RELIABILITY AND VALIDITY EVIDENCE FOR THE DUAL-DISORDER TREATMENT FIDELITY SCALE

A THESIS SUBMITTED TO THE GRADUATE DIVISION OF THE UNIVERSITY OF HAWAI'I IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

IN

PSYCHOLOGY

AUGUST 2005

By

Diane C. Wilson

Thesis Committee:

Annette S. Crisanti, Co-chairperson
Charles W. Mueller, Co-chairperson
A. Michael Wylie
ACKNOWLEDGMENTS

Without cooperation from the Adult Mental Health Division (AMHD) at the Hawai‘i Department of Health this research would not have been possible. I would like to thank the six participating programs as well as the AMHD Mental Health Services Research and Evaluation Center for their support of this project.
ABSTRACT

Psychometric properties of the Dual-Disorder Treatment Fidelity Scale (Mueser, Noordsy, Drake & Fox, 2003) were evaluated. Reliability was assessed through measures of inter-rater agreement, and validity was assessed through concurrent, predictive, and construct validation procedures. Two raters assessed 3 community mental health center (CMHC) and 3 dual diagnosis enhanced (DDE) programs for fidelity. Fifty-two consumers and 58 staff completed an outcomes survey and a measure of knowledge about the model, respectively. Inter-rater agreement on fidelity ratings across programs, within each program, and for most scale items was good. Total fidelity scores did not distinguish between CMHC and DDE program types; however, a cluster analysis revealed concurrent validity based on the item level ratings. Fidelity scores did not predict consumer outcomes, and fidelity was unrelated to staff knowledge of the integrated model. Conclusions drawn from item level ratings appear valid; the validity of conclusions based on total fidelity scores remains unknown.
TABLE OF CONTENTS

ACKNOWLEDGMENTS ........................................................................................................ iii
ABSTRACT ........................................................................................................................ iv
LIST OF TABLES .............................................................................................................. vii
LIST OF FIGURES ........................................................................................................ viii
CHAPTER ONE .................................................................................................................. 1
   Introduction................................................................................................................... 1
CHAPTER TWO ................................................................................................................. 4
   Literature Review......................................................................................................... 4
      Terminology.............................................................................................................. 4
      Prevalence............................................................................................................... 5
      Consequences and Costs......................................................................................... 6
      Historical Context ................................................................................................. 7
      Integrated Treatment ............................................................................................. 8
CHAPTER THREE ........................................................................................................... 20
   Methods....................................................................................................................... 20
      Study Design ......................................................................................................... 20
      Research Questions ............................................................................................... 22
         Item Analysis ...................................................................................................... 22
         Reliability ........................................................................................................... 22
         Validity ............................................................................................................... 22
      Participants .......................................................................................................... 23
         Programs ............................................................................................................. 23
         Individuals ......................................................................................................... 24
      Instrumentation ..................................................................................................... 24
         Dual-Disorder Treatment Fidelity Scale .............................................................. 24
         Consumer Survey ............................................................................................... 26
         Staff Survey ....................................................................................................... 28
      Procedure .............................................................................................................. 28
      Data Analysis ........................................................................................................ 30
         Item analysis ...................................................................................................... 30
         Reliability ........................................................................................................... 30
         Validity ............................................................................................................... 30
      Ethical Considerations .......................................................................................... 33
CHAPTER FOUR ............................................................................................................. 35
   Results......................................................................................................................... 35
      Fidelity Assessments .............................................................................................. 35
      Validity Analyses ................................................................................................. 37
CHAPTER FIVE .................................................................................................................. 45
   Discussion .................................................................................................................... 45
      Item Analysis ......................................................................................................... 45
      Reliability ............................................................................................................... 46
      Validity ................................................................................................................... 47
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Number and Percent of Program Staff and Consumers Responding to Study Surveys by Program</td>
<td>24</td>
</tr>
<tr>
<td>2. Methods and Instruments Used to Evaluate the Validity of the Dual-Disorder Treatment Fidelity Scale</td>
<td>26</td>
</tr>
<tr>
<td>3. Dual-Disorder Treatment Fidelity Scale Item Analysis</td>
<td>36</td>
</tr>
<tr>
<td>4. Intraclass Correlation Coefficients (ICC) for Fidelity Ratings by Program</td>
<td>37</td>
</tr>
<tr>
<td>5. Descriptive Statistics for Continuous Consumer Outcomes</td>
<td>40</td>
</tr>
<tr>
<td>6. Regression Analyses for Fidelity Predicting Consumer Outcomes</td>
<td>40</td>
</tr>
<tr>
<td>7. Differences Between Fully and Moderately Implemented Programs’ Continuous Consumer Outcomes</td>
<td>42</td>
</tr>
<tr>
<td>8. Descriptive Statistics for Dichotomous Consumer Outcomes</td>
<td>43</td>
</tr>
<tr>
<td>9. Differences between Fully and Moderately Implemented Programs’ Dichotomous Consumer Outcomes</td>
<td>43</td>
</tr>
<tr>
<td>10. Regression Analysis for Fidelity Explaining Staff Survey Scores</td>
<td>44</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Dual-Disorder Treatment Fidelity Scale Total Ratings by Program Type</td>
<td>38</td>
</tr>
<tr>
<td>2. Dual-Disorder Treatment Fidelity Scale Item Ratings by Program Type</td>
<td>39</td>
</tr>
</tbody>
</table>
CHAPTER ONE

Introduction

As many as half of all individuals who have severe mental illnesses (SMI) will also have a substance use disorder (SUD) in their lifetime (Regier et al., 1990). Substance use among people who have SMI has been associated with adverse social, medical, and legal consequences, high treatment costs, and poor treatment outcomes (RachBeisel, Scott & Dixon, 1999). Programs applying an integrated approach to substance abuse and mental health treatments, however, have been shown effective in improving consumer outcomes (Drake, Mercer-McFadden, Mueser, McHugo & Bond, 1998). Nevertheless, few programs attempt to provide integrated services, and those that do rarely measure program implementation, adherence, or fidelity to the integrated model (U.S. Department of Health and Human Services, 1999; Drake, Mercer-McFadden, Mueser, et al., 1998).

Measuring program fidelity is essential in dual disorders treatment outcome research. Without such measures there is no assurance that the treatment under study was actually implemented. Indeed, when fidelity has been measured, implementation of the integrated model has been inconsistent across programs (Teague, Drake & Ackerson, 1995; Jerrell & Ridgely, 1999). Furthermore, outcomes in dual disorders treatment and research have been associated with program fidelity to the model. Integrated dual disorder programs with higher fidelity have been associated with outcomes better than those with lower fidelity (Ho et al., 1999; McHugo, Drake, Teague & Xie, 1999).
Fidelity is an important factor also when comparing different treatment models. Jerrell and Ridgely (1995, 1999) compared three integrated treatment models (behavioral skills, case management, and 12 step) and later examined the impact of fidelity or “robustness of program implementation” on the comparative effectiveness of the models. Their initial conclusions about which model demonstrated the best overall outcomes changed when robustness of program implementation was included as a variable in the analysis (Jerrell & Ridgely, 1999). These results emphasize the importance of considering fidelity when evaluating the effectiveness of dual disorder services, and in order to examine program fidelity, it must be measured. For these reasons, measures of program fidelity to the integrated model are needed.

Three fidelity scales have been developed to meet this need; the Comorbidity Program Audit and Self-Survey (COMPASS; Minkoff & Cline, 2001), the Dual-Disorder Treatment Fidelity Scale (Mueser, Noordsy, Drake & Fox, 2003), and a later version of the Dual-Disorder Treatment Fidelity Scale, the Integrated Dual Disorder Treatment (IDDT) Fidelity Scale (Substance Abuse and Mental Health Services Administration [SAMHSA], 2002). The COMPASS was designed for use as a needs assessment, quality improvement, and auditing tool for program fidelity to integrated treatment (Minkoff & Cline, 2001). The COMPASS is one part of a comprehensive toolkit that also includes scales for system and clinician level fidelity assessments. Because fidelity measurement is divided into system, program, and clinician level assessments, the COMPASS alone does not cover all domains of the integrated model.
Both the Dual-Disorder Treatment Fidelity Scale and the Integrated Dual Disorder Treatment (IDDT) Fidelity Scale were developed as a part of a national demonstration project, the Evidence-Based Practices Project (Drake, Goldman et al., 2001), and are appropriate for assessment of fidelity at the program or clinic level (Mueser et al., 2003; SAMHSA, 2002). To measure fidelity to the full integrated model, however, the IDDT Fidelity Scale is designed for use in conjunction with a second measure, the General Organizational Index (SAMHSA, 2002). Therefore, the Dual-Disorder Treatment Fidelity Scale appears the most comprehensive stand alone program level scale.

Finally, there are no published studies of psychometric properties for any of the three scales. For a scale to serve its function as an objective measure, an evaluation of its technical merits in terms of reliability and validity is essential (Anastasi, 1982). For both research and practice, the use of a fidelity scale which proves to be unreliable or invalid can lead to errant conclusions and can be wasteful of often scarce resources. As a result, consideration of reliability and validity evidence is a professional and ethical responsibility of the scale user (American Psychological Association [APA], 2002). Accordingly, the purpose of the present study is to assess the reliability and validity of the Dual-Disorder Treatment Fidelity Scale.
CHAPTER TWO

Literature Review

Terminology

**Severe mental illnesses (SMI).** Severe mental illnesses refers to a group of major mental illnesses, such as schizophrenia, bipolar disorder, recurrent major depression, or borderline personality disorder, which are associated with long-term impairments and extensive service needs (Drake, Mercer-McFadden, McHugo, et al., 1998). The Hawai‘i Adult Mental Health Division (AMHD) uses the term severe and persistent mental illness (SPMI) which is defined in part by the following:

Persons with SPMI exhibit emotional, cognitive, or behavioral functioning which is so impaired as to interfere substantially with their capacity to remain in the community without supportive treatment or require services of a long-term or indefinite duration. In these persons, mental disability is severe and persistent resulting in long term limitation in their functional capacities for primary activities of daily living such as interpersonal relationships, self-care, homemaking, employment, and recreation.

**Substance Use Disorders (SUD).** In the *Diagnostic and Statistical Manual of Mental Disorders IV-TR*, Substance Use Disorders are divided into two categories; Substance Abuse and Substance Dependence. Substance Abuse is characterized by a maladaptive pattern of substance use which results in recurrent and significant adverse consequences related to the repeated use. Substance Dependence is characterized by a cluster of cognitive, behavioral, and physiological symptoms indicating continued use despite significant substance related problems. Dependence also has a pattern of use that
can result in tolerance, withdrawal, and compulsive drug-taking behavior. “Craving” (a strong subjective drive to use the substance) is likely to be experienced by most individuals with dependence (American Psychiatric Association, 2000).

Dual Disorders. Terminology for the co-occurrence of severe mental illnesses and substance use disorders abounds. Terms used to describe the population include, for example; mental illness/chemical abuse (MICA), substance abuse/mental illness (SAMI), co-occurring addiction and mental disorder (COAMD), mental illness/substance abuse (MI/SA), comorbidity, dual diagnosis, and dual disorders (Drake, Mercer-McFadden, McHugo, et al., 1998). Here, the co-occurrence of severe mental illnesses and substance use disorders is referred to as dual disorders.

Prevalence

Current estimates of the prevalence of SUD among individuals with SMI vary widely, with specific estimates ranging from 10% to over 80% (Blanchard, Brown, Horan & Sherwood, 2000; Mueser et al., 1995). Probably the most comprehensive study on the prevalence of dual disorders was conducted as a part of the Epidemiological Catchment Area (ECA) study which sampled 20,291 individuals from both community and institutional populations (Regier et al., 1990). Prevalence estimates from the ECA study for any comorbid SUD were 47% among persons with schizophrenia and 56% among persons with bipolar disorder (Regier et al.). In contrast, the lifetime prevalence for SUD among the general population is estimated to be 17% (Regier et al.).

Furthermore, individuals with a SUD are at higher risk for mental illness than individuals without a SUD. The ECA study investigators found that individuals with
alcohol use disorders (AUD) were twice as likely to have an affective disorder and three times more likely to have schizophrenia than those without an AUD (Regier et al., 1990). Individuals with drug use disorders (DUD) were five times more likely to have an affective disorder and six times more likely to have schizophrenia than those without a DUD (Regier et al.). Thus, SUD are common among individuals with SMI, and mental illnesses are common among individuals with a SUD.

Consequences and Costs

The consequences of substance abuse among individuals with severe mental illnesses are great. These consequences include rates of social relationship, family, and financial problems, homelessness, victimization, infectious disease, violence, and incarceration higher than among individuals with an SUD alone (Drake & Mueser, 2000). Substance abuse among individuals with SMI is also associated with greater medication noncompliance, symptom exacerbation, psychiatric rehospitalization, and overall poorer course of SMI than among individuals with SMI alone (RachBeisel, Scott, & Dixon, 1999).

RachBeisel, Scott, and Dixon (1999) also reviewed the cost of treatment for individuals with dual disorders. Despite their tendency to drop out of outpatient treatment, individuals with dual disorders have treatment costs that exceed those of treatment for individuals with either disorder alone (Dickey & Azeni, 1996). This difference is primarily due to individuals’ high use of acute and expensive emergency and hospital services (RachBeisel, Scott, & Dixon). In turn, the
High use of costly services may be related to the historical lack of attention paid to one disorder or the other.

