Follow-up Difficulty and Retention: Evaluating Potential Attrition Bias in a Longitudinal Substance Abuse Treatment Study

Erin N. Harrop

A thesis

submitted in partial fulfillment of the

requirements for the degree of

Master of Social Work

University of Washington

2014

Committee:

Elizabeth A. Wells

Sarah Bowen

Program Authorized to Offer Degree:

Social Work

DEDICATION

This work is dedicated to my top 20 list of phenomenal teachers. To Ms. Kristi Gustafson-Lin, Mrs. Kay Law, and Mrs. Nancy Gustafson, who believed in me in 1st, 2nd, 3rd, 4th, and 5th grade and encouraged me to be a winner, a writer, and an educator. To Ms. Kivonne Tucker, who helped me survive middle school and learn how to pass a volleyball. To Ms. Ellen Caster-Kendall-Eyre, Ms. Lindsay Aegerter, and Ms. Gallagher who taught me to use art, creative writing, and calculus to cope with adolescence. To my Pitzer College favorites: my psychology mentor, Dr. Richard Tsujimoto and English mentor, Dr. Jill Benton. To the best of UW psychology statistics and neuropsychology, Dr. Jaime Diaz and Dr. Geoffrey Loftus; I'm glad I never have to take another one of your exams ever again. To the prevention profs at the UW School of Social Work, Dr. Jenn Stuber, Dr. Kevin Haggerty, and Dr. Rico Catalano, who each taught with poise, passion, and intentionality. To Dr. Jenny Romich and James DeLong, MSW, for challenging me. To Sarah, Seema, and Joel for the opportunities, guidance, and encouragement. In loving memory of G. Alan Marlatt, my first mentor in the world of substance abuse research. It seems as if most students are blessed with a couple teachers whose dedication, investment, and compassion inspire them for a lifetime. I have been richly blessed and am in gratitude to you all.

ACKNOWLEDGEMENTS

Thanks to my thesis committee: Dr. Betsy Wells (chair), Dr. Sarah Bowen, and Dr. Rico Catalano. You have each inspired me in different ways, and I am grateful for the mentorship, encouragement, and challenges that you have presented to me.

Thanks to the relentless MBRP team, led by Sarah, Seema, and Alan at the ABRC and CSHRB. Working with this research team has been a wonderful experience, and I am grateful for our team—Joel, Haley, Matt, Michelle, David, Beth, Susan, Katie, Alan, Kathleen Lustyk, Sharon, Neha, Sara, Sarah, and Betsy. I also want to express my gratitude for the staff at RCKC and the remarkable clients whom I met there. Thanks also to the National Institute of Health/National Institute on Drug Abuse (NIH/NIDA; PIs: G. Alan Marlatt, Sarah Bowen) which provided participant incentives, partial support for researchers, and all costs associated with conducting the study. NIH supplemental grants (UL1RR025014 and 1K1BDA031464-01), National Institute of Alcohol Abuse and Alcoholism F31 Fellowship AA 019608-03, and NIDA grant 5ROIDA025764-02 provided support for researcher salaries.

Finally, I would like to thank my family. Thank you, Jared, for supporting me throughout my academic journey, often at the expense of time that we could have spent together. I would also like to thank Ty and Ian for hanging out with me during family homework time. Mom and Dad, you have believed in me from the start, and I am grateful for your support. Britt, Ryan, CoCo, Megan, Michael, Nataly, Dawn, Heather, Colin, KT, and Sid—I appreciate you all being patient with me as I devote so much of my time to my program.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

University of Washington

Abstract

Follow-up Difficulty and Retention: Evaluating Potential Attrition Bias in a Longitudinal Substance Abuse Treatment Study

Erin N. Harrop

Chair of the Supervisory Committee:
Elizabeth A. Wells, Ph. D.
Social Work

BACKGROUND: Participant attrition threatens validity in longitudinal research. Non-random attrition creates "attrition bias," that could compromise result generalizability. This study explores the impact of attrition on a longitudinal substance use study.

METHODS: Easy-to-contact (ETC) participants, difficult-to-contact (DTC) participants, and study non-completers were compared on demographic characteristics. Following this, outcome variables for ETC and DTC participants were compared, using DTC participants as theoretical proxies for non-completers.

RESULTS: DTC participants and non-completers differed on few demographic characteristics. However, DTC participants were more likely to have used a substance during the follow-up period compared to ETC participants.

DISCUSSION: These results suggest that attrition bias may result in lower reports of relapse. This bias may not be a major threat to validity in etiological studies with at least 65% retention. However, additional testing for attrition bias in efficacy studies is important, as differential attrition by condition (particularly among substance users) could threaten conclusion validity.

CHAPTER 1

LITERATURE REVIEW

Substance abuse is a major public health concern in the United States. National studies indicate that up to 14.6% of all Americans will suffer from a substance use disorder (SUD) at some point in their lifetimes (Kessler, et al., 2005). Furthermore, among individuals with SUDs, the median age of onset is estimated to be 20 years old (Kessler, et al., 2005). Thus, substance use is a major issue for our youth, and an important problem in both adolescence and adulthood.

In addition to these high prevalence rates, substance use also has considerable effects on society. Substance use is associated with lower income status and homelessness (Bassuk, Buckner, Perloff & Bassuk, 1998). Additionally, drug users are more likely to contract HIV compared to non-users (Battjes, Leukefeld, Pickens, & Haverkos, 1988; Teplin, et al., 2005). Those with SUDs are also more likely to suffer from other mental disorders including mood disorders, anxiety disorders, antisocial personality disorders, and conduct disorders (Merikangas, et al, 1998). Given the high impact on society, and the high cost to the affected individual, it is imperative to develop and assess effective interventions for this population. This thesis examines the methodology of one such treatment efficacy study, by exploring threats to its internal and external validity due to participant drop-out.

Longitudinal studies are essential to developing and improving SUD interventions, because they allow the researcher to examine changes in participants over time (Royce, 2008). Repeatedly examining the same participants over a follow-up period allows researchers to observe the long-term effectiveness of an intervention, as well as the temporal order of the

relationships among the variables of interest (immediate, short, medium, and long term outcomes). Thus, this type of research is particularly well-suited to treatment efficacy studies. However, despite the benefits, longitudinal studies also present researchers with the challenge of tracking the same participants over a longer period of time and locating and interviewing those same participants over multiple time points to ensure the completeness of the data set. "Retention" refers to the ability of the study to maintain consistent engagement with participants throughout the duration of the study. Conversely, "attrition" refers to the loss of participants (and therefore, data) over time. Maintaining high participant retention rates over a long period of time is essential for achieving the research and clinical benefits of longitudinal designs.

In conducting longitudinal research with SUD populations, Bootsmiller and colleagues (1998) note that one of the primary difficulties in maintaining high study retention is the highly mobile, unstable, and transient lifestyles of many participants. Hansen, Tobler and Graham (1990) and Ziek, Beardsley, Deren, and Tortu (1996) further point out that difficulty is influenced by the chaotic patterns that often characterize the lives of individuals abusing substances. Often, participants are reluctant to be contacted due to the fear of unwanted pursuits by police, creditors, probation officers, and street associates. This reluctance can further compromise the ability of researchers to track participants (Bootsmiller, et al., 1998). Brown-Peterside and colleagues (2001) noted other characteristics contributing to retention difficulties including: have a lower income, suicidality, and homelessness. Goshin and Byrne (2012) further identify that characteristics of the interviewer, such as race, may impact retention within a study since participants may relate best to an interviewer of a similar ethnic background. In sum, a variety of factors make retention in longitudinal SUD studies challenging.

The Problem of Attrition

Participant retention is critical to a longitudinal study's success and validity (Royce, 2008). Study validity can be divided into two sub-categories: internal and external. Internal validity pertains to the likelihood that the intervention's observable effects can indeed be attributed to that specific intervention (Flick, 1988). Thus it is essential to determine whether or not other variables (aside from the ones intended) influence study outcomes. The higher the internal validity of the study, the more confident the researcher can be in the efficacy of the given intervention.

Attrition is a threat to the internal validity of a study, because it is impossible to ascertain the outcomes of missing participants. Thus, when attrition is high, it is possible that the unavailable data points may contradict the findings of a study based only on those interviewed, thus making the study's conclusions less valid. This is particularly important when a study's findings are very close to the level of statistical significance; a little more data in either direction could change the conclusions of the study (Flick, 1988). Further, attrition that is differential between conditions may threaten the conclusions of a study, even if only a small number of participants are lost to follow-up. If people lost to attrition in the experimental group differ from those lost to attrition in the control group, (e.g. a higher percentage of those with mental health problems are not retained in the intervention condition), this could affect study outcomes and intervention efficacy findings. Thus, when those lost to attrition differ systematically from the rest of the sample, the effect of attrition bias is greater. To the extent that attrition is random, attrition bias is minimized.

This threat to internal validity can result in two types of bias: inflation bias and suppression bias. Inflation bias makes a truly ineffective intervention seem as if it is effective; suppression bias makes it appear that a truly effective intervention either harms or does not impact a participant (Graham, 2009). Thus, without adequate retention, studies risk inflating or suppressing the demonstrable effects of a given intervention (Cotter, Burke, Loeber, & Navratil, 2002). Though there are many statistical solutions available to help account for missing data, none represent perfect solutions to the problem of attrition (Flick, 1988).

External validity refers to the extent to which the findings of a particular study can be generalized to a given population (Flick, 1988). Attrition is a threat to the external validity of a study, because it is possible for those lost to attrition to differ systematically from those who complete a study. In such cases, the results of the study can only generalize to study completers—as opposed to the entire sample that was studied. Thus, in order to ensure that a study's results will be generalizable, it is important to ensure high levels of retention.

Taken together, these threats to the internal and external validity of a study, due to attrition, are called *attrition bias* (or non-response bias; Weisberg, 2009). The amount of attrition bias depends on three things: the attrition rate, the extent to which non-respondents differ from respondents in baseline variables and outcomes, and the extent to which attrition is differential between study conditions (Weisberg, 2009). According to Weisberg (2009), error due to attrition bias occurs when attrition and the study's variables have a common cause or relationship. Thus, it becomes impossible to infer whether or not demonstrable effects between located participants are due to the intervention or to another cause. It is important to note however that attrition bias

does not necessarily translate into meaningful rates of error. When study completers and non-responders do not differ in systematic ways, the error resulting from this attrition bias will be small (Weisberg, 2009). Thus it is important to explore the extent to which attrition bias may compromise the results of a particular study, before concluding that attrition (on its own) invalidates the findings. Though a traditional view of attrition bias required researchers to go to "heroic" efforts to reduce non-response, Weisberg suggests a revisionist view: lower retention rates may still lead to accurate results provided that these lower response rates do not translate into meaningful bias.

Attrition Rates in Longitudinal Substance Abuse Studies

Robinson, Dennison, Wayman, Pronovost & Needham (2007) estimated that within general longitudinal healthcare research studies, retention rates averaged approximately 86 percent. Unfortunately, however, retention rates for substance abuse studies often remain much lower. In a meta-analysis of 85 longitudinal SUD studies, Hansen, et al. (1990) found that attrition is highest at the first follow-up point, and gradually decreases after that. Hansen and colleagues (1990) reported the following mean follow-up rates across the studies: 81.4% retention after three months, 78.3% at six months, 73.4% at one year, 71.8% at 2 years, and 67.5% at three years. In response to these norms, Hansen et al. (1990) urges researchers to take steps to reduce attrition and document ways to increase participant retention. In the years that followed Hansen's review, a series of research reports was published suggesting strategies to help increase follow-up rates for "non-traditional" populations, and as a result of this collaborative research, follow-up rates have risen (Wright, Allen, & Devine, 1995). Also, there is

movement in the field to transition away from simply offering follow-up strategies, towards creating replicable models for consistently producing high follow-up rates in studies (Scott, 2004).

