilitation program. *Military Medicine, 169*(6), 470-474. https://doi.org/10.7205/MILMED.169.6.470

*Gizewski, E. R., Müller, B. W., Scherbaum, N., Lieb, B., Forsting, M., Wiltfang, J., Leygraf, N., & Schiffer, B. (2013). The impact of alcohol dependence on social brain function. *Addiction Biology, 18,* 109-120. https://doi.org/10.1111/j.1369-1600.2012.00437.x

Gove, W. R., Geerken, M., & Hughes, M. (1979). Drug use and mental health among a representative national sample of young adults. *Social Force, 58*(2), 572-590. https://doi.org/10.1093/sf/58.2.572

Grekin, E. R., Svikis, D. S., Lam, P., Connors, V., LeBreton, J. M., Streiner, D. L.,

Smith, C., & Ondersma, S. J. (2010). Drug use during pregnancy: Validating the Drug Abuse Screening Test against physiological measures. *Psychology of Addictive Behaviors, 24*(4), 719-723. https://doi.org/10.1037/a0021741

*Hall, J. R., Bernat, E. M., & Patrick, C. J. (2007). Externalizing psychopathology and the error-related negativity. *Psychological Science, 18*(4), 326-333. https://doi.org/10.1111/j.1467-9280.2007.01899.x

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Academic Press.

*Hormes, J. M., Coffey, S. F., Drobes, D. J., & Saladin, M. E. (2012). The Obsessive Compulsive Cocaine Use Scale: Development and initial validation of a self-rated instrument for the quantification of thoughts about cocaine. *Drug & Alcohol Dependence, 120*(1-3), 250-254. https://doi.org/10.1016/j.drugalcdep.2011.07.024

*Irving, M., & Schweiger, A. (1991). Individual and family characteristics of middle class adolescents hospitalized for alcohol and other drug abuse. *British Journal of Addiction, 86*(11), 1435-1447.

*Kaslow, N. J., Jacobs, C. H., Young, S. L., & Cook, S. (2006). Suicidal behavior among low-income African American women: A comparison of first-time and repeat suicide attempters. *Journal of Black Psychology, 32*(3), 349-365. https://doi.org/10.1177/0095798406290459

Kiluk, B. D., Fitzmaurice, G. M., Strain, E. C., & Weiss, R. D. (2018). What defines a clinically meaningful outcome in the treatment of substance use disorders: Reductions in direct consequences of drug use or improvement in overall functioning? *Addiction, 114*(1), 9-15. https://doi.org/10.1111/add.14289

*Kimbrel, N. A., Steffen, L. E., Meyer, E. C., Kruse, M. I., Knight, J. A., Zimering, R.

T., & Gulliver, S. B. (2011). A revised measure of occupational stress for firefighters: Psychometric properties and relationship to posttraumatic stress disorder, depression, and substance abuse. *American Psychological Association, 8*(4), 294-306. https://doi.org/10.1037/0025845

Lange, W. R., White, N., & Robinson, N. R. (1992). Medical complications of substance abuse. *Substance Abuse, 92*(3), 205-214. https://doi.org/10.1080/00325481.1992.11701450

*Levy B. (2013). Autonomic nervous system arousal and cognitive functioning in bipolar disorder. *Bipolar Disorders, 15*, 70-79. https://doi.org/10.1111/bdi.12028

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Sage.

*Looman, J. & Abracen, J. (2013). The Static-99R: Are there really differences between the normative groups? *International Journal of Offender Therapy and Comparative Criminology, 57*(7), 888-907. https://doi.org/10.1177/0306624X12443657

Macleod, J. (2010). Drug-taking and its psychosocial consequences. In S. MacGregor (Ed.), *Responding to drug misuse: Research and policy priorities in health and social care* (pp. 25-39). Routledge.

Maly, R. C. (1993). Early recognition of chemical dependence. *Substance Abuse, 20*(1), 33-50.