Historical Context

Following the era of deinstitutionalization, a new generation of individuals with SMI emerged in community mental health settings. These individuals were characterized by their young age, resistance to treatment, and disruptive behaviors. Substance use was cited often as one aspect of these consumers' difficult behavior (Osher & Drake, 1996). By the late 1980s, their substance use was recognized as a disorder outside of their mental illnesses, and, facilitated by the publication of the Diagnostic and Statistical Manual of Mental Disorders-III, which provided a format for documentation of multiple disorders, the concept of co-occurring severe mental illness and substance use disorders was established (Drake, Mueser, Clark & Wallach 1996).

Early conceptualizations of dual disorders emphasized two hypotheses. Substance abuse was thought to, one, result from individuals' coping with mental illness (self-medication hypothesis), or, two, mental illness or symptom exacerbation was thought to result from substance abuse (Blanchard, Brown, Horan & Sherwood, 2000). In each case one disorder was considered primary and the other secondary. Whether the mental illness or substance abuse was considered primary often depended on where the individual presented for treatment (Osher & Drake, 1996).

Individuals who had dual disorders traditionally received treatment for each disorder in separate service systems with separate funding mechanisms, training and credentialing procedures, and eligibility criteria (Mueser, Drake & Noordsy, 1998).
Treatment for each disorder occurred sequentially or in parallel in these separate systems with little collaboration between the two. Worse, individuals with dual disorders were often excluded from both systems as a result of one disorder or the other. Within both systems, access and adequate treatment were problems, and, as a result, the traditional approach to dual disorder treatment has had little positive impact on either the course of substance use disorders or the mental health outcomes of individuals with dual disorders (Drake & Mueser, 2000; Bartels, Drake & Wallach, 1995).

**Integrated Treatment**

Current conceptualizations of dual disorders emphasize a multifactorial etiology, including biological and psychosocial components (Drake & Wallach, 2000), as well as parallel processes of recovery for SMI and SUD (Minkoff, 1989). In contrast to the self-medication and mental illness corollary hypotheses, both disorders are considered primary, each carrying equal weight for treatment (Minkoff, 1989). Consistent with this conceptualization, treatment models that emphasize integration of mental health and SUD services into a single comprehensive program have emerged.

Integrated mental health and SUD services combine treatment for SMI and SUD. Interventions for both disorders and their related problems are delivered in a single program, by the same team of providers, and are modified to account for the presence of dual disorders (Drake, Mercer-McFadden, Mueser, et al., 1998). In their review of the effectiveness of integrated services in comparison to traditional nonintegrated services, Drake, Mercer-McFadden, Mueser, et al. found that integrated programs have gained support over traditional services in terms of consumers’ engagement in treatment,
reduction of substance use, and decrease in hospitalizations. Accordingly, integrated dual disorder treatment has been identified as an evidence-based practice (Drake, Essock, et al., 2001; U.S. Department of Health and Human Services, 1999).

The integrated model consists of several components, some of which are principles of treatment and others of which are specific interventions. These components have consistently demonstrated effectiveness among individuals with either SMI or SUD alone and, as part of comprehensive programs, have shown promise with the dual disorder population (Lehman et al., 2004; Miller & Wilbourne, 2002). While further research is needed to specify consistent definitions for each component and to determine the effectiveness of each component individually, programs showing high fidelity to a combination of the following components can be considered consistent with the evidence-based practice model (Drake, Essock, et al., 2001).

Assessment. Essential to consumer access to dual disorders services is accurate detection and assessment of substance use disorders among individuals who have SMI. Detection and assessment are first steps to effective treatment, yet, despite the high prevalence, SUD among individuals with SMI are often overlooked (Carey & Correia, 1998). Effective integrated programs regularly screen all consumers using instruments with demonstrated validity for the SMI population, such as the Dartmouth Assessment of Lifestyle Instrument (Rosenberg et al., 1998), and follow-up with comprehensive SUD assessment (Drake & Mueser, 2000). For accurate diagnosis these programs use a longitudinal approach, evaluating substance use and related problems over time, and gather information from collateral sources, such as lab tests, behavioral observations, and
family members (Carey & Correia, 1998). Mental health, substance use, and how the disorders interact are all comprehensively assessed (Mueser et al., 2003). Assessment occurs on an ongoing basis over the course of consumers’ involvement with the treatment program (Center for Mental Health Services [CMHS], 1998).

**Long-term perspective.** Both SMI and SUD are characterized by a chronic course often with multiple relapses, and recovery from dual disorders is a longitudinal process (Minkoff, 1989). Because recovery takes place over years, effective integrated programs typically apply a long-term perspective and serve consumers over years or their lifetime if needed (Drake, Essock, et al., 2001). The effect of program length on outcomes was examined by Brunette, Drake, Woods, and Hartnett (2001). They compared two integrated residential programs, one short-term program (average length of stay approximately two months) and one long-term program (average length of stay over a year) and found the long-term program to be more effective. Consumers in the long-term program became engaged in treatment, maintained abstinence after discharge, and maintained housing more often than consumers in the short-term program. While both these programs were residential, outpatient programs appear to be a more conducive and cost effective approach to long-term service provision (Drake & Mueser, 2000).

**Staged interventions.** Several authors have outlined stages of behavior change, treatment, or recovery. For example, Prochaska and DiClemente (1986) outlined five stages of change; precontemplation, contemplation, active change, maintenance, and relapse. Minkoff (1989) described parallel stages for recovery from both addiction and mental illness including; acute stabilization, engagement, prolonged stabilization, and
rehabilitation. Finally, Mueser et al. (1995) described eight stages of treatment; early and late stages of engagement, persuasion, active treatment, and relapse prevention.

In a staged approach, individual consumers at different stages in the change process receive interventions matched to their stage of readiness (Drake, Essock, et al., 2001). Integrated dual disorders programs shown to improve consumer outcomes had incorporated this principle into their treatment delivery (McHugo, Drake, Teague & Xie, 1999; Mueser, Drake & Miles, 1997). Implicit in the staged approach is an acknowledgment that consumers can cycle between stages throughout their recovery, and that consumers in the earlier stages should receive interventions aimed at increasing their motivation for change and decreasing the immediate harmful consequences of their problem behaviors (Bellack & DiClemente, 1999; Marlatt, 1998).

Harm reduction. Consumers with dual disorders often have a long history of failure in achieving their goals (Bellack & DiClemente, 1999). Particularly when in the early stages of change, consumers may not be willing to set a goal of abstinence. The harm reduction approach shifts the focus of treatment away from abstinence and judgments about substance use and toward the consequences that result from the substance use (Marlatt, 1998). Marlatt notes that the harm reduction approach recognizes abstinence as an ideal outcome, but accepts alternatives that reduce harmful consequences. For example, individuals' efforts to use less harmful substances or to use clean needles are considered treatment successes and are supported by the treatment team. Public health systems in countries such as the Netherlands, the United Kingdom,
Australia, and Canada have demonstrated the positive effects of a harm reduction approach (Marlatt, 1998).

Comprehensiveness. The most effective integrated programs are comprehensive, that is, they are inclusive of a full range of integrated services (Drake, Mercer-McFadden, Mueser, et al., 1998). Comprehensive programs consist of interventions to improve consumer access to and engagement in services and to treat the consumer from a biopsychosocial perspective. Consumers who have dual disorders have impairments in many domains, and effective dual disorders programs address each of these domains (Drake, Essock, et al., 2001). Comprehensive programs include interventions with demonstrated effectiveness for SMI (e.g., case management, housing/residential services, social skills training, and employment supports) as well as interventions with demonstrated effectiveness for SUD (e.g., motivational interviewing and cognitive-behavioral coping skills training; Drake, Mercer-McFadden, Mueser, et al., 1998; Lehman et al., 2004; Miller & Wilbourne, 2003). Integrated dual disorder treatment consists of the following specific interventions.

Pharmacological interventions. Optimal treatment for consumers who have SMI, including those who have co-occurring SUD, includes pharmacotherapy (U.S. Department of Health and Human Services, 1999). In effective programs, initial psychopharmacologic assessment occurs early, and prescription of psychiatric medications does not require consumers to be abstinent (CMHS, 1998). At the same time, consumers with dual disorders are less likely to adhere to their medication prescriptions (Fenton, Blyler & Heinssen, 1997). Therefore, effective pharmacological treatment for
individuals with dual disorders is performed in the context of an ongoing, empathetic clinical relationship through which adherence is addressed and close monitoring can occur (CMHS, 1998).

Advances in the pharmacology of substance abuse and schizophrenia offer some specific possibilities for the treatment of SUD in individuals with dual disorders (Wilkins, 1997). For example, some psychiatric medications (e.g., risperidone) have been associated with improvements in SUD (CMHS, 1998; Wilkins, 1997). Physicians or nurses prescribing medications may consider this knowledge when choosing a psychiatric medication for individuals with dual disorders (CMHS, 1998). CMHS also notes, for individuals with SMI and known substance dependence, prescription of addictive medications, such as benzodiazepines, is not recommended. Finally, individuals with dual disorders have been shown to benefit, in terms of relapse prevention, from medications developed to treat SUD, such as disulfiram (Wilkins, 1997). These medications are recommended as possible ancillary treatments to support a comprehensive integrated program of recovery (CMHS, 1998).

Assertive outreach. Many consumers with dual disorders drop out of treatment (Drake, Essock, et al., 2001). Assertive outreach involves the provision of interventions, such as medication management and skills training, in consumers’ natural settings, such as where they live or work (Drake, Bartels, Teague, Noordsy & Clark, 1993). The goal of outreach is to increase consumer access to and engagement in services.

Assertive outreach is one part of assertive community treatment (ACT). ACT has demonstrated effectiveness in reducing length of hospitalization and improving consumer
living conditions among individuals with SMI. Consequently, ACT is a recommended treatment of the Schizophrenia Patient Outcomes Research Team (Lehman et al., 2004). Several studies have shown these results to generalize to the dual disorder population. For example, Drake, McHugo and Noordsy (1993) demonstrated an association between assertive outreach, as part of a comprehensive integrated program, and the engagement and retention in treatment of all 18 study participants over the four year study period. All participants who attained remission remained actively engaged in treatment suggesting the assertive outreach contributed to their success. Jerrell and Ridgely (1995) also found assertive outreach, as part of a case management model, was associated with improvements in consumers' psychiatric symptoms and psychosocial outcomes.

*Family interventions.* Families are often the primary social supports of individuals who have dual disorders. Families endure a great amount of stress associated with both mental illness and substance use disorders (Sciacca & Hatfield, 1995). At the same time, social supports and supportive environments are important for the recovery of individuals with dual disorders (Drake, Essock, et al., 2001). Family interventions help maintain consumer/family relationships and supportive environments. In one study, 64% of families perceived a need for integrated treatment programs to provide family support groups as a service to consumers with dual disorders (Sciacca & Hatfield, 1995).

Family interventions inclusive of illness education, crisis intervention, emotional support, and coping skills training have consistently been found to help reduce relapse rates and family burden and improve family relationships among individuals with SMI (Lehman et al., 2004). Integrated programs that include family interventions tailored to
dual disorders show promising effectiveness (Mueser, Drake & Miles, 1997). In a randomized, controlled clinical trial, Barrowclough et al. (2001) showed that consumer participation in an integrated program inclusive of family interventions resulted in improvements in consumers' general functioning, reductions in positive symptoms, and an increase in days abstinent when compared with routine care alone. Family interventions appear effective in helping consumers maintain their immediate social supports.

Motivational interviewing. Most consumers with dual disorders enter treatment initially in the precontemplation stage; they are unmotivated to change their substance use behavior, to manage their psychiatric disorders, or to pursue functional goals (Drake, Essock, et al., 2001). Motivational interviewing is a goal directed counseling style used to increase motivation for change by helping consumers examine and resolve their ambivalence about their behavior. The approach is non-confrontational and emphasizes empathy, acceptance, and affirmation of the consumers’ autonomy (Rollnick & Miller, 1995).

Motivational interviewing has demonstrated effectiveness (i.e., reductions in substance use and improvements on social impact scales) in samples of individuals with SUD (Burke, Arkowitz, & Menchola, 2003). This finding appears to generalize to individuals with dual disorders. For example, Swanson, Pantalon, and Cohen (1999) found that consumers who participated in a single session of motivational interviewing at discharge from inpatient care attended their first outpatient appointment more often than consumers who received a standard discharge. In a randomized controlled comparison,
Barrowclough et al. (2001) found integrated treatment inclusive of motivational interviewing resulted in improvements in consumers' overall functioning greater than routine care. This finding is common to most integrated programs that include motivational interventions (Drake & Mueser, 2000).

**Substance abuse counseling.** Once consumers are motivated to set the goal of achieving abstinence they need to develop cognitive and behavioral skills to accomplish their goal and to maintain abstinence (Drake, Essock, et al., 2001). Substance abuse counseling for individuals with dual disorders serves to promote this development. Cognitive-behavioral coping skills training is another intervention that has well-established effectiveness among individuals with SUD (Miller & Wilbourne, 2003). Bellack and DiClemente (1999) adapted this model for individuals with dual disorders. The treatment includes education regarding the consequences of substance use, how to manage cravings, and includes skill building, such as social skills, problem solving skills for avoiding high-risk situations, drug and alcohol refusal skills, coping skills to deal with symptoms related to substance use, and relapse prevention (Bellack & DiClemente, 1999).