Taking into account the difficulty of follow-up within SUD populations, and the potential threat of attrition bias to invalidate study findings, one question persists: What is an acceptable retention rate? Obviously, higher rates are better, but Hansten, Downey, Rosengren, and Donovan (2000) pointedly ask, "When is 'enough' enough?" There is considerable debate in the literature. Polich et al., (1980) suggests a cut off of 60%, arguing that the gains in accuracy after 60-70% of the data are collected are insubstantial. The "rule of thumb" for many has been a retention rate of 70%, leading some researchers to affirm this standard (Digiusto, et al., 2006) and others to insist it must be higher (Scott, 2004). Others' findings suggest, that LaPorte et al.'s (1981) guideline of 65-80% is reasonable, and that the additional gains in accuracy with higher retention rates are not necessarily worth the effort and cost required (Hansten, et al., 2000). Generally, the most conservative researchers recommend at least an 80% retention rate (Nemes, Wish, Wraight, & Messina, 2000), while others suggest that lower rates may also be valid (Hansten, et al., 2000).

Strategies for Reducing Attrition

In order to address this problem of retention, researchers have been exploring various types of tracking strategies in order to improve follow-up rates. Baseline strategies involve suggestions that can be implemented at a baseline interview to proactively gather information to help locate individuals; similarly, follow-up strategies involve suggestions that can be utilized

while tracking participants after a baseline appointment (Lankenau, Sanders, Hathazi, & Jackson Bloom, 2010).

Based on extant literature, the following suggestions are posed as strategies for enhancing tracking of participants during follow up: obtain detailed locator information (Digiusto Panjari, Gibson, Rea, 2006; Scott, 2004; Twitchell, Hertzog, Klein, & Schuckit, 1992; Wright, et al., 1995); obtain contact information for family members, friends, and utilized service agencies (Haggerty, et al., 2008; Nemes, et al., 2000; Ribisl, et al., 1996); educate participants about the goal of the research and how their information will be used (Scott, 2004; Wright, et al., 1995); create a recognizable project identity (Jarlais, Perlis, & Settembrino, 2005; Ribisl, et al., 1996; Robinson, Dennison, Wayman, Pronovost, & Needham, 2007); use electronic transfers for payment (Jarlais, et al., 2005; Lankenau, et al., 2010); utilize electronic data bases and records (Haggerty, et al., 2008; Lyons, et al., 2004; Nemes, et al., 2000); send reminder mail and phone calls (Coen, Patrick, & Shern, 1996; Scott, 2004; Wutzke, Conigrave, Kogler, Saunders, Hall, 2000; Ziek, et al., 1996); offer free treatment as an incentive for follow-up completion (Cottler, Compton, Ben-Abdallah, Horne, & Claverie, 1996); have team meetings to problem solve problem cases (Cottler, et al., 1996; Walton, Romanathan, & Reischl, 1998); increase incentive payments as difficulty increases (Cottler, et al., 1996; Jarlais, et al., 2005); pay a finder's fee to family members (Cottler, et al., 1996); do fieldwork at participant hang-outs (Desmond, Maddux, Johnson, & Confer, 1995; Haggerty, et al., 2008; Lankenau, et al., 2010; Walton, et al., 1998); have a toll-free study number (Bootsmiller, et al., 1998; Jarlais, et al., 2005; Lankenau, et al., 2010); offer phone interviews (Bootsmiller, et al., 1998; Cotter, Burke, Strouthamer-Loeber, & Loeber, 2005; Cottler, et al., 1996; Lankenau, et al., 2010); use email addresses (Lankenau, et

al., 2010); emphasize the importance of retention to staff (Ribisl, et al., 1996; Cottler, et al., 1996); provide interviewer incentives (Capaldi, & Patterson, 1987; Walton, et al., 1998); use electronic databases to track participant data (Cotter, Burke, Loweber, & Navratil, 2002); and offer sufficient financial reimbursement for interviews (Festinger & Dugosh, 2012; Desmond, et al., 1995; Coen, et al., 1996).

Through the consistent application of these follow-up strategies, some SUD researchers have managed to achieve retention rates above 90% (Bale, Arnoldussen, & Quittner, 1984; Desmond, et al., 1995; Cottler, et al., 1996; Haggerty, et al., 2008; Hansten, et al., 2000; LaPorte, McLellan, Erdlen & Parente, 1981, Nemes, et al., 2000; Polich, Armor & Braiker, 1980; Scott, 2004; Sullivan, Rumptz, Campbell, Eby, & Davidson, 1996; Walton, et al., 1998). This suggests that despite the challenges inherent in tracking this unique population, high retention rates are possible given rigorous follow-up methods and adequate staffing and financial resources. However, it is notable that many of these strategies require significant financial and staff resources, to which not all researchers have access (Cotter, et al., 2005; Haggerty, et al., 2008; Hansten, et al., 2000). Thus, researchers must balance the cost of such expensive measures with the benefit of the added accuracy in reduced attrition.

Contact Difficulty and Potential Attrition Bias

Several researchers have begun exploring the potential benefits of higher retention rates by examining contact difficulty and its effect on potential attrition bias. When trying to study the extent to which attrition biases a study, there are several strategies that researchers employ. Typically researchers perform one of the following comparisons: participants who complete all

follow-ups versus those who miss one or more follow-ups, difficult-to-contact (DTC) participants versus easy-to-contact (ETC) participants, and participants who complete follow-ups versus those who refuse (Ribisl, et al., 1996). The present study will employ the second method, comparing DTC participants and ETC participants. This statistical technique treats DTC participants as proxies for those who were ultimately lost to follow-up, in an effort to examine to what extent these two groups systematically differ (Groves & Couper, 2012).

Research Comparing Difficult-to-Contact and Easy-to-Contact Substance Abusing Populations

Research comparing DTC participants with ETC participants has revealed mixed findings, and there is little consensus in the literature. Generally, these types of studies divide participants into at least two groups based on contact difficulty, and then compare the groups on a number of baseline characteristics and outcome variables.

Defining Difficulty

In reviewing this literature, it is important to note that each study defines "contact difficulty" in a specific, often unique, way. Some researchers use this term to refer to the amount of time it takes to for a participant to complete a follow-up interview (e.g., Bale, Arnoldussen, & Quittner, 1984; Cottler, et al., 1996; Hansten, Downey, Rosengren, & Donovan, 2000), while others refer to it as the number of contact attempts required prior to interview completion (e.g., Digiusto, Panjari, Bibson, & Rea, 2006; LaPorte, McLellan, Erdlen & Parente, 1981; Moos & Bliss, 1978; Nemes, Wish, Wraight, & Messina, 2000; Scott, 2004; Walton, Romanathan, & Reischl). Others consider the cost of location efforts (e.g., Haggerty, et al., 2008), and still others

create some composite variable incorporating multiple measures for contact difficulty (e.g., Polich, Armore & Braiker, 1980). Once contact difficulty has been measured, participants are divided into groups. This too can be a complex process. Most researchers divide participants into two groups (easy and difficult; e.g., Bale, et al., 1984; Cottler, et al., 1996; Hansten et al.; Scott, 2004; Walton, et al., 1998), while others use a more stratified approach, examining multiple groups, and sometimes comparing one group (or groups) to the whole sample (e.g., Digiusto, et al., 2006; Moos & Bliss, 1978; Polich, et al., 1980; LaPorte, et al., 1981; Nemes, et al., 2000). These methodological differences may account for some of the diversity in research findings.

Comparisons of DTC and ETC Substance Abusing Populations

Following these methodological considerations, ETC participants are compared to DTC participants on a number of baseline characteristics and outcome variables. When participants do not differ on baseline characteristic variables, it is generally thought that the sample is more generalizable, ensuring that the diversity of the study sample is preserved. However, Flick (1988) cautions researchers that this is not always the case. When participants do not differ on outcome variables, it is generally thought that the results of the study (regarding treatment outcomes) can be more confidently generalized to the entire sample. However, some researchers have concluded that the small differences in outcomes that have been found are not particularly meaningful (Bale, et al., 1984; Digiusto, et al., 2006; Hansten, et al., 2000; LaPorte, et al., 1981; Polich, et al., 1980). Table 1 summarizes the findings of 10 substance abuse treatment studies that compared ETC participants and DTC participants. Often, it is suggested that participants who are harder to locate have worse treatment outcomes (Moos & Bliss, 1978; Nemes, et al.,

2000; Scott, 2004; Walton, et al., 1998). Still others suggest that the opposite may in fact be true, and that those doing exceptionally well in their recovery may be more difficult to track, because they have successfully remade their lives and no longer use treatment services (Nordström, & Berglund, 1986).

Table 1. Substance abuse treatment studies examining the impact of contact difficulty

Authors, Date	Sample	Follow-Up Rates	Definition of Contact Difficulty	Comparison Groups	Baseline Characteristic Differences	Outcome Differences
Moos & Bliss, 1978	505 persons in 5 different residential drug treatment programs	85% at 6 months	Number of contact attempts prior to follow-up	0 attempts, 1-2 attempts, 3+ attempts	Difficult: higher residential instability, unmarried, less education, younger, less income	Difficulty related to worse alcohol outcomes, behavioral and physical impairment
Polich, Armor & Braiker, 1980	922 male alcoholics enrolled in 8 alcohol treatment centers	91% at 18 months; 85% at 4 years	Amount of fieldwork required, hours of locating time required, number of contacts, length of time to completion of follow-up	8 groups based on cumulative effort required	Not evaluated	Difficulty related to worse outcomes, but no meaningful differences after 60%.
LaPorte, McLellan, Erdlen & Parente, 1981	150 Veteran's Administration men who abuse drugs and alcohol	96% at 6 months	Number of contact attempts prior to follow-up	1 attempt, 2-5 attempts, 6-11 attempts, 12+ attempts	Difficult: some differences in drug use patterns, but no systematic differences related to drug use severity	Difficulty not related to treatment outcome.
Bale, Arnoldussen, & Quittner, 1984	107 Veteran's Administration male inpatients. 76 alcoholics. 31 drug addicts.	94% at 6 months	Length of time to completion of follow-up	1 month or less, more than 1 month	Difficult: criminality, previous treatment history	Difficulty related to use outcomes in non-linear fashion: Daily users and abstainers easier than

						moderate drinkers.
Cottler, Compton, Ben-Abdallah, Horne, & Claverie, 1996	479 persons abusing drugs in a HIV study	96.6% at 18 months	Length of time to completion of follow-up	Interview within 50 day interview window; outside window	Difficult: unemployed	Not evaluated.
Walton, Romanathan, & Reischl, 1998	96 persons in a relapse prevention treatment study. 66% alcoholic. 34% drug abusers.	93% at 3 months; 97% at 6 months	Number of contact attempts prior to follow-up	25th quartile, 75th quartile	Difficult: 11 measured; none correlated	Difficulty related to more substance reuse.
Hansten, Downey, Rosengren, & Donovan, 2000	654 persons seeking substance abuse treatment	90% at 3 months	Length of time to completion of follow-up	Easiest 60% of sample, entire sample	Difficult: younger age	Difficulty related to worse outcomes, but no meaningful differences after 60%.
Nemes, Wish, Wraight, & Messina, 2000	412 persons in residential substance abuse treatment	93% at 19 months	Number of contact attempts prior to follow-up	1 attempt, 2-4 attempts, 5-9 attempts, 10+ attempts	Difficult: younger age, number of prior arrests	Difficulty related to increased cocaine use.
Scott, 2004	Two studies: 2010 adults in substance abuse treatment; 632 adults and adolescents in substance abuse treatment	96.3% at 12 months; 92.9% at 12 months	Number of contacts prior to follow-up	Easiest 70% of sample, hardest 30% of sample	Difficult: Hispanic, weekly marijuana users, those in residential treatment, homeless, criminal justice system	Difficulty related to increased alcohol and cocaine use.
Digiusto, Panjari, Gibson, Rea, 2006	317 heroin users in an Australian drug treatment study	70% at 3 months	Number of contacts prior to follow-up	1 contact, 2 contacts, 3-4 contacts, 5-20 contacts. 1 contact, whole group	Difficult: no differences	Difficulty not related to differences in drug use outcomes.

Typically, a number of baseline characteristics are also examined to differentiate easy and hard-to-reach participants. Watson and Wooden (2009) review a number of longitudinal studies, suggesting a number of baseline characteristics that tend to make contact more difficult. They report that generally males are more difficult to engage than females. Overall, those who are young and very old are more difficult to retain than those who are middle-aged. Racial and ethnic minorities are sometimes more difficult to track, but English-speaking ability is more indicative of difficulty than race alone. Unmarried folks are sometimes more difficult to follow than married folks, and lower educational status is also sometimes associated with contact difficulty.