*Marshall, L. E., & Marshall, W. L. (2006). Sexual addiction in incarcerated sexual offenders. *Sexual Addiction & Compulsivity, 13*(4), 377-390. https://doi.org/10.1080/10720160601011281

Martino, S., Grilo, C. M., & Fehon, D. C. (2000). Development of the Drug Abuse

Screening Test for Adolescents (DAST-A). *Addictive Behaviors, 25*(1), 57-70. https://doi.org/10.1016/S0306-4603(99)00030-1

Minnich, A., Erford, B. T., Bardhoshi, G., & Atalay, Z. (2018). Systematic review of the Michigan Alcohol Screening Test (MAST). *Journal of Counseling & Development, 96*(4), 335-344. https://doi.org/10.1002/j.1556-6676.2014.00000.x

*Møller, T., & Linaker, O. M. (2010). Using brief self-reports and clinician scales to screen for substance use disorders in psychotic patients. *Nordic Journal of Psychiatry, 64*(2), 130-135. https://doi.org/10.3109/08039480903274423

*Morein-Zamir, S., Jones, P. S., Bullmore, E. T., Robbins, T. W., & Ersche, K, D. (2013). Prefrontal hypoactivity associated with impaired inhibition in stimulant-dependent individuals but evidence for hyperactivation in their unaffected siblings. *Neuropsychopharmacology,* 38(10), 1945-1953.

* Mowbray, C. T., Bybee, D., Oyserman, D., MacFarlane, P., & Bowersox, N. (2006). Psychosocial outcomes for adult children of parents with severe mental illnesses: Demographic and clinical history predictors. *Health & Social Work, 31*(2), 99-108.

*Murphy, C. M., Taft, C. T., & Eckhardt, C. I. (2007). Anger problem profiles among partner violent men: Differences in clinical presentation and treatment outcome. *American Psychological Association, 54*(2), 189-200. https://doi.org/10.1037/0022-0167.54.2.189

National Drug Intelligence Center. (2011). *The economic impact of illicit drug use on American society*. U.S. Department of Justice.

National Institute of Health. (2015). *10 percent of US adults have drug use disorder*

at some point in their lives. http://www.nih.gov/news-events/news-releases/10-percent-us-adults-have-drug-use-disorder-some-point-their-lives

National Institute on Drug Abuse. (2018). *Screening and assessment tools chart.* https://www.drugabuse.gov/nidamed-medical-health-professionals/screening-tools-resources/chart-screening-tools

National Institute on Drug Abuse. (2021a). *Trends & statistics: Costs of substance abuse.* https://www.drugabuse.gov/drug-topics/trends-statistics/costs-substance-abuse#supplemental-references-for-economic-costs

National Institute on Drug Abuse. (2021b). *Trends & statistics: National survey of drug use and health*. https://www.drugabuse.gov/drug-topics/trends-statistics/national-drug-early-warning-system-ndews/national-survey-drug-use-health

* Nelson, L. D., Patrick, C. J., & Bernat, E. M. (2011). Operationalizing proneness to externalizing psychopathology as a multivariate psychophysiological phenotype. *Psychophysiology, 48*, 64-72. https://doi.org/10.1111/j.1469-8986.2010.01047.x

Newcomb, M. D. (1997). Psychosocial predictors and consequences of drug use: A developmental perspective within a prospective study. *Journal of Addictive Diseases, 16*(1), 51-89. https://doi.org/10.1300/J069v16n01_05

Newcomb, M. D., & Bentler, P. M. (Eds.). (1988). *Consequences of adolescent drug use: Impact on the lives of young adults*. Sage Publications.

Newcomb M. D., & Locke T. (2005). Health, social, and psychological consequences of drug use and abuse. In Z. Sloboda (Ed.), *Epidemiology of drug abuse* (pp. 45-59). Springer. https://doi.org/10.1007/0-387-24416-6_4

Newcomb, M. D., Vargas-Carmona, J., & Galaif, E. R. (1999). Psychological distress

among a community sample of adults: Predictors, consequences, or confound? *Journal of Community Psychology, 27*(4), 405-429. https://doi.org/10.1002/(SICI)1520-6629(199907)27:4<405::AID-JCOP4>3.0.CO;2-2

*Nochajski, T. H., Stasiewicz, P. R., & Patterson, D. A. (2013). Depression, readiness for change, and treatment among court-mandated DUI offenders. *Journal of Dual Diagnosis, 9*, 139-148. https://doi.org/10.1080/15504263.2013.779092

*Nordfjaern, T. (2013). Prevalence of substance use and mental distress among patients on waiting lists for substance use disorder treatment. *Journal of Psychoactive Drugs, 45*(3), 233-240. https://doi.org/10.1080/02791072.2013.805977

*Perepletchikova, F., Ansell, E., & Axelrod, S. (2012). Borderline personality disorder features and history of childhood maltreatment in mothers involved with child protective services. *Child Maltreatment, 17*(2), 182-190. https://doi.org/10.1177/1077559512448471

*Peters, R. H., Greenbaum, P. E., Steinberg, M. L., Carter, C. R., Ortiz, M. M., Fry, B. C., & Valle, S. K. (2000). Effectiveness of screening instruments in detecting substance use disorders among prisoners. *Journal of Substance Abuse Treatment, 18*, 349-358.