Integrated programs with demonstrably positive consumer outcomes have provided substance abuse counseling (Drake, Essock, et al., 2001). Specifically, Jerrell and Ridgely (1995) found that an integrated behavioral skills model of counseling, inclusive of symptom monitoring, relapse prevention, problem solving, and social skills training, was associated with improvements in areas of psychosocial functioning such as independent living, social relations, and work productivity. Cognitive behavioral skill
training as outlined above appears an effective model of substance abuse counseling for individuals with dual disorders.

Self-help. An alternative model of treatment for substance abuse is participation in self-help groups. Alcoholics Anonymous (AA) is the most prevalent and widely studied self-help group. Though participation in AA is almost always associated with reductions in alcohol use, the causal connection remains unclear. Nevertheless, twelve-step facilitation, a psychosocial intervention designed to facilitate participation in AA, has demonstrated efficacy (i.e., reductions in alcohol use and alcohol related problems) among individuals with primary alcohol use disorders (Project Match Research Group, 1997).

Among individuals with dual disorders, both social support and participation in self-help groups (e.g., AA, Narcotics Anonymous [NA], or Double Trouble Recovery [DTR]) have been associated with decreased substance use and improvements in psychosocial functioning (Vogel, Knight, Laudet & Magura, 1998). Laudet, Magura, Vogel, and Knight (2002) found that both a high level of perceived social support and regular participation in a DTR self-help group predicted decreased use of alcohol or drugs within the preceding year. Jerrell and Ridgely (1995) found participation in self-help (AA or NA) groups in combination with clinical support for attending the group meetings was associated with improvements in areas such as work performance and social relations.

While involvement in self-help groups appears beneficial for individuals with dual disorders, only a minority participates in these groups (Noordsy, Schwab, Fox &
Drake, 1996). In a review of the role of self-help for consumers with dual disorders, Noordsy, Schwab, Fox, and Drake concluded effective integrated dual disorder programs connect consumers with self-help groups and support participation for those consumers who show interest. For consumers who resist participation, however, pressure from providers is contraindicated.

Group dual disorders treatment. Another treatment format which facilitates social support for individuals with dual disorders is group treatment. Group treatment can provide a setting for consumers to share experiences, support, and coping strategies. Because many individuals abuse substances in a social context, a social or group context for treatment may be particularly helpful to the recovery process (Mueser & Noordsy, 1996). In addition, because individuals with dual disorders can be at varying stages of treatment, effective programs provide a variety of groups which target different stages of change. Types of groups include, for example, education, persuasion, relapse prevention, and social skills training groups (Mueser & Noordsy, 1996). Though the research is limited by high dropout rates, consumers who do participate regularly in group treatment have been shown to achieve increases in abstinence and engagement in treatment, and decreases in hospital use (Drake, Mercer-McFadden, Mueser, et al. 1998).

Overall, programs demonstrating positive outcomes among consumers with dual disorders generally have applied the foregoing principles and interventions. They have been comprehensive in nature and have included a variety of components from the menu above (Drake, Mercer-McFadden, Mueser, et al., 1998). Again, while further research is needed to dismantle this package of services, thus determining the most important or
effective components, a practical next step is to facilitate implementation of the integrated model. In order to accomplish this, psychometrically sound measures of fidelity to the model are needed.
CHAPTER THREE

Methods

Study Design

The purpose of the present study was to assess the reliability of the Dual-Disorder Treatment Fidelity Scale ratings and the validity of conclusions drawn from those ratings. Because variability in item ratings was essential for the validity analysis, the dispersion of item ratings was examined first. Reliability was then assessed through measures of inter-rater agreement. Finally, validity was assessed using three validation procedures or designs; criterion-related concurrent or known groups validation, criterion-related predictive validation, and construct validation.

In criterion-related validation, performance on a test or rating on a scale is checked against a direct and independent measure of the construct the test or scale is designed to measure. Concurrent or known groups validation occurs when the test or scale scores are validated against a criterion measure that reflects the existing status of the criterion, for example the current status or type of program (Anastasi, 1982). In the context of fidelity, concurrent validity is “demonstrated when the fidelity measure under investigation distinguishes between two or more groups which are known to differ in fidelity” (Calsyn, 2000, p. 108). In the present study, Dual-Disorder Treatment Fidelity Scale scores were checked against two program types; specialized Dual Diagnosis Enhanced (DDE; American Society of Addiction Medicine, 2001) programs and non-specialized community mental health centers (CMHCs). The scale was expected to
differentiate between these two program types. Specifically, the CMHCs were expected to score overall lower than the DDE programs.

Predictive validation occurs when test or scale scores are validated against a criterion measure which reflects future outcomes. The criterion measure is usually obtained a designated amount of time after the test scores, but sometimes is obtained at the same time due to the impracticality of extending validation procedures over long time intervals (Anastasi, 1982). In the context of fidelity “predictive validity occurs when the fidelity measure correlates with one or more measures of program effectiveness” (Calsyn, 2000, p. 108). For this study, Dual-Disorder Treatment Fidelity Scale scores were validated against consumer outcomes. Programs with high levels of fidelity were expected to show consumer outcomes better than those with low levels of fidelity.

Construct validation “requires the measure under investigation to exhibit correlations with other variables in theoretically meaningful ways” (Calsyn, 2000, p.108). Predictions are made about the relationships between the construct a scale is purported to measure (in this case fidelity) and other variables. The predictions are tested, and the degree to which they are supported speaks to the construct validity of the measure under investigation (Streiner, 1993). To assess construct validity for the Dual-Disorder Treatment Fidelity Scale, the relationship between fidelity and program staff knowledge of the integrated model was examined. Fidelity was expected to relate positively to program staff’s knowledge about integrated treatment.
Research Questions

Item Analysis

1. What is the dispersion of each fidelity scale item rating?

Reliability

2. To what extent do two independent raters agree on the chart review ratings used to derive fidelity scale ratings?

3. To what extent do the raters agree on fidelity scale ratings across programs, within each program, and on each scale item?

Validity

4. Criterion-related concurrent (known groups) validity:
   a. Do the program total fidelity scores distinguish between specialized DDE programs and community mental health center (CMHC) programs?
   b. Do the fidelity scale implementation labels (i.e., “not,” “moderately,” or “fully implemented”) distinguish between DDE and CMHC programs?
   c. Do the scale item ratings distinguish between DDE and CMHC programs?

5. Criterion-related predictive validity:
   a. Does program total fidelity predict consumer outcomes?
   b. Do “fully implemented” programs have consumer outcomes different than programs that are “not implemented?”

6. Construct validity:
   a. Is program total fidelity linearly related to staff knowledge of the integrated model?
b. Do staff from programs scoring in the “fully implemented” range have more knowledge of the integrated model than staff from programs scoring in the “not implemented” range?

**Participants**

**Programs**

Fidelity to the integrated model is assessed on the level of a treatment program. Six Hawaii Adult Mental Health Division (AMHD) programs participated in the study between March and May of 2004. All six programs had dual disorder treatment groups. Five of the programs were located on Oahu and one was on Maui. Three of the programs were community mental health centers (CMHC 1, CMHC 2, and CMHC 3) that did not specialize in dual disorders treatment. The other three participating programs (DDE 1, DDE 2, and DDE 3) specialized in dual disorder treatment and were contracted to meet American Society of Addiction Medicine (2001) Dual Diagnosis Enhanced criteria.

The American Society of Addiction Medicine (ASAM) uses the term Dual Diagnosis Enhanced (DDE) to describe programs that (a) are staffed by psychiatric and mental health clinicians as well as addiction treatment professionals who have been crossed trained, (b) accommodate people with dual diagnosis in their policies and procedures, including their assessment, treatment, and discharge planning procedures, (c) provide dual diagnosis specific and mental health group treatment, (d) provide motivational enhancement therapies and, (e) provide or collaborate closely with a program that provides crisis backup services (ASAM, 2001).
Individuals

Fifty-two consumers who were enrolled in the dual disorder groups at the six programs and fifty-eight staff members from the six programs participated in the study. For both consumers and staff, participation in the study occurred between March and May of 2004 and involved responding to self-administered surveys. See Table 1 for response rates.

Table 1. 
Number and Percent of Program Staff and Consumers Responding to Study Surveys by Program

<table>
<thead>
<tr>
<th>Program</th>
<th>Staff</th>
<th>Players</th>
<th>Consumers</th>
<th>Players</th>
</tr>
</thead>
<tbody>
<tr>
<td>Community Mental Health Center (CMHC) 1</td>
<td>18</td>
<td>86%</td>
<td>11</td>
<td>5%</td>
</tr>
<tr>
<td>CMHC 2</td>
<td>10</td>
<td>77%</td>
<td>5</td>
<td>3%</td>
</tr>
<tr>
<td>CMHC 3</td>
<td>12</td>
<td>75%</td>
<td>5</td>
<td>1%</td>
</tr>
<tr>
<td>Dual-Diagnosis Enhanced (DDE) 1</td>
<td>5</td>
<td>83%</td>
<td>8</td>
<td>100%</td>
</tr>
<tr>
<td>DDE 2</td>
<td>9</td>
<td>100%</td>
<td>19</td>
<td>95%</td>
</tr>
<tr>
<td>DDE 3</td>
<td>4</td>
<td>100%</td>
<td>4</td>
<td>100%</td>
</tr>
</tbody>
</table>

Instrumentation

Dual-Disorder Treatment Fidelity Scale

The Dual-Disorder Treatment Fidelity Scale consists of 20 items (see Appendix A). Each item is scored on a rating scale ranging from 1 (not implemented) to 5 (fully implemented). For example, for item 19, stage-wise treatment, a rating of 1 indicates interventions provided are contrary to the consumers' stage of recovery while a rating of 5 indicates that 80% or more of the interventions provided are consistent with the consumers' stage of recovery. Item ratings are summed for a total fidelity score with a
higher score indicating higher fidelity. Total scores are then compared to benchmarks to determine the program’s level of model implementation or implementation label. Due to a misprint in the benchmarks provided in the protocol, 5 points should be added to the total score before determining the implementation label (K. T. Mueser, personal communication, July 15, 2004). Scores ranging from: (a) 79-105 indicate that the integrated model is “fully implemented,” (b) 50-78 indicate “moderately implemented,” and (c) 25-49 indicate that the integrated model is “not implemented.” Thus, the scale provides data on program fidelity at the level of individual items, total scores, and implementation labels. In the present study, the validity of conclusions drawn from each one of these levels was assessed.

Data sources for the fidelity ratings include semi-structured interviews with the program directors and clinicians and chart reviews. Six of the fidelity scale item ratings are based solely on the chart reviews. Two additional item ratings are based on the combination of the interviews and chart reviews, though primarily on the interviews. To further operationalize these items, a chart review worksheet was added to the scale protocol (see Appendix B). For the worksheet, unless otherwise specified in the protocol, a cutoff of 70% was used to determine if the charts met protocol criteria. For example, for Item 5, Integrated Treatment Plan, if 70% of treatment plans received a “yes” (Y) rating for “specific,” then the program was rated as specific on the item. If >75% of the program’s treatment plans also addressed both disorders, but only 30% received a “yes” (Y) rating for “integrated,” the item was rated a 4 on the scale anchor points (“Both disorders addressed in >75% of plans, plus good specificity”). For the present study,
reliability was assessed for both the fidelity scale ratings and the chart review worksheet ratings.

To help assess the validity of the Dual-Disorder Treatment Fidelity Scale, further data were collected using two self-administered surveys, a consumer survey and a staff survey. Each will be described in more detail below. See Table 2 for a summary of methods and instruments used to assess validity.

Table 2.  
Methods and Instruments Used to Evaluate the Validity of the Dual-Disorder Treatment Fidelity Scale

<table>
<thead>
<tr>
<th>Method</th>
<th>Instrument</th>
<th>Respondent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Criterion-related validation</td>
<td>N/A (programs are self-identified)</td>
<td>N/A</td>
</tr>
<tr>
<td>Concurrent (known groups)</td>
<td>Consumer Survey</td>
<td>Consumers</td>
</tr>
<tr>
<td>Predictive</td>
<td>MHSIP Outcomes domain</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Substance abuse items</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lehman QOLI items</td>
<td></td>
</tr>
<tr>
<td>Construct validation</td>
<td>Staff Survey</td>
<td>Program Staff</td>
</tr>
</tbody>
</table>

Consumer Survey

The Consumer Survey was used to assess predictive validity. The survey consists of (a) the Mental Health Statistics Improvement Program (MHSIP) Consumer Satisfaction Survey outcomes domain (MHSIP, 2000), (b) two additional questions targeting substance abuse outcomes, and (c) five items from the Lehman Quality of Life Interview (QOLI; Lehman, 1995). The MHSIP and a very brief version of the QOLI are
administered to consumers regularly by the AMHD and were chosen in part to ease communication among researchers and providers within the AMHD.