Watson and Wooden's review about longitudinal studies informs retention research pertaining to longitudinal substance use studies. As demonstrated in Table 1, substance use researchers frequently examined the characteristics Watson and Wooden identified as being correlated to contact difficulty; however, none consistently emerged as strong predictors across multiple studies. Table 2 shows characteristics typically associated with contact difficulty and demonstrates the conflicting findings of substance use studies. Of the 12 baseline characteristics listed, four do not show conflicting results. According to this collective data, it appears that those with more extensive criminal behaviors and those with fewer treatment hours tend to be more difficult to contact. Similarly, these studies do not reflect differences in contact difficulty based on gender or psychiatric commorbidity.

However, based on Watson and Wooden's review, eight categories show conflicting results: residential instability, marital status, educational attainment, age, income, race, drug of

choice, and employment. Overall, the studies did not differ significantly in incentive amount or modes of interviewing (in-person and phone). Several additional study factors could be the cause of these conflicting results: study population characteristics, statistical power, methods of defining difficulty, and study rigor. Only one study reported that marital status, educational attainment, and income were related to contact difficulty (Moos & Bliss, 1978). This study also failed to report detailed participant demographic information and study methods. Due to these factors, marital status, educational attainment and income may in fact be unrelated to contact difficulty. Secondly, Scott (2004) was the only investigator to find three of these characteristics (residential instability, race, and drug of choice) related to contact difficulty. However, it should be noted that Scott's sample (N=2,642) is 2.5 times to 26 times larger than all other studies in this literature review, thus potentially enabling this researcher to detect small differences between groups that require more statistical power. The remaining conflicting results are puzzling. More research is needed to confirm the relationship between these baseline characteristics and contact difficulty.

In addition to listing demographic characteristics that may be associated with contact difficulty, results reported in Table 2 also suggest that substance use outcomes may be related to difficulty. Substance use retention literature is clearly divided on this point. Four studies report that worse substance use outcomes are related to contact difficulty (Moos & Bliss, 1978; Nemes, Wish, Wraight, & Messina, 2000; Scott, 2004; Walton, Romanathan, & Reischl, 1998); four do not find a significant relationship (Digiusto, Panjari, Gibson, Rea, 2006; Hansten, Downey, Rosengren, & Donovan, 2000; LaPorte, McLellan, Erdlen & Parente, 1981; Polich, Armor &

Braiker, 1980). The extent to which contact difficulty relates to substance use outcomes is an important consideration in retention research; further investigation is needed to clarify this issue.

Table 2. Characteristics related to follow-up difficulty in longitudinal substance use studies.

Characteristic	Studies Finding Characteristic Related to Contact Difficulty	Studies Finding Characteristic Unrelated to Contact Difficulty
High Residential Instability	Moos & Bliss, 1978 Scott, 2004 Bale, Arnoldussen, & Quittner, 1984	Hansten, Downey, Rosengren, & Donovan, 2000
Unmarried Marital Status	Moos & Bliss, 1978	Hansten, Downey, Rosengren, & Donovan, 2000 Bale, Arnoldussen, & Quittner, 1984 Nemes, Wish, Wraight, & Messina, 2000
Lower Levels of Education	Moos & Bliss, 1978	Hansten, Downey, Rosengren, & Donovan, 2000 Bale, Arnoldussen, & Quittner, 1984 Nemes, Wish, Wraight, & Messina, 2000 Cottler, Compton, Ben-Abdallah, Horne, & Claverie, 1996
Younger Age	Moos & Bliss, 1978 Hansten, Downey, Rosengren, & Donovan, 2000	Nemes, Wish, Wraight, & Messina, 2000 Bale, Arnoldussen, & Quittner, 1984 Cottler, Compton, Ben-Abdallah, Horne, & Claverie, 1996
Lower Income	Moos & Bliss, 1978	Walton, Romanathan, & Reischl, 1998
Criminality Measures	Scott, 2004 Bale, Arnoldussen, & Quittner, 1984 Nemes, Wish, Wraight, & Messina, 2000	
Race and/or Ethnicity	Scott, 2004	Bale, Arnoldussen, & Quittner, 1984 Nemes, Wish, Wraight, & Messina, 2000 Cottler, Compton, Ben-Abdallah, Horne, & Claverie, 1996
Drug of Choice	Scott, 2004	Bale, Arnoldussen, & Quittner, 1984 Walton, Romanathan, & Reischl, 1998 LaPorte, McLellan, Erdlen & Parente, 1981
Unemployment	Cottler, Compton, Ben-Abdallah, Horne, & Claverie, 1996	Bale, Arnoldussen, & Quittner, 1984 Nemes, Wish, Wraight, & Messina,

		2000
Lower Treatment Hours	Digiusto, Panjari, Gibson, Rea, 2006	
Gender		Hansten, Downey, Rosengren, & Donovan, 2000 Nemes, Wish, Wraight, & Messina, 2000 Cottler, Compton, Ben-Abdallah, Horne, & Claverie, 1996
Psychiatric Comorbidity		Hansten, Downey, Rosengren, & Donovan, 2000 Nemes, Wish, Wraight, & Messina, 2000 Cottler, Compton, Ben-Abdallah, Horne, & Claverie, 1996
Worse Substance Use Outcomes	Moos & Bliss, 1978 Scott, 2004 Walton, Romanathan, & Reischl, 1998 Nemes, Wish, Wraight, & Messina, 2000	Hansten, Downey, Rosengren, & Donovan, 2000 LaPorte, McLellan, Erdlen & Parente, 1981 Digiusto, Panjari, Gibson, Rea, 2006 Polich, Armor & Braiker, 1980

Using Difficult-to-Contact Participants as Proxies for Non-Completers

As described by Groves and Couper (2012), studies sometimes use DTC participants as proxies for participants lost to attrition. Central to these methods, is the assumption that these two groups do not systematically differ in significant ways. Despite the importance of this assumption, few studies have evaluated the extent to which DTC participants differ from non-completers based on baseline characteristics (Nemes, Wish, Wraight, & Messina, 2000). Nemes et al. compared treatment completers to treatment non-completers, in addition to comparing ETC participants to DTC participants on a variety of baseline measures. Of these, only race and treatment completion were found to be significant, leading the authors to conclude that it would be difficult to predict who might fail to complete the study based on baseline and demographic characteristics alone.

No longitudinal substance use study comparing DTC participants to ETC participants was found that reported evaluating the similarity between DTC participants and non-completers. Studies routinely complete analyses of attrition comparing completers to non-completers. However, these studies may miss critical differences between DTC and non-completing participants. While testing the homogeneity of these samples may be less critical for studies achieving very high follow-up rates (where attrition is significantly minimized), it is important methodologically to evaluate this assumption. This present study specifically evaluates this assumption to ensure the validity of using DTC participants as proxies for participants lost to attrition.

CHAPTER 2

INTRODUCTION TO THE PRESENT STUDY

Given the mixed findings of previous substance abuse treatment studies that compare difficult-to-contact (DTC) participants with easy-to-contact (ETC) participants, it is important to continue to explore the potential differences between these two groups. Some studies report findings that suggest some demographic groups are more likely to be lost to follow-up, implying that study findings may not be generalizable to a broader population (Moos & Bliss, 1978; Scott, 2004). More consensus is needed in the literature. Other studies have found that DTC participants experienced worse treatment outcomes than ETC participants, suggesting that study interventions may be less effective with DTC participants or non-completers (Nemes, Wish, Wraight, & Messina, 2000; Scott, 2004; Walton, Romanathan, & Reischl, 1998).

In addition to this literature, there is also a collection of other studies that did not reveal significant differences in demographic characteristics or substance use outcomes of DTC versus ETC participants (Digiusto, Panjari, Gibson, Rea, 2006; Hansten, Downey, Rosengren, & Donovan, 2000; LaPorte, McLellan, Erdlen & Parente, 1981; Polich, Armor & Braiker, 1980). In response to the divergence of the literature, Hansten et al. asserts that failing to consider studies with lower retention rates (below 70 or 80 percent) for publication is not recommended, and "further investigation is needed to shed light on the important topic of when 'enough is enough.'" (p. 1416). Thus, it is hoped that further data collection may build more of a consensus in this literature. Additionally, more data exploring the assumption of similarity between DTC

participants and non-completers would lend support to using this model for future attrition studies.

The present study seeks to add to this literature by exploring the baseline and demographic differences between three groups: ETC participants, DTC participants, and non-completers. This study fills in a gap in the literature by specifically evaluating the potential differences between DTC participants and non-completers in an SUD sample. Following this, ETC participants will be compared to DTC participants on a series of outcome variables.

Study Purpose

The present study involves a secondary data analysis of a randomized controlled trial assessing SUD aftercare treatments in a community sample. Participants received baseline interviews prior to completion of an 8-week aftercare group, and then completed follow-up assessments following the aftercare groups, and at the 2-month, 4-month, 6-month, and 12-month time-points. This thesis examines the potential effects of attrition bias study, by exploring the extent to which attrition occurred in a non-random fashion.

Data analysis occurs in three phases. In Phase 1, study non-completers are compared to both DTC participants and ETC participants on a series of baseline characteristics (gender, age, race, marital status, educational attainment, history of incarceration, income, employment status, treatment hours, and baseline depression, and anxiety scores) and substance use history (drug of choice, baseline Short Inventory of Problems Scale (SIP-II; Miller, Tonigan, & Longabaugh, 1995), and use during the past month of treatment). In Phase 2, non-completers are compared to DTC participants based on the same baseline variables, to determine if DTC participants differ

from non-completers in meaningful ways. The extent that these two groups are similar, lends support to the methodology of using DTC participants as a proxy measure for study participants lost to attrition. The third phase compares DTC participants and ETC participants on substance use outcomes.

The purpose of this study is to evaluate the utility of using different retention rates (65% vs. 80%) in a longitudinal (12-month follow-up) substance use treatment study. This study also seeks to evaluate the methodology of using DTC participants as proxy measures for participants lost to attrition.

Objectives

The objectives of this study are threefold:

1. Determine to what extent non-completers differ from completers who are easy-to-contact and hard-to-contact on baseline characteristics.
2. Evaluate the extent to which hard-to-contact completers serve as an appropriate proxy for non-completers.
3. Examine how difficult-to-contact participants and easy-to-contact participants differ in substance use outcomes.

Hypotheses

Based on extant literature, this study examines the following hypotheses: gender, younger and older age, non-white race, unmarried status, low educational attainment, history of incarceration, lower income, unemployment, lower treatment hours, higher depression and anxiety scores would be correlated with the DTC and non-completer groups, compared to ETC

participants. Baseline substance use measures were also predicted to be related to increased contact difficulty: opiates or cocaine as drug of choice, higher SIP-II score, and use during the past month of treatment. Regarding outcomes, it was predicted that incidence of drinking, heavy drinking, or drug use post study-intervention would be correlated with increased contact difficulty. It was also predicted that DTC participants would use substances more frequently compared to ETC participants.

Due to the exploratory nature of this study and the conflicting findings of past studies, two-tailed significance tests are used.

CHAPTER 3

METHODS

This study is a secondary analysis of data collected from a large treatment efficacy study assessing Mindfulness-Based Relapse Prevention (MBRP). All phases of this study were approved by the Institutional Review Board (IRB) of the University of Washington. No additional information was collected from participants for the present study. Additional IRB approval was obtained to work with the secondary data. Pertinent information regarding this study is reviewed below. Additional details are reported elsewhere (Bowen, et al., 2014). See Appendix A for IRB approval.

MBRP Study

The MBRP study was a longitudinal National Institute of Health-funded study comparing individuals in two treatment programs (MBRP and relapse prevention; RP) to those in a treatment-as-usual control condition, who received standard aftercare at a community substance abuse treatment center. This study recruited 286 participants from two different sites to complete a randomized controlled trial of the MBRP treatment compared to RP and standard aftercare groups.

Prior to participation, all participants provided informed consent. Participants were recruited through posters at the treatment agency in addition to periodic recruitment presentations at aftercare groups. After providing verbal consent, interested participants completed an interview for eligibility. Participants were required to: (a) be over the age of 18, (b)

have completed intensive outpatient or inpatient treatment before study participation, (c) be fluent in English, (d) be medically cleared for outpatient care, (e) be able to attend the 8 weekly treatment sessions, (f) agree to random assignment to condition, and (g) be willing to complete follow-up questionnaires. Exclusion criteria were: (a) active psychotic disorder or dementia, (b) acute suicidality or imminent danger to others, or (c) inclusion in the pilot MBRP study. Overall, 84 interested participants did not meet study criteria, and 32 declined participation after completing the prescreen interview.