Piazza, N. J., Martin, M., & Dildine, R. J. (2000). Screening instruments for alcohol and other drug problems. *Journal of Mental Health Counseling, 22*(3), 218-227.

Plant, M. L., Plant, M. A., & Mason, W. (2002). Drinking, smoking and illicit drug use among British adults: Gender differences explored. *Journal of Substance Abuse, 7*(1), 24-33. https://doi.org/10.1080/14659890110110392

Rasmussen, S. (2000). *Addiction treatment: Theory and practice*. Sage Publications. http://doi.org/10.4135/9781452231877.n9

*Reviere, S. L., Farber, E. W., Twomey, H., Okun, A., Jackson, E., Zanville, H., & Kaslow, N. J. (2007). Intimate partner violence and suicidality in low-income African American women. *Violence Against Women, 13*(11), 1113-1129. https://doi.org/10.1177/1077801207307798

Robbins, C. (1989). Sex differences in psychosocial consequences of alcohol and drug abuse. *Journal of Health and Social Behavior, 30*(1), 117-130. https://doi.org/10.2307/2136917

*Rosenkranz, S. E., Henderson, J. L., Muller, R. T., & Goodman, I. R. (2012). Motivation and maltreatment history among youth entering substance abuse treatment. *Psychology of Addictive Behaviors, 26*(1), 171-177. https://doi.org/10.1037/a0023800

*Rosenkranz, S. E. Muller, R. T., & Henderson, J. L. (2012). Psychological maltreatment in relation to substance use problem severity among youth. *Child Abuse & Neglect, 36*(5), 438-448. https://doi.org/10.1016/j.chiabu.2012.01.005

Sabato, T. M., Burnett, A. J., Kerr, D. L., & Wagner, L. (2013). Examining behavioral and psychosocial predictors of antibody testing among college youth: Implications for HIV prevention education and testing. *American Journal of Sexuality Education, 8*(1-2), 56-72. https://doi.org/10.1080/15546128.2012.740893

Salehi, S. H., As'adi, K., Musavi, J., Ahrari, F., Nemazi, P., Kamranfar, B., Gaseminegad, K., Faramarzi, S., & Shoar, S. (2012). Assessment of substance

abuse in burn patients by using Drug Abuse Screening Test. *Acta Medica Iranica, 50*(4), 257-264.

*Saltstone, R., Halliwell, S., & Hayslip, M. A. (1994). A multivariate evaluation of the Michigan Alcoholism Screening Test and the Drug Abuse Screening Test in a female offender population. *Addictive Behaviors, 19*(5), 455-462.

*Schiffer, B., Müller, B. W., Scherbaum, N., Forsting, M., Wiltfang, J., Leygraf, N., & Gizewski, E. R. (2010). Impulsivity-related brain volume deficits in schizophrenia-addiction comorbidity. *Brain: A Journal of Neurology, 133*(10), 3093-3103. https://doi.org/10.1093/brain/awq153

Scholl, L., Seth, P., Kariisa, M., Wilson, N., & Baldwin, G. (2019). Drug and opioid-involved overdose deaths – United States, 2013-2017. *Morbidity and Mortality Weekly Report, 67*(5152), 1419-1427. http://doi.org/10.15585/mmwr.mm675152e1

*Scott, D., McGilloway, S., & Donnelly, M. (2009). The mental health needs of women detained in police custody. *Journal of Mental Health, 18*(2), 144-151. https://doi.org/10.1080/09638230701879193

Selzer, M. (1968). Michigan Alcoholism Screening Test (MAST): Preliminary report. *University of Michigan Medical Center Journal, 34*(3), 143-145.