The MHSIP outcomes domain is one of four domains on the MHSIP Consumer Satisfaction Survey, with the others being: access, appropriateness, and satisfaction (Altschul, 2003). The outcomes domain consists of eight items following the stem, “As a direct result of services I received.” For example, “As a direct result of services I received, I am better able to control my life.” Respondents are asked to rate each statement on a scale of agreement from 1 “strongly agree” to 5 “strongly disagree.” The eight items are averaged for the domain score.

The Consumer Survey also includes two items that were developed for this study to specifically target consumer substance use. These items measure consumers’ knowledge about substance use and their ability to remain abstinent. They follow the same stem as the MHSIP outcomes domain items, but were analyzed separately. The items are; “I know more about how my substance use affects my mental illness” and “I feel better able to stay clean and/or sober.” Response options range from 1 “strongly agree” to 5 “strongly disagree.”

Finally, the Consumer Survey includes five items from the Lehman QOLI. The QOLI contains measures of objective and subjective quality of life in nine domains; general life satisfaction, living situation, daily activities and functioning, family relations, social relations, finances, work and school, legal and safety issues, and health (Lehman, 1996). The five items included on the Consumer Survey were selected from the general life satisfaction, finances, legal and safety issues, and health domains as these domains
are not covered by the MHSIP outcomes domain items. An example from the health domain is, “In general, would you say your physical health is (circle one):” with response options ranging from “excellent” to “poor.” An example from the legal and safety issues domain is, “During the last 6 months, have you been arrested?” with response options of “yes” or “no.” For the complete Consumer Survey see Appendix C.

**Staff Survey**

The Staff Survey was developed to help assess the construct validity of the Dual-Disorder Treatment Fidelity Scale. The survey consists of 20 statements taken from the item definitions included in the Dual-Disorder Treatment Fidelity Scale protocol. The survey includes statements such as, “At admission for services, all consumers are screened for substance use using standardized (i.e., published) instruments.” The item response options are “true” or “false.” Seven statements were reversed such that the correct response is “false.” For example, “Psychotropic medications are prescribed for the treatment of mental illness in consumers who have active substance use disorders” was reversed to “Psychotropic medications are not prescribed for the treatment of mental illness in consumers who have active substance use disorders.” Correct responses are summed for a total score ranging from 0 to 20. For the complete Staff Survey see Appendix D.

**Procedure**

Evaluation of reliability and validity of the Dual-Disorder Treatment Fidelity Scale required data collection from a variety of sources. First, the evaluation required fidelity assessments of the six Hawaii Adult Mental Health Division (AMHD) programs.
These assessments included chart reviews and interviews with program staff, and were conducted by two undergraduate research assistants. Second, the evaluation required data that was separate from the fidelity assessments and was collected from consumers and staff at each program.

For the fidelity assessments, the two undergraduate research assistants were provided approximately 40 hours of rater training over a six week period. After the training, the two raters conducted fidelity assessments at each of the six AMHD programs. The raters followed the scale protocol; at each program they conducted a semi-structured group interview with the program director and clinicians and reviewed 10 charts randomly selected from 20 provided by the program. Each rater reviewed the same 10 charts. When the program served fewer than 10 consumers, as was the case with DDE 1 and DDE 3, all charts were reviewed. Eight charts were reviewed at DDE 1 and three charts were reviewed at DDE 3. The raters independently rated the interview data after each interview. Chart reviews including chart review worksheets were also completed by each rater independently.

To collect data from the program staff a visit was made to each program during their regularly scheduled staff meeting. Staff were introduced to the project, and those who consented to participate completed the Staff Survey. Visits were also made to the dual disorder groups at each program. Consumers enrolled in these groups were asked to participate in the study during a portion of their regular group time. The project was introduced, and consumers who consented completed the Consumer Survey. Consumers
who participated received a five dollar gift certificate from the AMHD. See Appendices E and F for the staff and consumer informed consent forms.

Data Analysis

Item analysis

The variability or dispersion of a scale item within a sample is one criterion of its utility (Bond, Wiliams et al., 2000). Accordingly, the dispersion of each Dual-Disorder Treatment Fidelity Scale item rating from each rater was examined and tabulated.

Reliability

Reliability was assessed through measures of inter-rater agreement. Chart review worksheet ratings (from which six of the fidelity scale item ratings were derived) were examined for percent agreement. To determine the inter-rater agreement of fidelity scale ratings across programs, within each program, and on each scale item, intraclass correlation coefficients were calculated. The intraclass correlation coefficient (ICC) is a measure of agreement between ratings on a continuous variable. Because the ICC is sensitive to systematic error in addition to differences in association and is particularly useful for small samples (where the Pearson correlation may be an overestimate; Streiner, 1993) it is well-suited to these reliability analyses.

Validity

Validity was assessed for conclusions drawn from item ratings, total scores, and implementation labels (i.e., “not,” “moderately,” and “fully implemented”). The validity of the item ratings was assessed through only the concurrent or known groups procedure, but the validity of the total scores and implementation labels was assessed using each of
the three validation designs described earlier; known groups, predictive, and construct validation.

Criterion-related concurrent (known groups) validity. To determine if the program total fidelity scores distinguished between DDE and CMHC programs, the scores were graphed by program type and compared visually. Means for each program type were also calculated, but because of the small number of programs in the analysis ($N = 6$) statistical significance testing was not appropriate. To determine if the fidelity scale implementation labels (i.e., “not,” “moderately,” or “fully implemented”) distinguished between DDE and CMHC programs, they were compared to the labels that were expected given the program type. The DDE programs were expected to be “fully implemented” and the CMHCs were expected to be “not implemented.

To determine if the scale item ratings distinguished between DDE and CMHC programs, item scores were graphed by program type and a K-means cluster analysis was employed. The K-means procedure splits objects into a specified number of groups based on their pattern of scores across variables. Distance scores were calculated from the raw data using Euclidean geometry. In this case, the two programs most distant from each other served as the initial cluster centers. The remaining programs were then assigned to one or the other cluster iteratively such that the within cluster variability was minimized while the between cluster variability was maximized (Bacher, 2002). Because the purpose of the analysis was not to generalize (i.e., identify reliable groupings that exist in the population of programs), but to determine if, based on the item scores, the two a
priori groups (DDE and CMHC) would be identified, the analysis was appropriate despite the small sample of programs ($N = 6$).

Criterion-related predictive validity. Prior to the predictive validity analyses, the continuous scores from the Consumer Survey were recoded such that higher numbers reflected better outcomes. To determine if program total fidelity scores predicted consumer outcomes, separate simple linear regression analyses were used for each of the five continuous outcome variables; the MHSIP outcomes domain (mean of the eight items), the substance use “knowledge” and “ability” items, and the QOLI overall life satisfaction and physical health items. For the analyses, all consumers within a program were assigned the same fidelity score, thus ignoring variability among consumers in the treatment they received. This procedure was appropriate because fidelity is a program level variable; however, it ignores the dependencies among the data from individuals within the same program. Still, regression analysis is generally robust to departures from this assumption (Pedhazur, 1997). An alpha level of .05 was used for this and all subsequent analyses involving statistical significance testing.

To determine the validity of the implementation labels (i.e., “not,” “moderately,” or “fully implemented”), comparisons between consumer outcomes from “fully implemented” and “not implemented” were planned. However, because no programs scored in the “not implemented” range, outcomes from the “moderately” and “fully implemented” ranges were compared. The data from the fully implemented programs were aggregated as were the data from the moderately implemented programs, again ignoring the dependencies inherent in the data from within each program as well as the
potential influence of program level variables other than fidelity. While statistical
techniques for analyzing clustered data are available (Hedeker, McMahon, Jason &
Salina, 1994; Paterson, 1991), the small number of programs in this case precluded their
use. So, for the continuous outcomes, $t$ tests were used to compare group means. For the
dichotomous outcomes, two by two contingency tables were examined through chi-
square tests of independence. When the expected count of a cell in the tables was less
than five, a continuity correction was applied.

*Construct validity.* The assessment of construct validity involved separate
analyses to assess the total scores and implementation labels. To determine if program
total fidelity was linearly related to staff knowledge of the integrated model, a simple
linear regression was employed. To assess the validity of the implementation labels,
comparisons between Staff Survey scores from “fully implemented” and “not
implemented” programs were planned. Because no programs scored in the “not
implemented” range, “moderately” and “fully implemented” programs were compared on
staff knowledge. Mean Staff Survey scores were calculated for the “moderately” and
“fully implemented” programs. A $t$ test was planned but, because Levene’s test indicated
violation of the homogeneity of variance assumption and the sample sizes were unequal
($n = 26$ and $n = 32$ respectively), a test of statistical significance (i.e., $t$ test) was not
appropriate (Boneau, 1960). Nonetheless, an effect size estimate was calculated.

*Ethical Considerations*

The present study was approved by the Institutional Review Boards of the
University of Hawai‘i and the State Department of Health through a joint review process
at the University of Hawai‘i Mānoa. All participants were treated in accordance with the

*Ethical principles of psychologists and code of conduct* (APA, 2002).
CHAPTER FOUR

Results

Fidelity Assessments

Item analysis. The dispersion of each Dual-Disorder Treatment Fidelity Scale item rating was examined. Table 3 shows the distribution of ratings by item (N = 240, i.e., two sets of twenty for each of six programs). Variability in the items was observed for 16 of the 20 scale items (Items 2, 3, 4, 5, 7, 8, 9, 10, 11, 13, 14, 15, 16, 18, 19, and 20). For example, on Item 2, Integrated Assessment, 17% of ratings were a “2,” 33% were a “3,” and 50% were a “4.” Four scale items showed no variability, Items 1, 6, 12, and 17. For example, on Item 1, Identification of Dual Disorders, 100% of the ratings (i.e., ratings from both raters across all six programs) were a “4.”

Reliability analyses. To assess reliability, both the chart review worksheets and fidelity scale scores were examined for inter-rater agreement. Chart review worksheet scores were inspected and compared between raters. A total of 49 charts were reviewed and rated in 11 areas. Agreement across all charts was 92%. Twenty-five of the 49 charts had no disagreement in ratings, seven had only one area of disagreement, 16 had two areas of disagreement, and one had three areas of disagreement. The remaining reliability analyses were based on the fidelity scale ratings.

Intraclass correlation coefficients were calculated for ratings across programs (ICC = 0.95, N = 120), within each program (see Table 4), and for each item (see Table 3). Overall, the coefficients were high with only CMHC 3 (ICC = .77; see Table 4) and
Items 19 (ICC = .57) and 20 (ICC = .79; see Table 3) below the standard of .80. For the remaining analyses, the raters' scores were averaged and these average ratings were used.

Table 3.
_Dual-Disorder Treatment Fidelity Scale Item Analysis_

<table>
<thead>
<tr>
<th>Item Label</th>
<th>ICC ((n = 6))</th>
<th>Distribution of Ratings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>1 Identification of Dual-Disorders</td>
<td>1.00</td>
<td>0%</td>
</tr>
<tr>
<td>2 Integrated Assessment</td>
<td>1.00</td>
<td>0%</td>
</tr>
<tr>
<td>3 Mental Health Assessment</td>
<td>1.00</td>
<td>0%</td>
</tr>
<tr>
<td>4 Substance Abuse Assessment</td>
<td>.98</td>
<td>17%</td>
</tr>
<tr>
<td>5 Integrated Treatment Plan</td>
<td>.97</td>
<td>17%</td>
</tr>
<tr>
<td>6 Integrated Crisis Plan</td>
<td>1.00</td>
<td>100%</td>
</tr>
<tr>
<td>7 Integration of Services</td>
<td>.98</td>
<td>8%</td>
</tr>
<tr>
<td>8 Comprehensiveness of Services</td>
<td>.87</td>
<td>8%</td>
</tr>
<tr>
<td>9 Time-Unlimited Services</td>
<td>1.00</td>
<td>17%</td>
</tr>
<tr>
<td>10 Outreach Capability</td>
<td>.99</td>
<td>42%</td>
</tr>
<tr>
<td>11 Client-to-Clinician Ratio</td>
<td>1.00</td>
<td>0%</td>
</tr>
<tr>
<td>12 Integrated Group Treatment</td>
<td>1.00</td>
<td>0%</td>
</tr>
<tr>
<td>13 Group Treatment</td>
<td>.92</td>
<td>0%</td>
</tr>
<tr>
<td>14 Motivational Interviewing</td>
<td>1.00</td>
<td>33%</td>
</tr>
<tr>
<td>15 Cognitive-Behavioral Counseling</td>
<td>1.00</td>
<td>0%</td>
</tr>
<tr>
<td>16 Family Interventions</td>
<td>.94</td>
<td>50%</td>
</tr>
<tr>
<td>17 Pharmacological Treatment</td>
<td>1.00</td>
<td>0%</td>
</tr>
<tr>
<td>18 Self-Help Liaison</td>
<td>.93</td>
<td>0%</td>
</tr>
<tr>
<td>19 Stage-Wise Treatment</td>
<td>.57</td>
<td>0%</td>
</tr>
<tr>
<td>20 Harm Reduction</td>
<td>.79</td>
<td>25%</td>
</tr>
</tbody>
</table>

_Note._ ICC = Intraclass Correlation Coefficient
Table 4.
*Intraclass Correlation Coefficients (ICC) for Fidelity Ratings by Program*

<table>
<thead>
<tr>
<th>Program</th>
<th>ICC (n = 20)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Community Mental Health Center (CMHC) 1</td>
<td>.96</td>
</tr>
<tr>
<td>CMHC 2</td>
<td>.99</td>
</tr>
<tr>
<td>CMHC 3</td>
<td>.77</td>
</tr>
<tr>
<td>Dual-Diagnosis Enhanced (DDE) 1</td>
<td>.96</td>
</tr>
<tr>
<td>DDE 2</td>
<td>.96</td>
</tr>
<tr>
<td>DDE 3</td>
<td>.99</td>
</tr>
</tbody>
</table>

*Validity Analyses*

*Criterion-related concurrent (known groups) validity.* Known groups validity was assessed through three procedures. First, CMHC and DDE programs were compared on the basis of total fidelity scores. Across programs, total scores ranged from 65.50 to 84.50. The CMHC program mean was 67.50 ($SD = 3.28$) and the DDE program mean was 76.50 ($SD = 7.55$). Because of the small number of programs in the analysis ($N = 6$) statistical significance testing was not appropriate. Instead, program total scores were graphed (see Figure 1) and CMHC programs were compared to DDE programs through visual examination. The CMHCs were expected to score lower than the DDE programs, and while this was true for two of the CMHC programs, one was rated higher than one of the DDE programs.