Clients in this community treatment center typically progressed through three phases of treatment: (a) intensive (inpatient or outpatient), (b) weekly aftercare, and (c) monthly aftercare over the period of 1 year on average. Participants in the MBRP study were recruited from the second phase of treatment, when clients were engaged in weekly aftercare. After giving written informed consent, participants were randomized. They were subsequently enrolled in one of the three treatment groups for 8 weeks, before returning to standard aftercare at the agency. Participants were evaluated at baseline, after the 8-week treatment group, and at 2-, 4-, 6- and 12-month follow-up interviews. Retention rates for these follow-ups were: 74.5%, 67.8%, 65.4%, 78.3%, 80.4% respectively, with follow-up rates generally increasing over time.

Participants

The majority of participants in the MBRP study were male (71.5%), white, (57.9%), employed (61.3%), had a high school diploma or equivalency (91.4%), and had an annual household income below $5,000 dollars. Participant ages ranged from 18-70 years. Participants also reported the use of multiple substances with the following substances listed as their primary

and/or secondary drug of choice: alcohol (71.6%), cocaine (34.0%), marijuana (34.5%), methamphetamine (18.9%), and opiates (22.7%).

Follow-Up Strategies

Given the transient nature of the study's participants, it was important to develop creative retention strategies over the course of the study. Some strategies were implemented during the entire study, and others were added as the study continued and new ideas were generated. The author of this thesis served as a research assistant for the main study, and was therefore directly involved in developing and applying these strategies. For the purpose of building on the previous literature exploring retention strategies, the following list of follow-up strategies is explained in greater detail in the sections to follow.

Database searches. Searching jail registers aided in tracking participants. Incarceration was an obstacle to many participants' participation in follow-up assessment; thus tracking participants during incarceration allowed study staff to make prompt calls upon a participant's release. Other public databases, such as newspaper obituaries and death records were also searched. Inner-agency detoxification records also facilitated the location of several participants.

Changes to follow-up protocol. In order to reduce assessment burden, two shorter surveys, each containing the main outcomes of interest, were offered to participants who were unable to attend the full 90 minute follow-up. In addition, the study obtained IRB approval to conduct interviews both online and over the phone, allowing participants to complete the interview from various locations. Other changes to follow-up protocol included alternate meeting places, alternate compensation types of the same amount, and providing childcare during

appointments. It is possible that these changes in protocol may be correlated with the rising retention rates over the course of the study.

Appointment strategies. Follow-up appointments offered an excellent opportunity to build rapport with participants. At each appointment the participant completed an updated locator form with the appropriate releases of information. Bonuses were offered to participants for treatment attendance, and after four consecutive follow-ups. Snacks were provided at each appointment. Advance appointment scheduling (and reminder cards) was also used.

Emails and texts. For some participants, email correspondence was preferable to phone communication. Additionally, though some participants did not respond to phone calls, several responded to texts, and scheduled all appointments in this fashion.

Use of personal collateral contact people. Personal collateral contacts provided by participants were used to track participants and schedule them for follow-ups. Types of personal collateral included: family members, close friends, service providers, counselors, probation officers, 12-step sponsors, employers, cultural centers, temporary homeless camps, and shelters. Maintaining regular contact with collateral contacts also expedited the relocation process.

Calling protocol. Generally, frequent calls between follow-up check-ins and reminder calls were used to increase contact with participants. Some participants also asked to be contacted via text. Calls were timed more heavily toward the beginning of the month to help reach clients whose phone minutes renewed on a monthly basis. Also, calling from non-agency numbers helped locate participants who were avoiding calls from the treatment agency.

Confidentiality

Given the sensitive nature of information in this study, confidentiality was paramount. The MBRP study acquired a Federal Certificate of Confidentiality that provided additional protection for study participants. This reduced the risk of researchers being forced to disclose sensitive information about study participants. Additionally, when attempting to contact difficult to reach participants, researchers identified themselves as researchers calling about the "University of Washington study." No other information was given to personal contacts.

Present Study

Defining Difficulty

As explored in the literature review, studies using this methodology measure difficulty in several ways. Some use number of contacts until the follow-up (e.g., Digiusto, Panjari, Gibson, Rea, 2006; Scott, 2004), others the number of contact attempts (e.g., LaPorte, McLellan, Erdlen & Parente, 1981; Nemes, Wish, Wraight, & Messina, 2000; Walton, Romanathan, & Reischl, 1998), and still others use the length of time until the follow-up is completed (e.g., Cottler, Compton, Ben-Abdallah, Horne, & Claverie, 1996; Hansten, Downey, Rosengren, & Donovan, 2000). Polich, Armor and Braiker (1980) used a composite measure which took into account multiple variables including contact attempts and amount of fieldwork required. Fleming et al. (unpublished) also took into account the cost per participant in follow-up-related expenditures (including staff time and incentives).

In the current study, contact difficulty was measured by the mean number of contact attempts per follow-up. Simply evaluating the total number of successful contacts, or the length

of time until follow-up completion, may not have captured the effort involved to achieve successful contact or follow-up completion. Thus, contact attempts were a better measure of the difficulty involved in reaching participants. Additionally, number of contact attempts is the most frequently used definition of difficulty in the literature (Digiusto, Panjari, Gibson, Rea, 2006; LaPorte, McLellan, Erdlen & Parente, 1981; Moos & Bliss, 1978; Nemes, Wish, Wraight, & Messina, 2000; Scott, 2004; Walton, Romanathan, & Reischl, 1998). For the purpose of this study, all of the following were considered contact attempts: phone calls, letters, in-person meetings, and emails.

This measure of difficulty brings up an important limitation of the methodology. Contact attempts were recorded in vivo using the online study database. However, contact attempts documented in this log are incomplete, and underestimate the actual number of attempts made (for example calling "episodes" were logged more frequently than each separate phone call, when the participant had multiple contact methods). Also, the log did not record other attempts made, such as those involving online searches or social service agencies. When faced with a similar situation in their own study, Hansten et al. (2000) concluded that the record, though incomplete, still suggested a link between effort and follow-up difficulty.

In addition to the measure of difficulty, it is also important to determine what qualifies (i.e. how many contact attempts) a participant as "difficult-to-contact" (DTC). Again, studies use different cut-offs to define difficulty based on their selected measure. Some divide the data into multiple categories of difficulty, sometimes up to eight groups (e.g., Polich, et al., 1980), others divide the entire sample into two dichotomous groups (e.g., Cottler, et al., 1996), and others look

at either end of their sample (e.g. 25th and 75th percentile; Walton, et al., 1998). Another group of researchers compare the easiest 60%, 70%, or 80% of their sample to the hardest 40%, 30%, or 20% (Fleming, et al., unpublished; Hanston, et al., 2000; Scott, 2004). In the case of these studies, these percentages were chosen so that researchers could compare a lower yet acceptable retention rate (such as 60% or 65%) with a much higher retention rate. These studies enable the researcher to judge how the study might have looked if a lower retention rate were achieved.

For the purposes of the current study, this final strategy was followed, using an 80/20% cut-off point as modeled by Fleming et al. (unpublished). Using this technique, participants who completed the final 12-month follow-up were divided into two groups: those who were easier-to-contact (ETC) and those who were more difficult (DTC; see *Figure 1*). Of the total study sample, 80.4% of participants completed the 12-month follow-up (230). The ETC group was defined as the first 80% of participants to complete the 12-month follow-up (the first 174 participants); the DTC group was defined as the last 20% of participants who required more contact attempts to complete the 12-month follow-up (last 56 participants). This 80/20% split enabled the researcher to compare the findings of the higher 80.4% retention rate of the MBRP study with the findings of the easier to locate 64% of the sample (toward the low end of acceptability). The extent to which these two samples are similar, determines the potential for attrition to bias the internal or external validity of study findings.

Participant Groups

Participants were divided into difficulty groups as follows: ETC participants, DTC participants, and non-completers. Figure 1 shows the study groups for the first and second phases

of the study, which involved a comparison of three groups. Of the 286 participants, 56 failed to complete the study ("non-completers"). Of the 230 participants who did complete the study, 8 participants were contacted beyond 2 months past the 3-week follow-up window. This issue systematically affected those enrolled in later cohorts due to the lessening of time for follow-ups at the completion of the study. For this reason, the 8 participants whose follow-up did not occur within 2 months of the follow-up window were excluded from analysis. The remaining 222 participants were classified as either ETC (the 80% requiring the least amount of contact attempts) or DTC (the 20% requiring the highest number of contact attempts). Figure 2 shows the study groups for the third phase of the study, involving only comparisons between ETC and DTC participants.

Figure 1. Phases 1 and 2: Study groups for demographic data comparison.