Selzer, M. (1971). The Michigan Alcoholism Screening Test: The quest for a new diagnostic instrument. *The American Journal of Psychiatry, 127*(12), 1653-1658. https://doi.org/10.1176/ajp.127.12.1653

*Shuggi, R., Mann, R. E., Zalcman, R. F., Chipperfield, B., & Nochajski, T. (2006). Predictive validity of the RIASI: Alcohol and drug use and problems six months

following remedial program participation. *American Journal of Drug and Alcohol Abuse, 32*(1), 121-133. https://doi.org/10.1080/00952990500328711

Skinner, H. A. (1982a). *Guide for using the Drug Abuse Screening Test (DAST).* http://www.emcdda.europa.eu/html.cfm/index3618EN.html

Skinner, H. A. (1982b). The Drug Abuse Screening Test. *Addictive Behaviors, 7*(4), 363-371. https://doi.org/10.1016/0306-4603(82)90005-3

*Skinner, H. A. (1984). Instruments for assessing alcohol and drug problems. *Bulletin of the Society of Psychologists in Addictive Behaviors, 3*, 21-33.

*Skinner, H. A. & Goldberg, A. E. (1986). Evidence for a drug dependence syndrome among narcotic users. *British Journal of Addiction, 81*(4), 479-484.

*St. Germain, S. A. & Hooley, J. M. (2012). Direct and indirect forms of non-suicidal self-injury: Evidence for a distinction. *Psychiatry Research, 197*(1-2), 78-84. https://doi.org/10.1016/j.psychres.2011.12.050

*Stuart, G. L. & Holtzworth-Munroe, A. (2005). Testing a theoretical model of the relationship between impulsivity, mediating variables, and husband violence. *Journal of Family Violence, 20*(5), 291-303. https://doi.org/10.1007/s10896-005-6605-6

*Sullivan, T., Cavanaugh, C., Ufner, M., Swan, S., & Snow, D. (2009). Relationships among women's use of aggression, their victimization, and substance use problems: A test of the moderating effects of race/ethnicity. *Journal of Aggression, Maltreatment, & Trauma, 18*(6), 646-666. https://doi.org/10.1080/10926770903103263

Tabachnick, B. G., & Fidell, L. S. (2019). *Using multivariate statistics* (7th ed.). Pearson.

*Teslak, A. G. (2000). The utility of the CAGE, MAST, DAST and SASSI-3 in assessing substance use/misuse in a psychiatric population. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 61*(5-B), 2814.

*Ting, L., Jordan-Green, L., Murphy, C. M., & Pitts, S. C. (2009). Substance use problems, treatment engagement, and outcomes in partner violent men. *Research on Social Work Practice, 19*(4), 395-406. https://doi.org/10.1177/1049731508329418

U.S. Department of Health and Human Services (HHS), Office of the Surgeon General. (2018). *Facing addiction in America: The Surgeon General's spotlight on opioids*. https://addiction.surgeongeneral.gov/sites/default/files/Spotlight-on-Opioids_09192018.pdf

U.S. Preventative Services Task Force (USPSTF). (2020). Screening for unhealthy drug use: U.S. Preventative Services Task Force recommendation statement. *JAMA, 323*(22), 2301-2309. https://doi.org/10.1001/jama.2020.8020

Visser, L. (1991). Drug abuse: Consequences in terms of family pathology and disintegration. *Medicine and Law, 10*, 209-212.

Walsh, C., MacMillan, H. L., & Jamieson, E. (2003). The relationship between parental substance abuse and child maltreatment: Findings from the Ontario Health Supplement. *Child Abuse & Neglect, 27*(12), 1409-1425. https://doi.org/10.1016/j.chiabu.2003.07.002

*Weber, G. N. (2008). Using to numb the pain: Substance use and abuse among lesbian, gay, and bisexual individuals. *Journal of Mental Health Counseling, 30*(1), 31-48.

*Weinborn, M., Woods, S. P., Nulsen, C., & Park, K. (2011). Prospective memory

deficits in Ecstasy users: Effects of longer ongoing task delay interval. *Journal of Clinical & Experimental Neuropsychology, 33*(10), 1119-1128. https://doi.org/10.1080/13803395.2011.614595

*Weinstein, D. W. (1999). Posttraumatic stress disorder, dissociation and substance abuse as long-term sequelae in a population of adult children of substance abusers. *Dissertation Abstracts International Section A: Humanities and Social Sciences, 59*(7-A), 2723.

Wilson, N., Kariisa, M., Seth, P., Smith IV, H., & Davis, N. L. (2020). Drug and opioid-involved overdose deaths – United States, 2017-2018. *Morbidity & Mortality Weekly Report, 69*(11), 290-297. https://doi.org/10.15585/mmwr.mm6911a4