Second, the CMHC and DDE programs were compared on the basis of implementation labels. The CMHCs were expected to fall in the “not implemented” range and the DDE programs were expected to fall in the “fully implemented” range. Counter to these expectations, two CMHCs and one DDE program fell into the “moderately
implemented” range (50-78), and two DDE programs and one CMHC fell into the “fully implemented” range (79-105; see Figure 1).

![Graph showing Dual-Disorder Treatment Fidelity Scale Total Ratings by Program Type]

Figure 1.
*Dual-Disorder Treatment Fidelity Scale Total Ratings by Program Type*

The foregoing examinations ignored variability among programs at the item level. Consequently, the third procedure for assessing known groups validity was an analysis at the item level. Item scores for each program type were averaged and graphed (see Figure 2). No program type (DDE nor CMHC) scored consistently (i.e., across items) higher or lower than the other. Rather, the DDE and CMHC program ratings overlapped on some items (i.e., Items 1, 6, 12, 17, and 2), were close on some (e.g., Items 18, 16, 15, and 8),
and diverged on others (e.g., Items 10, 20, 3, 13, and 4). Furthermore, CMHC programs were rated higher than DDE programs on some items (e.g., Items 10, 20, and 9) and lower on others (e.g., Items 5, 14, 3, 13, and 4).

![Mean Rating vs. Item Number Graph]

Figure 2. *Dual-Disorder Treatment Fidelity Scale Item Ratings by Program Type*

To test the possibility that DDE and CMHC programs had distinct patterns of item scores, a K-means cluster analysis was employed. The groupings derived from the cluster analysis were compared to the groupings based on program type. The groupings matched (i.e., the DDE programs clustered together and the CMHCs clustered together), indicating that the fidelity scale distinguished between DDE and CMHC programs at the item level.

*Criterion-related predictive validity.* Three statistical procedures were used to determine the relationship between fidelity and consumer outcomes. First, simple linear regression was used to determine if fidelity was linearly related to the consumer outcomes measured on continuous scales. Descriptive statistics for the continuous consumer outcomes are summarized in Table 5, and the regression analysis results for each of the five variables are summarized in Table 6. No statistically significant (i.e., alpha < .05) relationships were found.
Table 5.
Descriptive Statistics for the Continuous Consumer Outcome Variables

<table>
<thead>
<tr>
<th>Outcome Variable</th>
<th>Range</th>
<th>Mean (SD)</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>MHSIP Outcomes</td>
<td>2.13 to 5.00</td>
<td>3.90 (.75)</td>
<td>3.88</td>
</tr>
<tr>
<td>Substance use outcomes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Knowledge</td>
<td>1.00 to 5.00</td>
<td>4.08 (.97)</td>
<td>4.00</td>
</tr>
<tr>
<td>Ability</td>
<td>3.00 to 5.00</td>
<td>4.33 (.77)</td>
<td>5.00</td>
</tr>
<tr>
<td>Overall life satisfaction</td>
<td>1.00 to 6.00</td>
<td>4.28 (1.50)</td>
<td>4.00</td>
</tr>
<tr>
<td>Physical health</td>
<td>1.00 to 5.00</td>
<td>2.98 (1.15)</td>
<td>3.00</td>
</tr>
</tbody>
</table>

- `a` Mental Health Statistics Improvement Project 2000 Outcomes domain; possible scores range from 1 to 5, recoded such that higher scores indicate better outcomes.
- `b` Possible scores range from 1 to 5, recoded such that higher scores indicate better outcomes.
- `c` From the Quality of Life Interview (Lehman, 1995); possible scores range from 1 to 6, with higher scores indicating greater satisfaction.
- `d` From the Quality of Life Interview (Lehman, 1995); possible scores range from 1 to 5, recoded such that higher scores indicate better health.

Table 6.
Regression Analyses for Fidelity Predicting Continuous Consumer Outcome Variables

<table>
<thead>
<tr>
<th>Outcome Variable</th>
<th>df</th>
<th>B</th>
<th>SE B</th>
<th>β</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>MHSIP Outcomes</td>
<td>1, 50</td>
<td>.01</td>
<td>.02</td>
<td>.10</td>
<td>.48</td>
</tr>
<tr>
<td>Substance use outcomes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Knowledge</td>
<td>1, 50</td>
<td>.02</td>
<td>.02</td>
<td>.11</td>
<td>.42</td>
</tr>
<tr>
<td>Ability</td>
<td>1, 49</td>
<td>.02</td>
<td>.02</td>
<td>.15</td>
<td>.30</td>
</tr>
<tr>
<td>Overall life satisfaction</td>
<td>1, 48</td>
<td>.07</td>
<td>.04</td>
<td>.25</td>
<td>.08</td>
</tr>
<tr>
<td>Physical health</td>
<td>1, 50</td>
<td>.02</td>
<td>.03</td>
<td>.10</td>
<td>.47</td>
</tr>
</tbody>
</table>

- `a` Mental Health Statistics Improvement Project 2000 Outcomes domain; possible scores range from 1 to 5, recoded such that higher scores indicate better outcomes.
- `b` Possible scores range from 1 to 5, recoded such that higher scores indicate better outcomes.
- `c` From the Quality of Life Interview (Lehman, 1995); possible scores range from 1 to 6, with higher scores indicating greater satisfaction.
- `d` From the Quality of Life Interview (Lehman, 1995); possible scores range from 1 to 5, recoded such that higher scores indicate better health.
The remaining two predictive validity analyses were applied to test the implementation labels, that is, the ranges provided in the scale protocol classifying programs as “not,” “moderately,” or “fully implemented.” The planned analyses involved comparisons between consumer outcomes from programs scoring in the “not implemented” range to those from programs scoring in the “fully implemented” range. However, because no programs scored in the “not implemented” range, outcomes from the “moderately” and “fully implemented” ranges were compared. With this modification, statistical power was decreased; the consumer sample size had been based on expected large effects versus the small differences that might be expected between programs closer in their levels of fidelity.

The results for the continuous outcomes are summarized in Table 7; the table includes means, standard deviations, and the t test results. Descriptive statistics for the dichotomous outcomes are reported in Table 8, and the results of the chi square analyses are summarized in Table 9. No differences were statistically significant.

Because statistical power was a concern, post hoc power analyses were conducted for each consumer outcome variable. The estimates indicated that statistical power was indeed low (see Tables 7 and 9), with the exception of one dichotomous variable, “arrest” (for which it was .95). Consequently, standardized effect sizes (i.e., Cohen’s d for the continuous outcome variables and Cohen’s w or Phi for the dichotomous outcome variables) were calculated (Cohen, 1977). While the small sample did not bias the effect size estimates, it lessened their precision (demonstrated by the wide confidence intervals reported in Table 7).
Table 7.
Differences between Fully and Moderately Implemented Dual-Disorder Treatment Programs’ Continuous Consumer Outcomes

<table>
<thead>
<tr>
<th>Variable</th>
<th>Full (n = 38)</th>
<th>Moderate (n = 14)</th>
<th>t</th>
<th>df</th>
<th>p</th>
<th>1-β&lt;sup&gt;a&lt;/sup&gt;</th>
<th>d (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MHSIP Outcomes&lt;sup&gt;b&lt;/sup&gt;</td>
<td>3.94 (.74)</td>
<td>3.81 (.80)</td>
<td>.52</td>
<td>50</td>
<td>.61</td>
<td>.08</td>
<td>.17 (-.44 - .78)</td>
</tr>
<tr>
<td>Substance use outcomes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Knowledge&lt;sup&gt;c&lt;/sup&gt;</td>
<td>4.13 (.93)</td>
<td>3.93 (1.07)</td>
<td>.67</td>
<td>50</td>
<td>.51</td>
<td>.10</td>
<td>.20 (-.41 - .82)</td>
</tr>
<tr>
<td>Ability&lt;sup&gt;d&lt;/sup&gt;</td>
<td>4.39 (.72)</td>
<td>4.15 (.90)</td>
<td>.98</td>
<td>49</td>
<td>.33</td>
<td>.16</td>
<td>.31 (-.32 - .94)</td>
</tr>
<tr>
<td>Overall life satisfaction&lt;sup&gt;e,f&lt;/sup&gt;</td>
<td>4.36 (1.46)</td>
<td>4.07 (1.64)</td>
<td>.61</td>
<td>48</td>
<td>.55</td>
<td>.09</td>
<td>.19 (-.43 - .81)</td>
</tr>
<tr>
<td>Physical health&lt;sup&gt;g&lt;/sup&gt;</td>
<td>3.00 (1.14)</td>
<td>2.93 (1.21)</td>
<td>.20</td>
<td>50</td>
<td>.84</td>
<td>.05</td>
<td>.06 (-.55 - .67)</td>
</tr>
</tbody>
</table>

<sup>a</sup> Statistical power (Cohen, 1977).
<sup>b</sup> Mental Health Statistics Improvement Project 2000 Outcomes domain; possible scores range from 1 to 5, recoded such that higher scores indicate better outcomes.
<sup>c</sup> Possible scores range from 1 to 5, recoded such that higher scores indicate better outcomes.
<sup>d</sup> Moderate (n = 13).
<sup>e</sup> From the Quality of Life Interview (Lehman, 1995); possible scores range from 1 to 6, with higher scores indicating greater satisfaction.
<sup>f</sup> Full (n = 36).
<sup>g</sup> From the Quality of Life Interview (Lehman, 1995); possible scores range from 1 to 5, recoded such that higher scores indicate better health.
Table 8.
**Descriptive Statistics for the Dichotomous Consumer Outcomes**

<table>
<thead>
<tr>
<th>Outcome Variable</th>
<th>Number (%) responding Yes</th>
<th>Number (%) responding No</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Legal and safety</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arrest</td>
<td>10 (19%)</td>
<td>42 (81%)</td>
</tr>
<tr>
<td>Victim</td>
<td>8 (15%)</td>
<td>44 (85%)</td>
</tr>
<tr>
<td><strong>Finances</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Food</td>
<td>40 (78%)</td>
<td>11 (22%)</td>
</tr>
<tr>
<td>Clothing</td>
<td>31 (61%)</td>
<td>20 (39%)</td>
</tr>
<tr>
<td>Housing</td>
<td>39 (78%)</td>
<td>11 (22%)</td>
</tr>
<tr>
<td>Medical</td>
<td>38 (79%)</td>
<td>10 (21%)</td>
</tr>
<tr>
<td>Transportation</td>
<td>28 (55%)</td>
<td>23 (45%)</td>
</tr>
</tbody>
</table>

*Note.* Variables from the Quality of Life Interview (Lehman, 1995); possible responses were Yes or No, with a No response indicating a better outcome on the Legal and Safety variables and a Yes response indicating a better outcome on the Finances variables.

Table 9.
**Differences between Fully and Moderately Implemented Dual-Disorder Treatment Programs' Dichotomous Consumer Outcomes**

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>$X^2$</th>
<th>df</th>
<th>p</th>
<th>$w^a$</th>
<th>$1-\beta^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Legal and safety</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arrest</td>
<td>52</td>
<td>3.02</td>
<td>1</td>
<td>.08</td>
<td>.24</td>
<td>.95</td>
</tr>
<tr>
<td>Victim</td>
<td>52</td>
<td>.00</td>
<td>1</td>
<td>1.00</td>
<td>.00</td>
<td>.05</td>
</tr>
<tr>
<td>Finances</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Food</td>
<td>51</td>
<td>.00</td>
<td>1</td>
<td>1.00</td>
<td>.00</td>
<td>.05</td>
</tr>
<tr>
<td>Clothing</td>
<td>51</td>
<td>.11</td>
<td>1</td>
<td>.74</td>
<td>.05</td>
<td>.10</td>
</tr>
<tr>
<td>Housing</td>
<td>50</td>
<td>.10</td>
<td>1</td>
<td>.75</td>
<td>.04</td>
<td>.10</td>
</tr>
<tr>
<td>Medical</td>
<td>48</td>
<td>.00</td>
<td>1</td>
<td>1.00</td>
<td>.00</td>
<td>.05</td>
</tr>
<tr>
<td>Transportation</td>
<td>51</td>
<td>.19</td>
<td>1</td>
<td>.67</td>
<td>.06</td>
<td>.14</td>
</tr>
</tbody>
</table>

*a Effect size index for dichotomous variables, equivalent to Phi (Cohen, 1977).