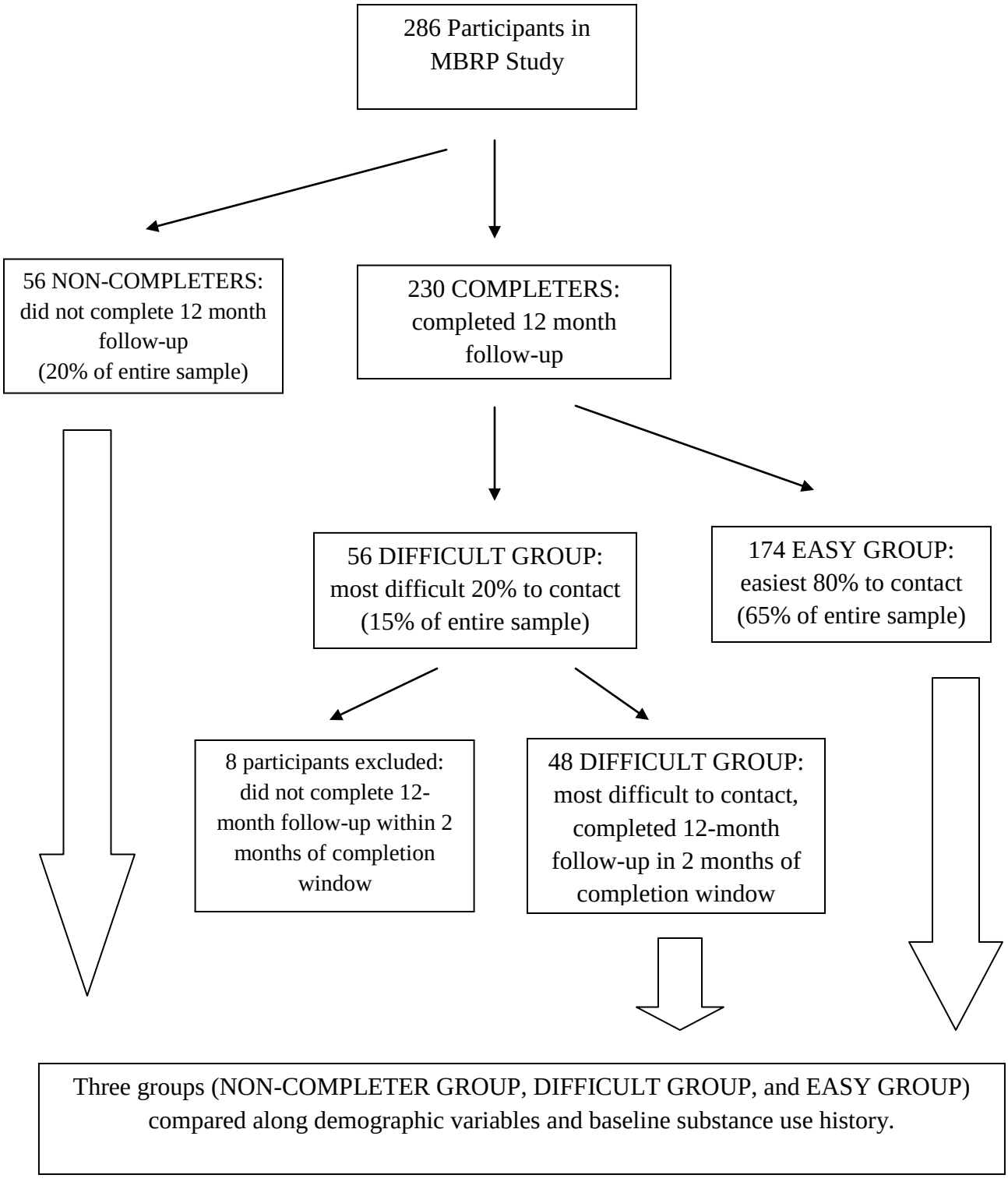
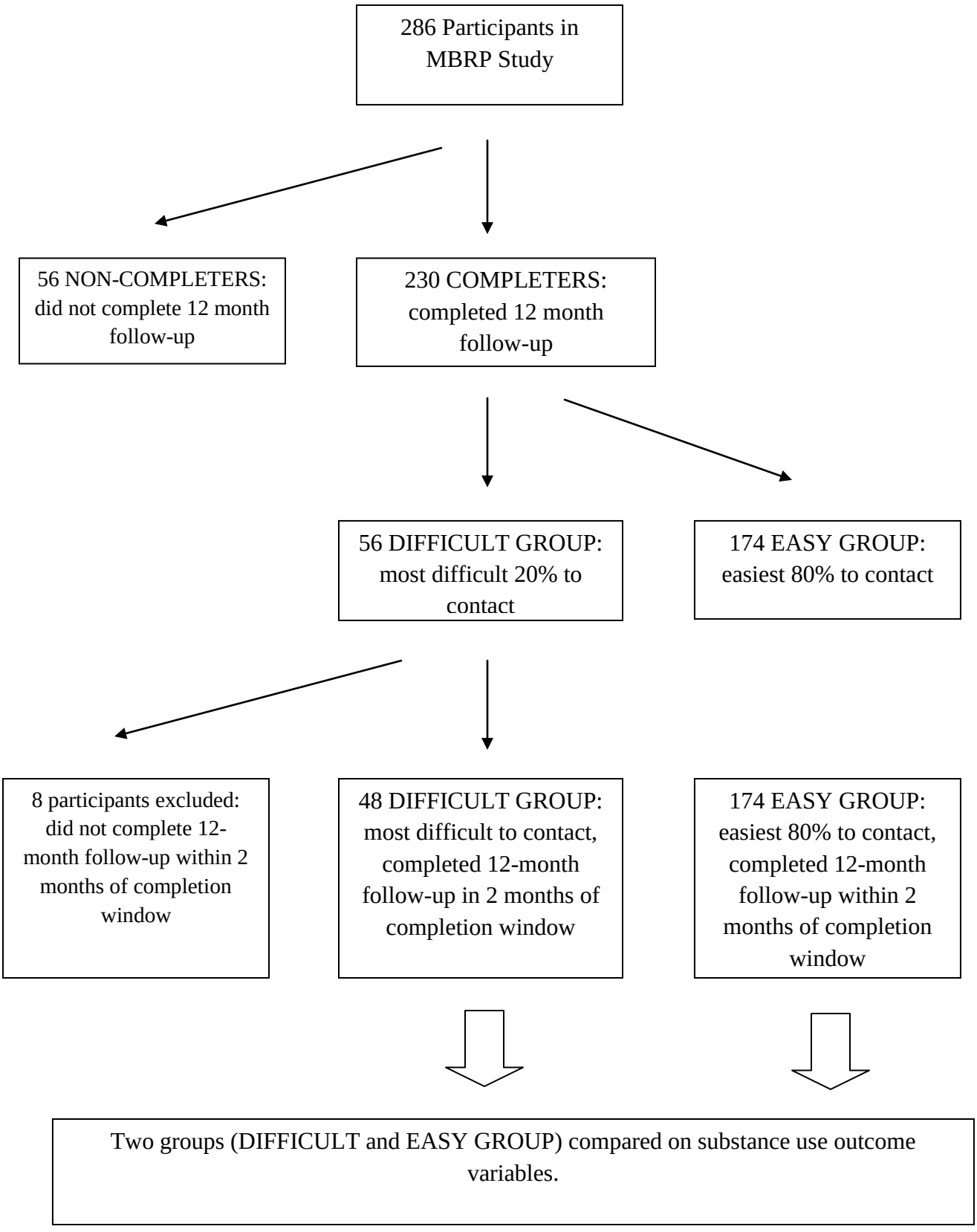


Figure 2. Phase 3: Study groups for outcome variable comparison.



Measures

Demographic variables. Phases 1 and 2 of the study required the collection of certain demographic variables including age, race (white, African American, Asian, Latino/a, Pacific Islander, Native American, Mixed, other), gender (male/female), employment status (employed, unemployed), marital status (single/partnered), and income. Annual income was measured using an ordinal variable of different income brackets.

Baseline characteristics. Phases 1 and 2 also used additional measures which included: total Beck Depression Inventory (BDI) score at baseline, total Beck Anxiety Inventory (BAI) score at baseline, and history of incarceration at baseline (Beck, Steer & Carbin, 1988; Beck, Epstein, Brown, & Steer, 1988). The BDI is a short 21-measure for depression symptoms. It has consistently shown high internal consistency across decades of administration in both clinical and non-clinical populations (Beck, Steer & Carbin, 1988). Additionally, it is able to differentiate between depression and anxiety. The BAI is a short 21-item measure that assesses for anxiety symptoms, and has shown high internal validity, and high test-retest reliability (Beck, Epstein, Brown, & Steer, 1988). History of incarceration was measured by a single binary self-report question, where the participant endorsed previous incarceration (jail or prison) or did not.

Baseline substance use measures. Baseline substance use measures included the following: primary and secondary drugs of choice, self-reported substance use in the past 30 days of treatment, and total Short Inventory of Problems (SIP-II) score (Miller, Tonigan, Longabaugh, 1995). The SIP-II is a 15-item measure that measures problems associated with substance use, with higher scores reflecting that a person is experiencing more problems as a result of his/her substance use. It has shown high reliability and internal consistency in large

populations of substance users (Forcehimes, Tonigan, Miller, Kenna, & Baer, 2007; Kiluk, Dreifuss, Weiss, Morgenstern, & Carroll, 2013). Participants were able to pick both a primary and secondary drug of choice. Many participants were polysubstance users, and this measure helped capture this important information. Substance use was collected using the Timeline Follow Back (TLFB), a measure in which the participant self-reports use of various substances for each day in a given time period (Sobell & Sobell, 1992). The TLFB has demonstrated good reliability and validity when delivered in person, by phone or online (Sobell, Brown, Leo, & Sobel,1996; Sobell & Sobell). Treatment hours were collected from agency treatment records, in order to total the number of client contact hours the participant had with the agency during the period of the 8-week study intervention.

Substance abuse outcomes. Incidence of drinking, heavy drinking, and drug use, in addition to frequency of use over the 12-month follow-up period was measured using the TLFB. Heavy drinking was defined as 4 or more drinks for women and 5 or more drinks per day for men (National Institute of Alcohol Abuse and Alcoholism, 2004).

Statistical Analysis

Phases 1 and 2: Baseline demographics, characteristics and substance use. A series of ANOVA and Chi-Square tests were run to assess relationships between participant demographics, baseline characteristics, baseline substance use severity, and tracking difficulty groups (non-completers, ETC, and DTC). Substance use variables were recoded from continuous to categorical variables due to the skew and kurtosis of the data, which was not significantly alleviated with data transformation (e.g., any use, average use once per week, average use twice

per week, average use 3 times per week). Variables that were found to be significantly correlated with difficulty were then entered into generalized linear models as predictors, to investigate if these variables were still correlated with the primary outcome of contact difficulty when controlling for other variables. All three groups were compared for Phase 1, and DTC participants were compared to non-completers in Phase 2.

Phase 3: Outcome measures. Incidence of drinking, heavy drinking, and drug use, in addition to frequency of use (weekly, twice weekly, thrice weekly) over the 12-month follow-up period were assessed as primary outcomes in binary generalized linear models. These models assessed associations between difficulty group and substance use outcomes, while controlling for demographic variables related to difficulty group.

Due to the exploratory nature of this study, significance value was set at $p < .05$. These results using this *p*-value should be interpreted with caution, due to the use of multiple statistical tests. Using the Bonferroni correction for multiple tests would suggest that findings found at the $p = .001$ level are less likely to be accounted for by beta errors.

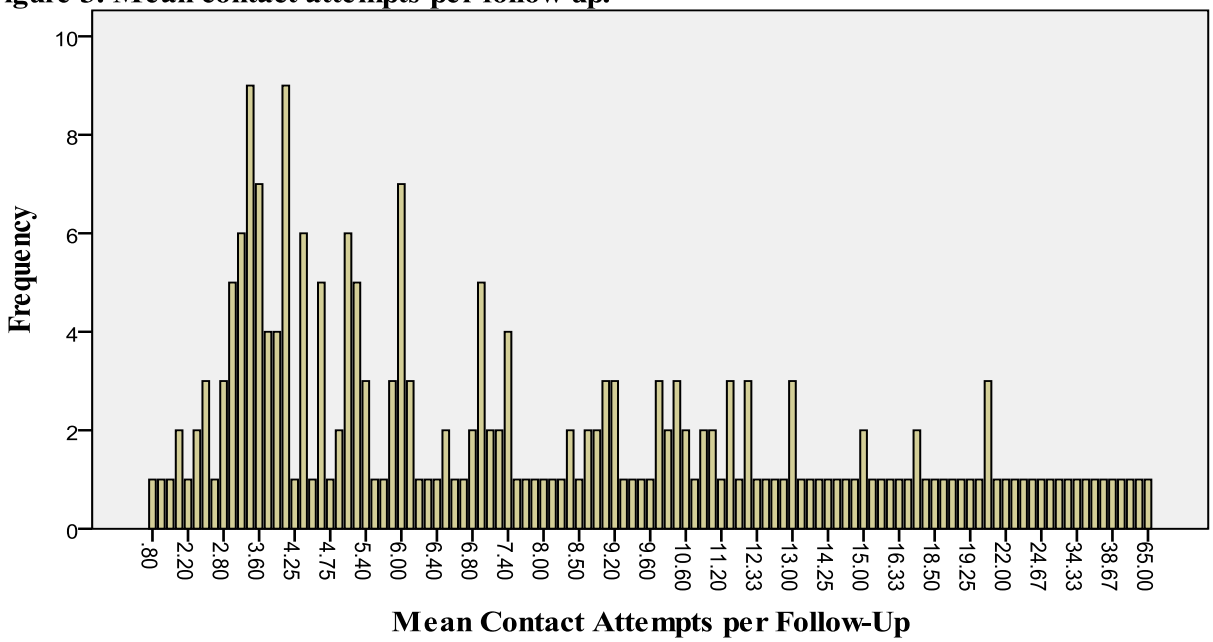
CHAPTER 4

RESULTS

Contact Difficulty

Contact difficulty was measured by calculating the mean number of contact attempts required for each follow-up interview. Participants could complete up to five follow-ups in the course of the study (post-course, 2-month, 4-month, 6-month, and 12-month). Mean contact attempts ranged from 0.8 to 65 attempts per follow-up. The data demonstrated a distinctly positive skew, indicating that a small proportion of study completers needed significantly more contact attempts before completing a follow-up. Figure 3 shows the frequency distribution of this variable. Fifty percent of study completers finished follow-up interviews after an average of 7 contact attempts per follow-up, 70% finished follow-ups after 11 contact attempts, and 90% finished after 19 contact attempts.

Figure 3. Mean contact attempts per follow-up.



The 80/20 split in the data occurred at 13 contact attempts. Thus, participants who required less than 13 contact attempts were defined as easy-to-contact (ETC); those who required more were defined as difficult-to-contact (DTC).

Phase 1: Comparing Baseline Demographics, Characteristics, and Substance Use, between ETC, DTC and Non-Completers

Phase 1 of the study compared the three different difficulty groups (ETC, DTC, and non-completers) to each other. Table 3 shows the comparisons between these groups and their baseline demographics, characteristics, and substance use measures. Variables found to be significant in the three-group model were then compared in two groups: study completers (both ETC and DTC) compared to non-completers, in order to examine how attrition may have impacted the characteristics of study completers. Table 4 shows the comparisons between these two groups.