*b Statistical power (Cohen, 1977).

*c From the Quality of Life Interview (Lehman, 1995); possible responses were Yes or No.*
Construct validity. The assessment of construct validity involved two analyses. First, the relationship between fidelity and staff knowledge of the integrated model was examined using a simple linear regression (see Table 10). No statistically significant relationship was found.

Table 10. Regression Analysis for Fidelity Explaining Staff Survey Scores

<table>
<thead>
<tr>
<th>df</th>
<th>B</th>
<th>SE B</th>
<th>β</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>1, 56</td>
<td>.07</td>
<td>.06</td>
<td>.16</td>
<td>.22</td>
</tr>
</tbody>
</table>

Note. Staff Survey scores ranged from 6 to 19; $M = 14.72$ ($SD = 2.88$).

Similar to the predictive validity analyses, the second planned analysis of construct validity was an examination of the scale ranges (i.e., “not,” “moderately,” and “fully implemented”). The mean Staff Survey score from the “moderately implemented” programs ($M = 14.85, SD = 3.62$) appeared similar to that from the “fully implemented” programs ($M = 14.56, SD = 2.24$). Because Levene’s test indicated violation of the homogeneity of variance assumption and the sample sizes were unequal ($n = 26$ and $n = 32$ respectively) a test of statistical significance (i.e., $t$ test) was not appropriate (Boneau, 1960). Nonetheless, an effect size estimate was calculated ($d = -.10, CI = -.62 - .42$). Though imprecise, the estimate indicates a generally trivial effect in the opposite direction than expected.
CHAPTER FIVE

Discussion

The purpose of this study was to evaluate the reliability and validity of the Dual-Disorder Treatment Fidelity Scale. Overall, the results indicate good inter-rater reliability. The results of the validity analyses were mixed, with evidence for the validity of conclusions drawn from item level ratings greater than for those drawn from total fidelity scores or implementation labels. The study also highlights a number of methodological issues that deserve consideration in future work in this area. The item, reliability, and validity analyses are discussed more completely below.

Item Analysis

Four scale items did not vary across programs. These included Items 1, Identification of Dual-Disorder Clients, 6, Integrated Crisis Plan, 12, Integrated Group Treatment for Dual Disorders, and 17, Pharmacological Treatment of Mental Illness. The lack of variance on these items suggests they may not be useful; however, as explained in more detail below, the methodology used in this study accounts for this finding for at least two of the items.

One of the selection criteria for program participation in the study was that the program had a dual disorder group. As a result, Item 12, Integrated Group Treatment for Dual Disorders, did not show variability across programs. Additionally, Item 1, Identification of Dual-Disorder Clients, may not have varied because of the chart review procedure. The scale protocol suggests a review of ten randomly selected charts of clients with dual disorders. Consequently, for Item 1 the charts reviewed were charts of clients
RECEIVED
AS
FOLLOWS
CHAPTER FIVE

Discussion

The purpose of this study was to evaluate the reliability and validity of the Dual-Disorder Treatment Fidelity Scale. Overall, the results indicate good inter-rater reliability. The results of the validity analyses were mixed, with evidence for the validity of conclusions drawn from item level ratings greater than for those drawn from total fidelity scores or implementation labels. The study also highlights a number of methodological issues that deserve consideration in future work in this area. The item, reliability, and validity analyses are discussed more completely below.

Item Analysis

Four scale items did not vary across programs. These included Items 1, Identification of Dual-Disorder Clients, 6, Integrated Crisis Plan, 12, Integrated Group Treatment for Dual Disorders, and 17, Pharmacological Treatment of Mental Illness. The lack of variance on these items suggests they may not be useful; however, as explained in more detail below, the methodology used in this study accounts for this finding for at least two of the items.

One of the selection criteria for program participation in the study was that the program had a dual disorder group. As a result, Item 12, Integrated Group Treatment for Dual Disorders, did not show variability across programs. Additionally, Item 1, Identification of Dual-Disorder Clients, may not have varied because of the chart review procedure. The scale protocol suggests a review of ten randomly selected charts of clients with dual disorders. Consequently, for Item 1 the charts reviewed were charts of clients
who already had been identified as having dual disorders. This would not have affected ratings for the DDE programs, as entry into the program required dual disorders, but may have inflated the ratings for the CMHC programs. Using a random sample of charts from all consumers served by the program would alleviate this problem.

None of the charts reviewed in any of the participating programs had a written crisis plan; therefore all six programs received the lowest rating on Item 6, Integrated Crisis Plan. In contrast, all six programs received the highest rating on Item 17, Pharmacological Treatment of Mental Illness. Investigation across a wider variety of programs (e.g., primary substance abuse programs in addition to primary mental health programs) is needed to determine the utility of these items.

Reliability

Overall, independent raters tended to agree on Dual-Disorder Treatment Fidelity Scale ratings. The results indicate that the scale can be implemented reliably by individuals having received approximately 40 hours of training. Though adequate, the reliability of the ratings from the first program assessed (CMHC 3) was low (ICC = .77) relative to the other programs. This finding may reflect rater inexperience and suggests incremental value of adding the completion of a full assessment to the training protocol.

At the item level, the reliability of ratings on Item 19, Stage-Wise Treatment, was an exception (ICC = .57) to otherwise high inter-rater agreement. Ratings for this item are determined primarily through a semi-structured interview, and the interview questions for the item are less direct than those for other items, requiring greater inference from the raters and, perhaps, accounting for the discrepancy. The addition of more direct interview
questions to this section of the scale protocol may improve the inter-rater reliability of ratings for the item.

Validity

Validity was assessed for conclusions drawn from item ratings, total scores, and implementation labels (i.e., “not,” “moderately,” and “fully implemented”). Concurrent (known groups), predictive, and construct validation procedures were used. The results were mixed. The item level ratings demonstrated validity, but the analyses of total scores and implementation labels were less conclusive primarily due to methodological problems that limit the validity of inferences drawn from the predictive and construct validity analyses. A discussion of each validity analysis follows.

Criterion-related concurrent (known groups) validity. The Dual-Disorder Treatment Fidelity Scale was expected to differentiate between DDE and CMHC programs. On the basis of total fidelity scores, DDE programs were expected to rate higher than CMHC programs. While this was true for four of the programs, one of the CMHCs received a total score higher than one of the DDE programs. Unlike the other two CMHCs, this center had an established alcohol and drug abuse program that was unexpectedly consistent with the integrated model. Overall, the DDE programs were expected to fall in the “fully implemented” range and the CMHCs were expected to fall in the “not implemented” range. Instead, no programs fell in the “not implemented” range, and only two of the three DDE programs fell in the “fully implemented range.” Thus, no clear and consistent differentiation between program types emerged on the basis of programs’ total fidelity scores or implementation labels.
Conversely, an examination of program ratings at the item level revealed differences between program types based on relative strengths and limitations of each. For example, the CMHC programs were strong on Item 10, outreach capability, whereas the DDE programs were relatively limited in that area. In contrast, the DDE programs were strong on the assessment items (Items 3 and 4) and the CMHCs relatively limited in that area. The scale differentiated between program types at the item level. These findings provide some support for the validity of conclusions drawn from item level ratings. At the same time, they raise larger questions about the construct of fidelity and the level at which fidelity to the integrated treatment model is assessed.

Integrated dual disorder treatment is a multidimensional construct. While some of the Dual-Disorder Treatment Fidelity Scale items overlap (e.g., the assessment items) others tap distinct dimensions (e.g., time-unlimited services and motivational interviewing). The procedure of summing scores for a composite fidelity score may be misleading. In this process the items are weighted equally suggesting that they are equally important. At this stage, however, we know very little about the importance or effectiveness of each item or component individually (Drake, Mercer-McFadden, Mueser et al., 1998). Much more study is needed to dismantle the integrated model, identify the active ingredients, and refine the model definition. Nevertheless, the development and examination of a fidelity scale for the model represents an important step in this direction.

A second question raised by the known groups analysis regards the level at which fidelity to the integrated model is assessed. The Dual-Disorder Treatment Fidelity Scale
was designed for program level assessments. While implementation of the integrated model at the program level is consistent with what the literature suggests is effective (i.e., the consumer need only engage with one comprehensive program; Drake, Mercer-McFadden, Mueser et al., 1998), it is not the only approach to integrated service delivery.

In their Comprehensive Continuous Systems of Care (CCISC) model, Minkoff and Cline (2001) suggest that programs within a system of care (e.g., a county operated system or a state funded system) may strive for varying levels of fidelity to the integrated model. For example, some programs within a system may have limited dual diagnosis capacity (e.g., meet ASAM [2001] Dual Diagnosis Capable criteria), and others may specialize in dual disorder treatment (e.g., meet DDE criteria). These programs may have different strengths and limitations (as was true of the CMHC and DDE programs assessed in this study) that complement each other in the larger system of care. The system may then be considered comprehensive rather than each individual program, and fidelity would be assessed at the system level. While Minkoff and Cline offer a toolkit containing fidelity scales for system, program, and clinician level assessments, there are no published data on the effectiveness of this type of integrated service delivery or on the psychometric properties of the scales. Exploration of these matters may aid in the development of a clearer definition for the integrated model.

Criterion-related predictive validity. The results of the predictive validation procedure were inconclusive. The expected positive relationship between fidelity and consumer outcomes was not supported by the data. Furthermore, consumers served by programs classified by the scale as "moderately implemented" did not show outcomes
different from "fully implemented" programs. Four important methodological problems, however, prevent sound conclusions on the basis of these results.

The first problem involves the analysis of clustered data. Fidelity is a program level variable, and consumer outcomes are individual level variables that are clustered or nested within program. Two issues arise from this situation. First, if the consumer outcome data are used at the individual level (i.e., not aggregated to the program level), they violate the assumption of independence when compared across programs. This problem arose in the regression analyses. Second, if not included in the analysis, program level variables (other than fidelity) have unknown effects on the analysis. For example, in the comparisons between "moderately" and "fully implemented" programs, the influence of cluster level variables, such as program or program type, remains unknown. In the regression analyses, this problem manifest in the perfect confounding of fidelity with program, resulting in ambiguity about any conclusions drawn.

These issues have recently been discussed by Mowbray, Holter, Teague and Bybee (2003) in the specific context of validating fidelity measures. While these authors offer multilevel statistical techniques (e.g., hierarchical linear modeling, random-effects regression models, or moderated multiple regression) as a solution, they acknowledge that the program sample is usually of inadequate size and variability for their use. This was true for this study and may be equally true for future studies in Hawaii. For future work in this area, procedures other than predictive validation may prove more feasible and conclusive.
The foregoing issue aside, a second problem emerges. Both predictive validity analyses were impacted by the unexpectedly restricted range of total fidelity scores (scores ranged from 65.5 to 84.5 out of 105 possible points). Restriction of range attenuates standardized regression coefficients and consequently impacts significance testing (Pedhazur, 1997). In this case, the standardized coefficients suggest small effects in the expected direction; however, it is not clear whether or not these effects would have shown statistical significance had the range not been restricted.

One advantage of regression is that the unstandardized coefficients are not affected by the restriction of range (Pedhazur, 1997). Interpretation of these coefficients, however, is less clear. In this case, consumer outcomes were measured primarily on a five point scale of agreement. For example, the substance abuse outcome “ability,” (i.e., “I feel better able to stay clean and sober”) was rated for agreement from 1 to 5. For this item, the unstandardized coefficient was .02. Is an increase of .02 for each point increase in fidelity a meaningful gain on this scale of agreement? A 30 point increase in fidelity can bring a program from “not implemented” to “fully implemented.” According to the coefficient for this item, a 30 point increase in fidelity would correspond with .6 of a point increase in agreement on this outcome measure (i.e., “I feel better able to stay clean and sober” rated for agreement from 1 to 5). This does not seem to be a meaningful difference.

The restricted range of total fidelity scores also affected the analysis of the scale ranges or implementation labels (i.e., “not,” “moderately,” and “fully implemented”), though indirectly. When the comparison of programs changed from “not” versus “fully
implemented" to "moderately" versus "fully implemented," the expected effect size changed from large to medium or small. As a result, the sample size and subsequent statistical power of the analyses were rendered inadequate.

In effort to recover some information from the analysis, standardized effect size estimates were calculated. The continuous consumer outcome variables (i.e., the MHSIP outcomes domain, substance abuse questions, and QOLI overall life satisfaction and physical health items) showed effects that were generally consistent with those that, by convention, are small but not trivial (Cohen, 1977). This is what we might expect from comparisons between programs so close in fidelity. Note that there was an exception; the "physical health" variable showed a trivial effect \( d = .06 \). The estimates for the dichotomous variables, on the other hand, (i.e., the remaining QOLI items) suggest generally trivial effects or, in the case of the "arrest" variable, a small effect in the opposite direction than expected \( w = .24 \). It is not clear whether the inconsistent pattern of observed results is due to a lack of effect or a lack of precision in the effect size estimates. So, despite this further analysis, the results remain inconclusive. Future efforts to determine predictive validity may benefit from a program sample having a larger range of fidelity scores and/or a larger consumer sample.

A third important methodological problem involves the consumer sampling procedure. Consumers were recruited from the dual disorder groups at each program. This process resulted in differential response rates between DDE and CMHC programs. While virtually all consumers who were enrolled in the groups participated in the study, not all consumers identified as having dual disorders were actually enrolled. In fact, the
Proportion of consumers enrolled in the groups varied widely between program types. All of the consumers enrolled in DDE programs were also enrolled in the corresponding groups, but only 1-5% of consumers identified as having dual disorders at the CMHCs were enrolled in the CMHC groups. The consumers served by the CMHC programs, identified as having dual disorders, and not enrolled in the groups, may have differed from those who were enrolled. These potential differences in functioning, for example, may have biased the Consumer Survey scores, thereby making the validity of the analysis even more questionable.

The disparate group enrollment rate for consumers served by the CMHC versus DDE programs also raises a question regarding the content validity of the Dual-Disorder Treatment Fidelity Scale. While content validity was not directly assessed, and dual disorder groups are only one dimension of the integrated model, the contrasting group enrollment rates suggest differences in coverage or penetration (i.e., the degree to which the services reach the target population) that were not detected by the scale. Considering that treatment drop-out is cited as a significant problem for individuals with dual disorders and that engagement in services is cited as an important outcome in the integrated treatment literature (Drake, Mercer-McFadden, Mueser et al., 1998), penetration seems an essential aspect of the model. Accordingly, the IDDT Fidelity Scale used in conjunction with the General Organizational Index (GOI; SAMHSA, 2002) includes a measure of penetration. While the utility of the IDDT Fidelity Scale plus GOI as a whole remains unknown, the measure of penetration may be a useful addition to the Dual-Disorder Treatment Fidelity Scale.
The fourth limitation of the predictive validation procedure concerns the consumer outcomes measure. The Consumer Survey consisted of broad outcome measures that are distal to the specific interventions included in the integrated model. While measures of quality of life, for example, can provide practical and meaningful data, treatment effects are more likely to be detected by outcome measures more specific to the target of an intervention (Matt & Navarro, 1997). Furthermore, integrated treatment has most consistently shown effectiveness in increasing treatment engagement and decreasing substance use (Drake, Mercer-McFadden, Mueser et al., 1998). The Consumer Survey did not include a direct measure of engagement, and while two questions about substance use were included, these items have not been validated. Finally, while the Lehman QOLI has demonstrated sound psychometric properties, the Consumer Survey included only a small number of QOLI items. The validity of these items independent of the full measure is unknown. The use of standardized measures that are proximal to the intervention targets (e.g., treatment engagement, stage of change measures, and drinking frequency/amount) in addition to the broad measures used in the present study may have strengthened the study design.

_Construct validity._ Program staff knowledge about the integrated model was expected to relate positively to program total fidelity scores. In addition, staff from the “fully implemented” programs were expected to demonstrate more knowledge about the model than staff from the “moderately implemented” programs. Neither of these expectations was supported by the data. However, the construct validity analysis suffered methodological problems similar to those of the predictive validity analysis.
The regression analysis of fidelity as an explanatory variable for staff knowledge was impacted by the restricted range of fidelity total scores. An examination of the unstandardized regression coefficient, however, suggests little linear relationship. For example, a change from a 50% to an 80% score on the Staff Survey would require an increase of 114 fidelity points, more than are on the scale. In terms of the implementation labels, staff from the "moderately implemented" programs were not more knowledgeable of the model than staff from the "fully implemented" programs. This analysis, however, was also limited by the restricted range of fidelity scores; it was additionally hindered by the heterogeneity of variance of the Staff Survey scores in that statistical comparison of means was not appropriate.

Like the Consumer Survey data, the Staff Survey data were nested within program, thus casting doubt on the validity of these analyses. The influence of program variables other than fidelity (e.g., program type) on the differences in staff knowledge between "moderately" and "fully implemented" programs is unknown. Furthermore, the explanatory power of fidelity outside of the specific programs remains unspecified.

A final limitation of the construct validity analysis involves the Staff Survey as an outcome measure. Because there were no existing instruments that measure staff knowledge of the integrated model, the Staff Survey was developed specifically for use in this study. While the range and dispersion of scores on the Staff Survey were adequate it has not been evaluated in terms of reliability and validity.

The study of the reliability of measurement instruments and the validity of conclusions drawn from those instruments is an ongoing process (Streiner, 1993). This
study of the Dual-Disorder Treatment Fidelity Scale was confined to the examination of one form of reliability, inter-rater reliability, and three categories of validity, concurrent, predictive, and construct validity. The study was also confined to six state funded programs in Hawai‘i, thus limiting the generalizability of the findings within and especially beyond Hawai‘i. In spite of these limitations, the study as a whole provides evidence that the scale can be implemented reliably in Hawai‘i AMHD programs. It also provides initial evidence for the validity of conclusions drawn from the item level ratings. Information from item level ratings (e.g., program strengths and needs for training or resources) may be particularly useful to programs in the initial stages of integrated model implementation. Making decisions (e.g., programming, policy, or funding decisions) on the basis of total fidelity scores or implementation ratings, however, is not recommended at this stage in the process of scale validation.
<table>
<thead>
<tr>
<th>Item</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Identification of Dual-Disorder Clients</td>
<td>No systematic method to identify dual-disorder clients.</td>
<td>Clients are screened for dual-disorder problems when substance abuse is suspected, but the screening is inconsistent.</td>
<td>Clients are screened for dual-disorder problems when substance abuse is suspected, and the screening process is standardized.</td>
<td>Every new client admitted to the agency is screened to identify dual-disorder problems, but the screening is inconsistent.</td>
<td>Every new client admitted to the agency is systematically screened to identify dual-disorder problems.</td>
</tr>
<tr>
<td>2. Integrated Assessment of Dual-Disorder Clients</td>
<td>Neither disorder assessed with specificity.</td>
<td>One disorder assessed with some specificity and the other disorder not assessed.</td>
<td>Both disorders assessed, with good specificity of one disorder.</td>
<td>Both disorders assessed with good specificity but no integration.</td>
<td>Both disorders assessed with good specificity, and integration is mentioned.</td>
</tr>
<tr>
<td>3. Comprehensive Mental Health Assessment (See protocol for list of areas)</td>
<td>No comprehensive assessment available (fewer than half of the areas assessed).</td>
<td>Comprehensive assessment available (at least half of the areas assessed), but not updated (&gt;1 year old).</td>
<td>Comprehensive and updated assessment covers 5-6 areas.</td>
<td>Comprehensive and updated assessment covers 7-8 areas.</td>
<td>Comprehensive and updated assessment covers at least 9 areas.</td>
</tr>
<tr>
<td>4. Comprehensive Substance Abuse Assessment (See protocol for list of areas)</td>
<td>No comprehensive assessment available (fewer than half of the areas assessed).</td>
<td>Comprehensive assessment available (at least half of the areas assessed), but not updated (&gt;1 year old)</td>
<td>Comprehensive and updated assessment covers 3-4 areas.</td>
<td>Comprehensive and updated assessment covers 5-6 areas.</td>
<td>Comprehensive and updated assessment covers 7-8 areas.</td>
</tr>
<tr>
<td>5. Integrated Treatment Plan</td>
<td>Both disorders addressed in &lt;25% of treatment plans.</td>
<td>Both disorders addressed in 25 - 75% of the treatment plans.</td>
<td>Both disorders addressed in &gt;75% of plans, but plans lack specificity and integration.</td>
<td>Both disorders are addressed in &gt;75% of plans, plus good specificity.</td>
<td>Both disorders addressed in &gt;75% of plans, plus good specificity and integration.</td>
</tr>
<tr>
<td>Item</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>------</td>
<td>------------------------------------------------------------------</td>
<td>------------------------------------------------------------------</td>
<td>------------------------------------------------------------------</td>
<td>------------------------------------------------------------------</td>
<td>------------------------------------------------------------------</td>
</tr>
<tr>
<td>6. Integrated Crisis Plan</td>
<td>Fewer than 25% of dual-disorder clients have a written crisis plan (for either disorder).</td>
<td>25-79% of dual-disorder clients have a written crisis plan for at least one disorder.</td>
<td>Crisis plan present for 80% of dual-disorder clients, but plans target both substance abuse and mental illness in &lt;25% of the charts.</td>
<td>Crisis plan present for 80% or more of dual-disorder clients, and plans target both substance abuse and mental illness up to &gt;75% of charts.</td>
<td>Crisis plan present for 80% or more of dual-disorder clients, and plans target both substance abuse and mental illness &gt;75% of charts.</td>
</tr>
<tr>
<td>7. Integration of Services</td>
<td>Clinicians who treat both disorders see &lt;25% of dual-disorder clients.</td>
<td>Clinicians who treat both disorders see 25-49% of dual-disorder clients.</td>
<td>Clinicians who treat both disorders see 50-69% of dual-disorder clients.</td>
<td>Clinicians who treat both disorders see 70-89% of dual-disorder clients.</td>
<td>Clinicians who treat both disorders see 90% or more of dual-disorder clients.</td>
</tr>
<tr>
<td>8. Comprehensiveness of Services (# accessible to dual-disorder clients within 2 months):</td>
<td>One or none of listed services are accessible.</td>
<td>Two of these services are accessible.</td>
<td>Three of these services are accessible.</td>
<td>Four of these services are accessible.</td>
<td>Five of these services are accessible.</td>
</tr>
<tr>
<td></td>
<td>• Residential</td>
<td>• Family</td>
<td>• Illness Self-Management</td>
<td>• Vocational</td>
<td>• ACT or ICM</td>
</tr>
<tr>
<td>9. Time-Unlimited Services</td>
<td>Specific time limits of up to a year are placed on dual-disorder services.</td>
<td>Specific time limits of 1-2 years are placed on dual-disorder services.</td>
<td>Specific time limits of more than 2 years are placed on dual-disorder services.</td>
<td>No specific time limits on dual-disorder services, but there is pressure for clients to move out of these services.</td>
<td>No specific time limits on dual-disorder services, and no pressure for clients to move out of these services.</td>
</tr>
<tr>
<td>Item</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>----------------------------------------</td>
<td>----------------------------------------</td>
<td>----------------------------------------</td>
<td>----------------------------------------</td>
<td>----------------------------------------</td>
</tr>
<tr>
<td>10. Outreach Capability (i.e., home and community visits for clients who need it)</td>
<td>Outreach rarely if ever done.</td>
<td>Outreach done for emergency purposes.</td>
<td>Outreach is done for emergency purposes and medication and symptom monitoring.</td>
<td>Outreach is done for emergency purposes, medication/ symptom monitoring, and to attend to basic needs (food, clothing, shelter).</td>
<td>Outreach is done to develop and maintain therapeutic alliance, in addition to the other purposes specified.</td>
</tr>
<tr>
<td>11. Client-to-Clinician Ratio (excluding psychiatrist)</td>
<td>Over 50 clients per clinician.</td>
<td>41-50 clients per clinician.</td>
<td>31-40 clients per clinician.</td>
<td>21-30 clients per clinician.</td>
<td>20 or fewer clients per clinician.</td>
</tr>
<tr>
<td>12. Integrated Group Treatment for Dual-Disorders</td>
<td>No groups are offered for dual-disorder clients.</td>
<td>Groups are offered for only one of the two disorders.</td>
<td>Separate groups for each disorder are offered but no integration of the disorders in the groups.</td>
<td>Separate groups for each disorder, but some discussion of the other disorder does take place.</td>
<td>Integrated groups where both disorders are the focus of the treatment.</td>
</tr>
<tr>
<td>13. Group Treatment</td>
<td>No groups.</td>
<td>One group type.</td>
<td>Two group types.</td>
<td>Three group types.</td>
<td>Four or more group types.</td>
</tr>
<tr>
<td>• Education</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Persuasion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Active Treatment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Combined Persuasion and Active Treatment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Social Skills Training</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Relapse Prevention</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>14. Individual Motivational Interviewing</td>
<td>Motivational interviewing is done with &lt;20% of dual-disorder clients.</td>
<td>Motivational interviewing is done with 20-39% of dual-disorder clients.</td>
<td>Motivational interviewing is done with 40-59% of dual-disorder clients.</td>
<td>Motivational interviewing is done with 60-79% of dual-disorder clients.</td>
<td>Motivational interviewing is done with 80% of dual-disorder clients.</td>
</tr>
<tr>
<td>15. Individual Cognitive-Behavioral Counseling (CBC)</td>
<td>CBC is done with &lt;20% of dual-disorder clients.</td>
<td>CBC is done with 20-39% of dual-disorder clients.</td>
<td>CBC is done with 40-59% of dual-disorder clients.</td>
<td>CBC is done with 60-79% of dual-disorder clients.</td>
<td>CBC is done with 80% of dual-disorder clients.</td>
</tr>
<tr>
<td>16. Family Interventions</td>
<td>Fewer than 20% of families in weekly contact with clients are receiving services.</td>
<td>20 – 39% of families in weekly contact with clients are receiving services.</td>
<td>40 – 59% of families in weekly contact with clients are receiving services.</td>
<td>60% or more of families in weekly contact with clients are receiving services, but no standard curriculum or manual is used.</td>
<td>60% or more of families in weekly contact with clients are receiving services, and a manual or curriculum is used.</td>
</tr>
<tr>
<td>17. Pharmacological Treatment of Mental Illness</td>
<td>Pharmacological treatment provided to &lt;25% of clients with an active substance use disorder.</td>
<td>Pharmacological treatments provided for 25-49% of clients with an active substance use disorder.</td>
<td>Pharmacological treatments provided for 50-79% of clients with an active substance use disorder.</td>
<td>Pharmacological treatments provided for 80-89% of clients with an active substance use disorder.</td>
<td>Pharmacological treatments provided for 90% of clients with an active substance use disorder.</td>
</tr>
<tr>
<td>18. Self-Help Liaison</td>
<td>No referral of dual-disorder clients to self-help in the community.</td>
<td>Occasional referral of clients to self-help.</td>
<td>Clients are routinely referred to self-help groups, but clinicians do not attend these groups with clients.</td>
<td>Clients routinely referred to self-help groups, and clinicians sometimes attend these groups with clients.</td>
<td>Clients routinely referred to self-help groups, clinicians frequently attend these groups with clients, and agency has liaison.</td>
</tr>
<tr>
<td>Item</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>----------------------------------------</td>
<td>----------------------------------------</td>
<td>----------------------------------------</td>
<td>----------------------------------------</td>
<td>----------------------------------------</td>
</tr>
<tr>
<td>19. Stage-Wise Treatment</td>
<td>Interventions contrary to stages.</td>
<td>Some (&lt;40%) interventions are consistent with clients' motivational stages.</td>
<td>Many (40-59%) interventions consistent with motivational stages.</td>
<td>Most (60-79%) interventions consistent with motivational stages.</td>
<td>Most (80% or more) interventions are consistent with motivational stages and <em>explicitly</em> reflect stages of treatment.</td>
</tr>
<tr>
<td>20. Reducing Negative Consequences</td>
<td>No awareness of this principle.</td>
<td>Staff shows little awareness of principle, although some interventions are consistent with it.</td>
<td>Agency endorses principle, but staff only occasionally acts on this principle.</td>
<td>Agency endorses principle, but staff does not routinely implement strategies consistent with this philosophy.</td>
<td>Staff routinely implements strategies consistent with this philosophy.</td>
</tr>
</tbody>
</table>
## APPENDIX B