ANOVA and Chi-Square Comparisons

Age. An analysis of variance showed that age was significantly related to difficulty group, F (2, 269) = 4.538, p = .012. Post hoc analyses using the Tukey HSD post hoc test for significance indicated that the average age of non-completers (M = 35.2, SD = 10.10) was significantly lower than ETC (M = 39.9, SD = 11.06).

Race. A chi-square test of independence was performed to examine the relationship between race and contact difficulty. The relationship between these variables was not significant, X^2 (14, N = 276) = 15.959, p = .316. Due to low numbers of certain racial groups, (Latino/a, Native American, Asian, and Pacific Islander), additional chi-square tests for independence were

performed to assess whether white race (compared to all others) or African-American race (compared to all others) was significantly related to difficulty. White race was not significantly related to difficulty group (p = .308). However, African-American race reached significance when comparing completers to non-completers, X^2 (1, N = 276) = 4.301, p = .038, OR = .434. Thus, the odds of study completion for a non-African American participant would be 43.4% that of an African American participant.

Gender. A chi-square test of independence was performed to examine the relationship between gender and contact difficulty. The relationship between these variables was significant, X^2 (2, N = 272) = 6.031, p = .049, such that males were more difficult to contact compared to females. A second, two group comparison looking at study completers and non-completers, females were more likely to complete the study as compared to males X^2 (1, N = 272) = 5.859, p = .038, OR = .318. The odds of a male completing the study were 34.0% that of a female completing the study.

Employment Status, Marital Status, Annual Income, and History of Incarceration. Chi-square tests of independence were performed to examine the relationships between contact difficulty and several other variables including: employment status, marital status, annual income, and history of incarceration. These relationships were all non-significant. See Table 3 for specific results. Chi-square tests of independence were performed to examine the relationships between study completion and employment status, marital status, annual income, and history of incarceration. These relationships were all non-significant. See Table 4 for specific results.

Total BDI Score, Total BAI Score, and SIP-II Score. An analysis of variance showed that total BDI score at baseline was not significantly related to contact difficulty, F (2, 272) = .505, p = .604 (see Table 3). An analysis of variance showed that BAI score at baseline was not significantly related contact difficulty, F (2, 272) = .065, p = .937. An analysis of variance showed that the SIP-II score at baseline, reporting levels of problems associated with substance use, was not significantly related to contact difficulty, F (2, 273) = 0.466, p = .628. Subsequent analyses of variance showed that neither BDI nor BAI nor SIP-II were significantly related to study completion (see Table 4).

Treatment Hours. An analysis of variance showed that the number of treatment hours (while in the 8-week study period) was significantly related to contact difficulty, F (2, 274) = 8.548, p < .001 (see Table 3). Post hoc analyses using the Tukey HSD post hoc test for significance indicated that ETC participants (M = 13.14, SD = 6.71) had significantly more treatment hours compared to DTC participants (M = 9.39, SD = 7.84) and non-completers (M = 9.74, SD = 6.76) by an average of 3.74 and 3.39 hours respectively. In a second analysis, treatment hours remained significant when comparing study completers to non-completers, F (1, 276) = 4.301, p < .038 (see Table 4).

Drug of Choice. A chi-square test of independence was performed to examine the relationship between drugs of choice (primary or secondary) and contact difficulty (see Table 3). The relationship between alcohol and contact difficulty was not significant X^2 (2, N = 276) = 839, p = .657, as was the case with marijuana, X^2 (2, N = 276) = 1.252, p = .535, and methamphetamines, X^2 (2, N = 276) = 4.815, p = .090. The relationship between cocaine and

contact difficulty was significant, X^2 (2, N = 276) = 5.878, p = .053, such that those who reported cocaine as a primary or secondary drug of choice were easier to contact compared to those who did not. On the other hand, those who reported opiates as their primary or secondary drug of choice were more difficult to contact, X^2 (2, N = 276) = 6.746, p = .034.

When comparing completers and non-completers, both cocaine and opiates remained significant (see Table 4). The relationship between cocaine and study completion was significant, X^2 (1, N = 276) = 3.763, p = .052, OR = .506, with cocaine users 50.6% more likely to complete the study compared to participants who did not use cocaine. The relationship between opiates and study completion was also significant, X^2 (1, N = 276) = 6.103, p = .052, OR = 2.268, with the odds of completing the study being 2.268 higher if the participant did not use opiates.

Use in the Past 30 Days. A chi-square test of independence was performed to examine the relationship between substance use in the past 30 days and contact difficulty. The relationship between these variables was significant, X^2 (2, N = 276) = 10.385, p = .006. When comparing study completers and non-completers, those who used 30 days prior to study participation were no more likely to complete the study than abstainers, X^2 (1, N = 276) = .165, p = .685.

Table 3. Three group comparison of baseline characteristics.

	Pearson Chi-Square Value	*F*	*p* value
Age		4.538	.012
Race	15.959		.316
African American	4.301		.038
Gender	6.031		.049
Employment Status	0.921		.631
Marital Status	0.598		.454
Annual Income	10.894		.779
History of Incarceration	3.164		.206
BDI total		0.505	.604
BAI total		0.065	.937
SIP-II Score		0.466	.628
Treatment Hours		8.548	.001
Alcohol Primary or Secondary	0.839		.657
Marijuana Primary or Secondary	1.252		.535
Cocaine Primary or Secondary	5.878		.053
Methamphetamine Primary or Secondary	4.815		.090
Opiates Primary or Secondary	6.746		.034
Use in 30 days prior to study	10.385		.006

Table 4. Completers vs. non-completers.

	Pearson Chi-Square Value	*F*	*p* value	*OR*
Age		6.213	0.013	
Race	12.940		.880	16.970
African American	4.301		0.015	.434
Gender	5.859		0.038	.318
Employment Status	.600		.439	.609
Marital Status	.568		.505	.543
Annual Income	3.698		.814	3.434
History of Incarceration	2.760		.108	2.984
BDI Total		.087	.769	
BAI Total		.051	.821	
SIP-II Score		.370	.544	
Treatment Hours		6.062	0.014	
Alcohol Primary or Secondary	.497		.515	.489
Marijuana Primary or Secondary	.179		.751	.178
Cocaine Primary or Secondary	3.763		0.052	.506
Methamphetamine Primary or Secondary	2.900		.124	2.713
Opiates Primary or Secondary	6.103		0.013	2.268
Use in 30 days prior to study	0.165		0.685	

Generalized Linear Models

After completing chi-square and ANOVA comparisons for the 18 variables listed in Table 3 and Table 4, several variables emerged as associated with contact difficulty: gender, age, African-American race, opiate use, cocaine use, treatment hours, and use during treatment. To further explore these associations, generalized linear models (GLMs) were fit to the data to determine how each predictor performed when controlling for all other predictors. First, more complex GLMs were run using all 18 variables of interest, to confirm that no new variables

emerged as related to contact difficulty once other variables were controlled. These models produced no new significant findings. Following this, simpler GLM models were fit to the data using only the variables of interest (those variables which were associated with contact difficulty in chi-square and ANOVA comparisons) to allow for more degrees of freedom in analysis and provide the best fitting model with the simplest number of predictors. Results of the generalized linear model involving all three contact difficulty groups (DTC, ETC, non-completers) are found in Table 5. Results of the generalized linear model involving completers and non-completers are found in Table 6.

After controlling for the factors mentioned above, age, African-American race, cocaine use, and use during treatment failed to have a significant effect on difficulty group or study completion. Only gender emerged as a consistent predictor of difficulty group and study completion. Male gender was associated with higher contact difficulty (B = .219, p = .036; see Table 5). A second generalized linear model (controlling for the same factors) comparing completers to non-completers showed that women were more likely to complete the study than men (B = -1.139, p = .020; see Table 6). Treatment hours remained associated with contact difficulty (B = -.019, p = .003; see Table 5), but not study completion (p = .078; see Table 6) after controlling for other factors. Opiate use remained associated with study completion (p = .009; see Table 6), but not difficulty group (p = .071; see Table 5).

In sum, after controlling for other variables, the only factors that were significantly associated with difficulty group were treatment hours and male gender, such that fewer treatment hours and male gender were associated with greater contact difficulty. Similarly, after controlling

for other variables, the only factors that were significantly associated with higher study completion rates were less opiate use and female gender.

Table 5. Generalized linear model for three study groups: ETC, DTC, and non-completers.

Parameter Estimates

Parameter	*B*	Std. Error	95% Wald Confidence Interval		Hypothesis Test		
			Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	.974	.2677	.450	1.499	13.252	1	.000
Non African American	.006	.1205	-.231	.242	.002	1	.963
African American	0[a]	.	.	.	.	.	.
Opiates not DOC	-.215	.1191	-.449	.018	3.268	1	.071
Opiates DOC	0[a]	.	.	.	.	.	.
Cocaine not DOC	.087	.1141	-.136	.311	.584	1	.445
Cocaine DOC	0[a]	.	.	.	.	.	.
Male	.219	.1047	.014	.424	4.388	1	.036
Female	0[a]	.	.	.	.	.	.
No use in past 30 days	-.098	.1366	-.365	.170	.512	1	.474
Use in past 30 days	0[a]	.	.	.	.	.	.
Age in years	-.005	.0045	-.014	.004	1.307	1	.253
Treatment Hours	-.019	.0065	-.032	-.007	8.898	1	.003
(Scale)	.537[b]	.0475	.452	.638			

Dependent Variable: Difficulty group

Predictor Variables: Race, opiates as drug of choice, cocaine as drug of choice, gender, use during treatment, age, and treatment hours.

a. Set to zero because this parameter is redundant.

b. Maximum likelihood estimate.

Table 6. Generalized linear model for two study groups: completers vs. non-completers.

Parameter Estimates

Parameter	*B*	Std. Error	95% Wald Confidence Interval		Hypothesis Test		
			Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	1.401	1.0746	-.705	3.507	1.700	1	.192
Non African American	-.254	.4797	-1.194	.686	.280	1	.597
African American	0[a]	.	.	.	.	.	.
Opiates not DOC	1.076	.4108	.270	1.881	6.855	1	.009
Opiates DOC	0[a]	.	.	.	.	.	.
Cocaine not DOC	-.092	.4489	-.971	.788	.042	1	.838
Cocaine DOC	0[a]	.	.	.	.	.	.
Male	-1.139	.4901	-2.099	-.178	5.397	1	.020
Female	0[a]	.	.	.	.	.	.
No use 30 days prior to study	-.550	.5623	-1.652	.552	.957	1	.328
Use 30 days prior to study	0[a]	.	.	.	.	.	.
Age in years	.013	.0177	-.022	.047	.503	1	.478
Treatment hours	.050	.0283	-.006	.105	3.105	1	.078
(Scale)	1[b]						

Dependent Variable: non-completers vs. completers

Predictor Variables: Race, opiates as drug of choice, cocaine as drug of choice, gender, use during treatment, age, and treatment hours.

a. Set to zero because this parameter is redundant.

b. Fixed at the displayed value.

Phase 2: Comparison between Non-Completers and Difficult-to-Contact Participants

Phase 2 of the study compared the DTC to participants to non-completers to determine the similarities and differences between the two groups. The extent to which these two groups are similar lends support to the method of using DTC participants as proxy measures for those

who are not reached for final follow-ups. Table 7 shows the comparison of these groups and their baseline demographics, characteristics, and substance use measures.

Table 7. DTC vs. Non-completers.

	Pearson Chi-Square Value	*F*	*p* value	*OR*
Age		0.605	0.439	
African American	0.135		0.135	
Gender	0.026		0.026	.340
Treatment Hours		0.06	0.807	
Cocaine Primary or Secondary	0.305		0.581	
Opiates Primary or Secondary	5.120		0.024	2.268
Use in 30 days prior to study	3.803		.051	.335

When comparing study non-completers to DTC, females were more likely than males to be DTC, X^2 (1, N = 100) = 5.120, p = .026, OR = .340. Opiate users were more likely than non-users to be non-completers than DTC, X^2 (1, N = 276) = 6.103, p = .024, OR = 2.268, with the odds of completing the study as a DTC participant being 2.268 higher if the participant did not use opiates as a primary or secondary drug of choice. Additionally, participants who did not use in the past 30 days were more likely than people who used to be non-completers X^2 (1, N = 276) = 3.803, p = .051, OR = .335. In sum, DTC participants did not differ from non-completers on many of the baseline characteristics. However, non-completers were more likely than DTC participants to be male, opiate users, and less likely to have used substances in the 30 days prior to the study. When entered into a generalized linear model, controlling for other variables, these differences between groups remained significant. Table 8 shows the *p*-values for this model.

Table 8. Generalized linear model for two groups: DTC and non-completers.

Parameter Estimates

Parameter	*B*	Std. Error	95% Wald Confidence Interval		Hypothesis Test		
			Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	1.310	1.4572	-1.546	4.166	.808	1	.369
Non African American	-.921	.7012	-2.295	.453	1.725	1	.189
African American	0[a]	.	.	.	.	.	.
Opiates not DOC	1.641	.6343	.398	2.885	6.694	1	.010
Opiates DOC	0[a]	.	.	.	.	.	.
Cocaine not DOC	.551	.6887	-.799	1.901	.641	1	.423
Cocaine DOC	0[a]	.	.	.	.	.	.
Male	-1.398	.6414	-2.656	-.141	4.753	1	.029
Female	0[a]	.	.	.	.	.	.
No use 30 days prior to study	-1.391	.6691	-2.703	-.080	4.324	1	.038
Use 30 days prior to study	0[a]	.	.	.	.	.	.
Age in years	.006	.0253	-.043	.056	.066	1	.798
Treatment hours	-.025	.0344	-.092	.042	.534	1	.465
(Scale)	1[b]						

Dependent Variable: DTC versus non-completers

Predictor Variables: Race, opiates as drug of choice, cocaine as drug of choice, gender, use during treatment, age, and treatment hours.

Phase 3: Comparison between Difficult-to-Contact Participants and Easy-to-Contact Participants on Outcome Variables

Phase 3 compared the substance use outcomes of ETC participants to the DTC participants. Results from the generalized linear models (controlling for age, gender, race, drug

of choice, pre-study use, and treatment hours) are provided in Table 9. Contact difficulty was significantly related to having at least one incidence of drinking (p = .005), at least one incidence of heavy drinking (p = .013), and at least one use of any substance (p = .001).

When examining frequency of use (use monthly, weekly, twice weekly), only frequency of alcohol use was related to difficulty group. On the whole, participants who drank at least once a week (p = .003), or drank heavily at least once a week (p = .008), were more likely to be DTC (p = .008). Thus, participants who drank weekly, were more difficult to contact. Frequency of drug use was unrelated to difficulty group. However, it should also be noted that use twice a week of any substance was not correlated with contact difficulty. Thus, it appears that frequency of drinking is related to contact difficulty to a point, but more habitual use of alcohol is not correlated with contact difficulty.

Table 9. Substance use outcomes in binary logistic model

	B	*p* value
Any drinking in follow-up	1.105	0.005
Drinking once a month	9.582	0.002
Drinking once a week	8.65	0.003
Drinking twice per week	122.616	0.977
Any heavy drinking in follow-up	0.961	0.013
Heavy drinking once a month	1.51	0.008
Heavy drinking once a week	7.542	0.006
Heavy drinking twice per week	49.774	0.99
Any drug use in follow-up	0.686	0.069
Drug use once a month	1.042	0.066
Drug use once a week	1.354	0.245
Drug use twice per week	1.943	0.128
Any drinking or drug use in follow up	11.328	.001

CHAPTER 5

DISCUSSION

Studies utilizing difficult-to-contact (DTC) participants as proxy measures for those lost to attrition rarely compare these groups on baseline characteristics to determine the extent to which they differ. The present study analyzed baseline differences between non-completers, DTC, and easy-to-contact (ETC) participants, in addition to differences between completers and non-completers. Regarding the extent to which DTC and non-completers differed, this study found few differences in baseline characteristics. DTC participants differed from non-completers on three measures, after controlling for other factors, such as age, race, and number of treatment hours. Non-completers were more likely than DTC participants to be male, use opiates as a primary or secondary drug of choice, and to have used substances in the 30 days prior to the study. However, it should be noted that none of these differences would remain if using a Bonferroni correction for multiple tests (in this case, $p < .001$). Thus, these results should be interpreted in an exploratory manner. Essentially, this study suggests that DTC participants may be more likely to be female, abstinent at the start of the study, and have a drug of choice other than opiates. Therefore, studies using DTC participants as proxies for non-completers may not capture the full variability of those lost to attrition, who may be more likely to be male and use opiates. However, the data do clearly suggest that DTC participants and non-completers are more alike on baseline characteristics than different. Additional research is warranted to support this methodology approach.

Analysis comparing completers to those lost to attrition revealed similar results. Those lost to attrition only differed from completers at baseline on gender and opiate-use status. Non-completers were more likely to be male and use opiates as a primary or secondary drug of choice. However, these differences between groups also disappeared using the Bonferroni correction. Taken together, these results suggest that contact difficulty and attrition does not appear to meaningfully limit the generalizability of this study's findings. These results suggest that attrition bias may not be a major threat to validity in etiological studies with at least 65% retention. However, additional testing for attrition bias in efficacy studies is important, as differential attrition by condition (particularly among substance users) could threaten conclusion validity.

Baseline Characteristics

Prior research analyzing the degree to which participants differ based on contact difficulty has left conflicting results. Some researchers reported demographic differences (i.e. race, age, employment status, etc.), while others either found no differences between difficulty groups, or their differences were so small that the researchers deemed them unimportant. The results of this study may help clarify some of this literature. In the research addressing marital status, educational attainment, and income, only one group of researchers (e.g. Moos & Bliss 1978) have reported finding meaningful differences between groups based on these factors. The current study also failed to find support for the hypothesis that these variables are correlated with contact difficulty. Thus, these results bolster the literature supporting the assertion that DTC participants do not differ meaningfully from ETC participants in terms of education, income, and

marital status. Similarly, researchers have found that ETC and DTC participants do not differ based on psychiatric comorbidities; this study again supported this conclusion.

The findings of this study diverge from the literature in three areas: measures of criminality, gender, and treatment hours. Three studies previously found criminal history to be connected with contact difficulty (Bale, Arnoldussen, & Quittner, 1984; Nemes, Wish, Wraight, & Messina, 2000; Scott, 2004). However, this study did not. These differences in findings could be due to the gross distinction made in this study's measure of criminality (i.e., dichotomous answer: "Have you ever been in jail or prison?"), and the high number of participants who had been in jail or prison (73.7% of the sample). Variables which more fully capture degrees of criminality (e.g., length of incarceration, type of offenses, etc.) may be useful measures in clarifying this issue. Also, different studies had different protocols when participants became incarcerated. Some received permission to continue collecting personal interviews from incarcerated participants, while others did not, thus affecting data completeness.

When evaluating gender, previous studies did not report finding gender differences between ETC and DTC participants. In a report focused on attrition bias, Watson and Wooden (2009) reported that non-completers may be more likely to be male in longitudinal studies. However, this assertion was not specifically referring to studies of substance use. Given the more significant amount of literature that does not report finding an effect of gender on contact difficulty, and the possibility of alpha inflation, one should interpret the results of this study with caution. More specific research on female populations may help clarify this issue, as several studies excluded or recruited fewer women.

Few studies have examined the impact of treatment hours on contact difficulty. Digiusto, Panjari, Gibson and Rea (2006), found that lower treatment hours were correlated with greater contact difficulty. The present study found treatment hours to be unrelated to contact difficulty after controlling for age, gender, race, and drug of choice, suggesting that the association between treatment hours and contact difficulty is accounted for by these variables. It is also possible that another factor, which contributes to participants' engagement in treatment, may also affect their engagement in research participation. Based on these two studies, we can conclude that treatment hours may not help with retention, but they do not appear to hurt. Future studies addressing additional components of treatment engagement may help to clarify this issue.

Regarding age, race, and employment status, most studies report that these factors are not significantly related to contact difficulty. The present study supports these assertions, not having found significant evidence to support the influence of these factors on contact difficulty. However, it is worth mentioning that while previous studies link non-white races with being more difficult to contact (Scott, 2004; Watson & Wooden, 2009), the present study found that African American participants may be more likely to complete. However, as seen with other variables, these effects did not remain when controlling for other demographic factors.

The literature has also yielded mixed results in regard to drug of choice. The current study found no impact of alcohol, marijuana, or methamphetamines, on contact difficulty. However, when participants listed cocaine as a primary or secondary drug of choice, some differences were observed. Cocaine users were more likely to complete the study and more likely to be ETC. However, these effects did not remain consistent when controlling for race. It is

possible that the effects seen in correlation with cocaine use could be confounded by other demographic factors. Opiates, on the other hand, when identified as a primary or secondary drug of choice, were significantly associated with being DTC, and this effect remained after controlling for demographic variables. However, this effect was not significant using the Bonferroni correction. There was also a strong association between race and drug of choice in this sample: African Americans were significantly more likely to report using cocaine as a drug of choice, and significantly less likely to report using heroin or methamphetamine ($p < .05$). Similarly, those identifying as white were significantly less likely to report using cocaine as a drug of choice, and significantly more likely to report using heroin or methamphetamine ($p < .05$). Further studies could help to clarify the relationships between race, drug of choice, and contact difficulty.

Overall, the low incidence and significance in differences between ETC and DTC populations lend support to the assertion that participants do not differ systematically based on baseline characteristics, thus supporting the external validity of this study. While some factors remained significant after controlling for demographic variables (opiates, gender, use in the past 30 days, and treatment hours), no baseline characteristics achieved significance rates using the Bonferroni correction. Thus, while these results should be interpreted with caution, it appears that baseline characteristics may be largely unrelated to participant contact difficulty.

Outcome Variables

The main divide in the literature surrounding contact difficulty revolves around how contact difficulty may impact substance use outcomes. Several authors, running rigorous studies

have found that contact difficulty is related to worse substance use outcomes; other rigorous studies have revealed no significant effect of contact difficulty on outcomes. Substance use, as defined dichotomously (i.e. any drug use or drinking during the 12-month follow-up) was the only factor of the current study found to significantly differentiate DTC from ETC participants, using the Bonferroni correction. This suggests that while DTC participants may be less likely to be completely abstinent, DTC differences in substance use frequency (and perhaps, consequences associated with substance use frequency) may not be meaningfully different from ETC participants. Similarly, any incident of drinking, and any incident of heavy drinking was also found to be significant in the generalized linear model, although these did not meet the needed significance levels for the more stringent Bonferroni correction. Additionally, drinking or heavy drinking at least once a week were also associated with contact difficulty, though not meeting significance at the $p = .001$ level. In and of themselves, these results seem to suggest that worse substance use outcomes are in fact correlated with increased contact difficulty.

This outcome data appears to show a distinct trend toward contact difficulty being related to increased substance use over the course of the study. However, we must interpret the frequency data with caution. Only the dichotomous variable of "any use/no use" achieved a high enough level of significance according to the Bonferroni correction. Beyond this, the data trend toward suggesting that increased frequency in use is correlated with contact difficulty. The outcome data in this study (as is common in substance use studies) was significantly positively skewed, with a high preponderance of zero values, thus hindering more precise measurements for use frequency. Furthermore, only alcohol use outcomes were related to contact difficulty, not drug use outcomes. Central to interpreting the results of this study, is the question of how much

drinking may be considered problematic in a study such as this. Participants who reported drinking once a week appear to be more difficult to contact; this likelihood disappears when the drinking rate is increased to two times per week. The meaningfulness of this data is unclear. However it is also worth noting that participants who reported heavy drinking at least once a week, were also more likely to be DTC, lending support to the theory that worse substance use outcomes are associated with increased follow-up difficulty. However, due to the lack of a high level of significance in these trends, it is recommended that further investigation continue before consistently dismissing studies with less than a 65% retention rate. That being said, it is prudent to explore the impact of attrition bias on intervention studies, to ensure that differential attrition by condition does not impact efficacy findings.

It is also interesting to note how steeply effort to reach participants increases for the most difficult to contact participants. While nearly half of the sample is captured after 7 contact attempts, the last 10% required more than 20 attempts, and the last 5% required 30-65 attempts. Given the amount of resources needed to ensure very high retention rates, it is important to take into account the time and financial resources of the study when determining a target retention rate. While higher retention rates are universally preferred, it is prudent to consider the cost (time and resources) and benefit (of some increased accuracy) of pursuing these very high retention rates.