**Chart Review Worksheet**

<table>
<thead>
<tr>
<th>Program</th>
<th>Date</th>
<th>Rater</th>
</tr>
</thead>
</table>

### Item 1
Screen? (y/n)

### Item 2
Rating (1-5)

### Item 3
# of areas (1-11)
Updated? (y/n)

### Item 4
# of areas (1-8)
Updated? (y/n)

### Item 5
Both d/o? (y/n)
Specific? (y/n)
Integrated? (y/n)

### Item 6
Plan? (y/n)
Target both? (y/n)

### Item 8 Comments

### Item 16 Comments

<table>
<thead>
<tr>
<th>Chart #</th>
<th>Chart #</th>
<th>Chart #</th>
<th>Chart #</th>
<th>Chart #</th>
</tr>
</thead>
</table>

### Item 1
Screen? (y/n)

### Item 2
Rating (1-5)

### Item 3
# of areas (1-11)
Updated? (y/n)

### Item 4
# of areas (1-8)
Updated? (y/n)

### Item 5
Both d/o? (y/n)
Specific? (y/n)
Integrated? (y/n)

### Item 6
Plan? (y/n)
Target both? (y/n)

### Item 8 Comments

### Item 16 Comments

Item 1 total # of charts = _______ (______ %)  
Item 2 mean rating = _______  
Item 3 mean # of areas = _______  
Item 4 mean # of areas = _______  
Item 5 total # of plans both d/o = _______ (______ %)  
Item 5 # specific = _______ (______ %)  
Item 5 # integrated = _______ (______ %)  
Item 6 total # charts with plan = _______ (______ %)  
Item 6 # target both = _______ (______ %)
APPENDIX C

Consumer Survey

Please indicate your agreement/disagreement with each of the following statements by circling the response that best fits your opinion.

As a direct result of services I received:

1. I deal more effectively with daily problems........................................ 1 2 3 4 5
2. I am better able to control my life.................................................... 1 2 3 4 5
3. I am better able to deal with crisis................................................... 1 2 3 4 5
4. I am getting along better with my family.......................................... 1 2 3 4 5
5. I do better in social situations........................................................... 1 2 3 4 5
6. I do better in school and/or work....................................................... 1 2 3 4 5
7. My housing situation has improved.................................................. 1 2 3 4 5
8. My symptoms are not bothering me as much.................................... 1 2 3 4 5
9. I know more about how my substance use affects my mental illness. 1 2 3 4 5
10. I feel better able to stay clean and sober......................................... 1 2 3 4 5

Please answer the remaining questions by circling the response that best fits your opinion.

11. How do you feel about your life in general? (CIRCLE ONE)
    1 2 3 4 5 6
    Terrible Unhappy Mostly Dissatisfied Satisfied Pleased Delighted

12. In general, would you say your physical health is (CIRCLE ONE):
    1-Excellent 2-Very Good 3-Good 4-Fair 5-Poor

13. During the last 6 months, have you been arrested? (CIRCLE ONE)
    1-Yes 2-No 3-Don't Know

14. During the last 6 months, were you a victim of any crimes, such as assault, rape, mugging, or theft of your property or money? (CIRCLE ONE)
    1-Yes 2-No 3-Don't Know

15. During the past 6 months, did you generally have enough money each month to cover:
    a. Food?................................................................. 1-Yes 2-No
    b. Clothing?.......................................................... 1-Yes 2-No
    b. Housing?.......................................................... 1-Yes 2-No
    c. Medical Care?................................................... 1-Yes 2-No
    d. Traveling around the island for things like shopping, medical appointments, or visiting friends or relatives?......... 1-Yes 2-No
APPENDIX D

Staff Survey

The evidence-based practice for consumers who have dual diagnosis is integrated mental health and substance abuse treatment. Your program may or may not follow the integrated treatment model. From what you know about integrated treatment, please indicate whether you think each statement is True or False of the integrated model (not of your program).

T  F  1.  At admission for services, all consumers are screened for substance use with standardized (i.e., published) instruments.

T  F  2.  Mental health and substance abuse treatment services are provided by separate specialized clinicians.

T  F  3.  Mental health needs are comprehensively assessed and updated at least yearly.

T  F  4.  Substance use behaviors and patterns are comprehensively assessed and updated at least yearly.

T  F  5.  Assessment includes behavioral details of how one disorder influences the other disorder.

T  F  6.  Crisis plans that address mental health or substance abuse related crises are developed based on consumer need.

T  F  7.  Separate treatment plans address mental health and substance abuse treatment needs.

T  F  8.  Residential services, family services, illness self-management, assertive or intensive case management, and vocational services are all available within 2 months of referral.

T  F  9.  The provision of services is limited to 3 years.

T  F  10.  Services are provided outside of the clinic for new but not previously admitted consumers.

T  F  11.  The ratio of consumers to clinicians within the program is 20 or fewer per clinician.

T  F  12.  Groups designed to address both mental health and substance abuse problems at the same time are provided.

T  F  13.  Education, persuasion, active treatment, social skills training, and relapse prevention groups are all provided.

T  F  14.  Motivational interviewing is systematically done with all consumers with dual disorders.

T  F  15.  Cognitive-behavioral counseling for substance abuse is provided to consumers with dual disorders.

T  F  16.  When possible consumer family members are educated about dual disorders and trained in communication, problem solving, stress reduction, and relapse prevention skills.

T  F  17.  Psychotropic medications are not prescribed for the treatment of mental illness in consumers who are actively abusing substances.

T  F  18.  Consumers are connected with consumer run self-help groups such as AA or NA.

T  F  19.  Relapse prevention skills are taught to consumers from the start of their treatment.

T  F  20.  Efforts are made to reduce the negative consequences of substance abuse via methods other than substance use reduction itself.
APPENDIX E
Staff Informed Consent Form

CASE MANAGER AGREEMENT TO PARTICIPATE IN

Reliability and Validity Evidence for the Dual-Disorder Treatment Fidelity Scale

Principal Investigator: Diane Wilson
2800 Woodlawn Dr. Suite #120
Honolulu, HI 96822
dsimonds@hawaii.edu
(808) 539-3939

Description of Project: This is a research study of integrated mental health and substance abuse (dual diagnosis) treatment programs. Integrated treatment has been shown effective in helping consumers achieve recovery. A measurement tool has been developed to determine how well programs are providing integrated treatment. The purpose of this study is to determine if this measurement tool is useful for evaluating Hawai‘i Adult Mental Health Division programs.

Invitation to Participate: You are being invited to take part in this research study because you provide services to consumers who have dual diagnosis at ___________ which is one program we are studying. You do not have to be part of the study and your participation will not affect your employment with the center. Before agreeing to be part of this study, please read and/or listen to the following information carefully and feel free to ask any questions you might have.

Description of Procedures: If you decide to be in this study, you will be asked to complete a questionnaire. The time needed to complete the questionnaire is about 20 minutes. The questionnaire will include questions about your knowledge of integrated treatment.

Risks and Inconveniences: There is a possibility that responding to the questionnaire may make you feel uncomfortable or inconvenienced. If this happens you can choose not to answer certain questions or you can choose to stop your participation.

Benefits: This study is not being done to help you, personally. What we learn from you may help others in the future by making services and programs better.

Confidentiality: You will not be asked to write your name on the questionnaire. Your name will not be recorded. Any information obtained from the questionnaire will be kept strictly confidential and will not be accessible to anyone other than the primary investigator and her supervisor. Your identity will not be linked to any data which means than you will not be identified in anything published as a result of this project.
**Voluntary Participation:** Your participation in this study is entirely voluntary. Refusal to participate in any part of the study will not affect your employment with the center. You can stop being in the study at any time without affecting your employment with the center.

**Questions:** Please feel free to ask any questions about anything that seems unclear to you. You can ask the principal investigator questions about this project at any time. You can also contact Annette Crisanti, Ph.D., the supervisor of this project, at any time to ask questions about the research. Her phone number at the Adult Mental Health Division is (808) 539-3959. You may also contact the Committee on Human Studies, University of Hawaii, 2540 Maile Way, Honolulu, Hawaii, 96822. Phone (808) 956-5007 if you feel that you have been treated unfairly in any way relating to this study or have any complaints or questions regarding your rights as a participant in this study.

**Consent:** By agreeing to stay and complete the questionnaire you are giving your consent to participate in the study.
APPENDIX F

Consumer Informed Consent Form

CONSUMER AGREEMENT TO PARTICIPATE IN

Reliability and Validity Evidence for the Dual-Disorder Treatment Fidelity Scale

Principal Investigator: Diane Wilson
2800 Woodlawn Dr. Suite #120
Honolulu, HI 96822
dsimonds@hawaii.edu
(808) 539-3939

Description of Project: This is a research study of integrated mental health and substance abuse (dual diagnosis) treatment programs. Integrated treatment has been shown effective in helping consumers achieve recovery. A measurement tool has been developed to determine how well programs are providing integrated treatment. The purpose of this study is to determine if this measurement tool is useful for evaluating Hawai‘i Adult Mental Health Division programs.

Invitation to Participate: You are being invited to take part in this research study because you are in a dual diagnosis group at which is one program we are studying. You do not have to be part of the study to stay in this program, or to get any other services you may be getting for mental health or substance abuse. Before agreeing to be part of this study, please read and/or listen to the following information carefully and feel free to ask any questions you might have.

Description of Procedures: If you decide to be in this study, you will be asked to complete a questionnaire. The time needed to complete the questionnaire is about 20 minutes. The questionnaire will include questions about how your life is going in terms of living situation, health, finances, mental health, substance use, and treatment.

Risks and Inconveniences: There is a possibility that some of the questions in the questionnaire may make you feel uncomfortable. If this happens you can choose not to answer certain questions or you can choose to stop your participation. If you wish, we can call your clinician, another staff member or concerned others to make sure you have someone to talk with about your feelings.

Benefits: This study is not being done to help you, personally. What we learn from you may help others in the future by making services and programs better.

Confidentiality: You will not be asked to write your name on the questionnaire. Your name will not be recorded. Any information obtained from the questionnaire will be kept strictly confidential and will not be accessible to anyone other than the primary investigator and her supervisor. Your identity will not be linked to any data which means than you will not be identified in anything published as a result of this project.
Compensation: You will receive a $5.00 gift certificate to Foodland or Sack 'n Save Foods for your participation.

Voluntary Participation: Your participation in this study is entirely voluntary. Refusal to participate in any part of the study will not affect the services that you are receiving now or in the future. You can stop being in the study at any time without affecting your treatment or other services.

Questions: Please feel free to ask any questions about anything that seems unclear to you. You can ask the principal investigator questions about this project at any time. You can also contact Annette Crisanti, Ph.D., the supervisor of this project, at any time to ask questions about the research. Her phone number at the Adult Mental Health Division is (808) 539-3959. You may also contact the Committee on Human Studies, University of Hawaii, 2540 Maile Way, Honolulu, Hawaii, 96822. Phone (808) 956-5007 if you feel that you have been treated unfairly in any way relating to this study or have any complaints or questions regarding your rights as a participant in this study.

Consent: By agreeing to stay and complete the questionnaire you are giving your consent to participate in the study.
REFERENCES