Limitations and Future Recommendations

When interpreting this study's data, it is important to bear in mind some important limitations. First, the outcome data was significantly positively skewed, making it more difficult

to discern use differences between the groups. While the frequencies and amounts of alcohol were recorded, only the frequency of other substances was noted; furthermore, all non-alcohol substance use was coded the same. More sophisticated measures, such as a measure quantifying type and amount of substance use (in addition to the number of alcoholic drinks), could convey additional information about the severity of use, which could aid in the interpretation of the results. Second, aside from African American and white participants there were not enough participants of other racial groups to do a more thorough analysis of the effect of race or ethnicity on contact difficulty. Research specifically recruiting other populations in high numbers could strengthen the literature. Third, the number of contact attempts is most certainly an underestimate, due to data collection procedures. It would be beneficial if future studies developed standard protocols around documenting retention efforts, from the beginning of the study. In order for this methodology to be further validated, better data tracking procedures are needed. Fourth, it should be noted that the retention rate for the present study was 80.4% after one year of following participants. Other studies in this field have managed to get higher retention (e.g. Cottler, Compton, Ben-Abdallah, Horne, & Claverie, 1996; Hansten, Downey, Rosengren, and Donovan, 2000; Scott, 2004). Achieving these high rates would allow researchers a more thorough view of the data. Finally, it must also be recognized that the study staff in charge of retention became more efficient and learned more techniques as the study progressed, as evidenced in the rising retention rates over time. It is possible that participants recruited at the beginning of the study may have been pursued with less intentionality compared to participants near the end of the study. Thus, there is a possibility of a cohort effect in this data.

Significance

There are several strengths of this study which make it well-suited to improve upon the literature. First, only four of the 10 studies examined in the literature review included follow-up lengths of one year or longer. Also, these studies each found different results. Given the increasing difficulty in tracking participants with longer follow-up windows, it is particularly important to explore potential attrition bias in studies with longer follow-up periods. Second, this study recruited participants from a community treatment center that predominantly serves low-income clients of all genders, thus allowing its results to be generalizable to other such populations. Additionally, participants reported a range of drugs of choice, including marijuana and methamphetamine which have received less attention in the literature. Also adding to the richness of this data is collection of up to two drugs of choice for participants who reported using multiple substances. Finally, it is also important to note that the follow-up strategies employed in this study did not require excess financial investment (except the time of the researcher), thus making these strategies accessible to studies with budgetary concerns.

Conclusion

The present secondary data analysis offers evidence that DTC participants can indeed be used as proxy measures for participants lost to attrition, though they may differ slightly in terms of gender make-up and use history. There was a trend in the data suggesting that DTC participants may be more likely to be male and use opiates compared to ETC participants. However this data is more exploratory in nature and is not significant in a Bonferroni model. Similarly, DTC participants were more likely to have at least one incidence of drug or alcohol

use; however, a clear association between frequency and severity of use with contact difficulty was not established. These results suggest that retention rates as low as 65% may still lead to meaningful data, and excessive effort to pursue the more difficult to contact participants may not be worth the time and energy, due to the potentially low impact of attrition on study results. However, these results should be interpreted with caution, as there were trends towards increased substance use being related to contact difficulty, and this could introduce attrition bias to samples with lower retention rates.

REFERENCES

Bale, R. N., Arnoldussen, B. H., & Quittner, A. M. (1984). Follow-up difficulty with substance abusers: Predictors of time to locate and relationship to outcome. *Substance Use & Misuse*, *19*(8), 885-902.

Bassuk, E. L., Buckner, J. C., Perloff, J. N., & Bassuk, S. S. (1998). Prevalence of mental health and substance use disorders among homeless and low-income housed mothers. *American Journal of Psychiatry*, *155*, 1561-1564.

Battjes, R. J., Leukefeld, C. G., Pickens, R. W., & Haverkos, H. W. (1988). The acquired immunodeficiency syndrome and intravenous drug abuse. *Bulletin on Narcotics*, *40*, 21-34.

Beck, A. T., Epstein, N., Brown, G., & Steer, R. A. (1988). An inventory for measuring clinical anxiety: Psychometric properties. *Journal of Consulting and Clinical Psychology*, *56*(6), 893.

Beck, A. T., Steer, R. A., & Carbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review*, *8*, 77-100.

Bootsmiller, B. J., Ribisl, K. M., Mowbray, C. T., Davidson, W. S., Walton, M. A., & Herman, S. E. (1998). Methods of ensuring high follow-up rates: Lessons from a longitudinal study of dual diagnosed participants. *Substance Use & Misuse*, *33*(13), 2665-2685.

Brown-Peterside, P., Rivera, E., Lucy, D., Slaughter, I., Ren, L., Chiasson, M. A., & Koblin, B. A. (2001). Retaining hard-to-reach women in HIV prevention and vaccine trials: Project ACHIEVE. *Journal Information*, *91*(9), 1377-1379.

Bowen, S., Witkiewitz, K., Clifasefi, S. L., Grow, J., Chawla, N., Hsu, S. H…. & Larimer, M. E. (2014). Relative efficacy of Mindfulness-Based Relapse Prevention, standard relapse prevention and treatment as usual for substance use disorders. *Journal of the American Medical Association--Psychiatry, 71*(5), *547-556.* doi:10.1001/jamapsychiatry.2013.4546

Capaldi, D., & Patterson, G. R. (1987). An approach to the problem of recruitment and retention rates for longitudinal research. *Behavioral Assessment, 9*(2), 169-177.

Coen, A. S., Patrick, D. C., & Shern, D. L. (1996). Minimizing attrition in longitudinal studies of special populations: An integrated management approach. *Evaluation and Program Planning*, *19*(4), 309-319.

Cotter, R. B., Burke, J. D., Loeber, R., & Navratil, J. L. (2002). Innovative retention methods in longitudinal research: A case study of the developmental trends study. *Journal of Child and Family Studies*, *11*(4), 485-498.

Cotter, R. B., Burke, J. D., Stouthamer-Loeber, M., & Loeber, R. (2005). Contacting participants for follow-up: How much effort is required to retain participants in longitudinal studies? *Evaluation and Program Planning*, *28*(1), 15-21.

Cottler, L. B., Compton, W. M., Ben-Abdallah, A., Horne, M., & Claverie, D. (1996). Achieving a 96.6 percent follow-up rate in a longitudinal study of drug abusers. *Drug and Alcohol Dependence*, *41*(3), 209-217.

Desmond, D. P., Maddux, J. F., Johnson, T. H., & Confer, B. A. (1995). Obtaining follow-up interviews for treatment evaluation. *Journal of Substance Abuse Treatment*, *12*(2), 95-102.

Digiusto, E., Panjari, M., Gibson, A., & Rea, F. (2006). Follow-up difficulty: Correlates and relationship with outcome in heroin dependence treatment in the NEPOD study. *Addictive Behaviors*, *31*(7), 1201-1210.

Festinger, D. S., & Dugosh, K. L. (2012). Paying substance abusers in research studies: Where does the money go? *The American Journal of Drug and Alcohol Abuse*, *38*, 43-48.

Fleming, C. B., Marchesini, G., Haggerty, K. P., Hill, K. G., & Catalano, R. F. (unpublished). The importance of high completion rates: Evidence from two longitudinal studies of the etiology and prevention of problem behaviors.

Flick, S. N. (1988). Managing attrition in clinical research. *Clinical Psychology Review*, *8*(5), 499-515.

Forcehimes, A. A., Tonigan, J. S., Miller, W. R., Kenna, G. A., & Baer, J. S. (2007). Psychometrics of the drinker inventory of consequences (DrInC). *Addictive Behaviors*, *32*(8), 1699-1704.

Goshin, L. S., & Byrne, M. W. (2012). Predictors of post-release research retention and subsequent reenrollment for women recruited while incarcerated. *Research in Nursing and Health*, *35*, 94-104.

Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology*, *60*, 549-576.

Groves, R. M., & Couper, M. P. (2012). *Nonresponse in household interview surveys*. Hoboken, NJ: John Wiley & Sons, Inc..

Haggerty, K. P., Fleming, C. B., Catalano, R. F., Petrie, R. S., Rubin, R. J., & Grassley, M. H. (2008). Ten years later: Locating and interviewing children of drug abusers. *Evaluation and Program Planning*, *31*, 1-9.

Hansen, W. B., Tobler, N. S., & Graham, J. W. (1990). Attrition in substance abuse prevention research a meta-analysis of 85 longitudinally followed cohorts. *Evaluation Review*, *14*(6), 677-685.

Hansten, M. L., Downey, L., Rosengren, D. B., & Donovan, D. M. (2000). Relationship between follow-up rates and treatment outcomes in substance abuse research: More is better but when is" enough" enough? *Addiction*, *95*(9), 1403-1416.

Jarlais, D. C. D., Perlis, T. E., & Settembrino, J. M. (2005). The use of electronic debit cards in longitudinal data collection with geographically mobile drug users. *Drug and Alcohol Dependence*, *77*, 1-5.

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., & Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Archive of General Psychiatry, 62*(6): 593-602. doi:10.1001/archpsyc.62.6.593.

Kiluk, B. D., Dreifuss, J. A., Weiss, R. D., Morgenstern, J., & Carroll, K. M. (2013). The Short Inventory of Problems – Revised (SIP-R): Psychometric properties within a large, diverse sample of substance use disorder treatment seekers. *Psychology of Addictive Behaviors*, *27*, 307-314. doi:10.1037/a0028445.

Lankenau, S. E., Sanders, B., Hathazi, D., & Jackson Bloom, J. (2010). Recruiting and retaining mobile young injection drug users in a longitudinal study. *Substance Use & Misuse*, *45*(5), 684-699.

LaPorte, D. J., McLellan, A. T., Erdlen, F. R., & Parente, R. J. (1981). Treatment outcome as a function of follow-up difficulty in substance abusers. *Journal of Consulting and Clinical Psychology*, *49*(1), 112.

Lyons, K. S., Carter, J. H., Carter, E. H., Rush, K. N., Stewart, B. J., & Archbold, P. G. (2004). Locating and retaining research participants for follow-up studies. *Research in Nursing & Health*, *27*, 63-68.

Merikangas, K. R., Mehta, R. L., Molnar, B. E., Walters, E. E., Swendsen, J. D., Aguilar-Gaziola, S., ... & Kessler, R. C. (1998). Comorbidity of substance use disorders with mood and anxiety disorders: Results of the International Consortium in Psychiatric Epidemiology. *Addictive Behaviors*, *23*(6), 893-907.

Miller, W.R., Tonigan, J.S., Longabaugh, R. (1995). In: The Drinker Inventory of Consequences (DrInC), Project MATCH Monograph Series. Mattson Margaret E., editor. Vol. 4. Rockville, MD: National Institute on Alcohol Abuse and Alcoholism; 1995.

Moos, R., & Bliss, F. (1978). Difficulty of follow-up and outcome of alcoholism treatment. *Journal of Studies on Alcohol and Drugs*, *39*(3), 473.

National Institute of Alcohol Abuse and Alcoholism. (2004). NIAAA Council approves definition of binge drinking. *NIAAA Newsletter*, *3*, 3.

Nemes, S., Wish, E., Wraight, B., & Messina, N. (2002). Correlates of treatment follow-up difficulty. *Substance Use & Misuse*, *37*, 19-45.

Nordström, G., & Berglund, M. (1986). Successfully adjusted alcoholics lost to long-term follow-up. A prospective study. *Drug and Alcohol Dependence*, *18*, 11-22.

Polich, J. M., Armor, D. J., & Braiker, H. B. (1980). *The course of alcoholism: Four years after treatment.* Santa Monica, CA: Rand Corporation.

Ribisl, K. M., Walton, M. A., Mowbray, C. T., Luke, D. A., Davidson, W. S., & Bootsmiller, B. J. (1996). Minimizing participant attrition in panel studies through the use of effective retention and tracking strategies: Review and recommendations. *Evaluation and Program Planning*, *19*, 1-25.

Robinson, K. A., Dennison, C. R., Wayman, D. M., Pronovost, P. J., & Needham, D. M. (2007). Systematic review identifies number of strategies important for retaining study participants. *Journal of Clinical Epidemiology*, *60*(8), 757-e1.

Robles, N., Flaherty, D. G., & Day, N. L. (1994). Retention of resistant subjects in longitudinal studies: Description and procedures. *The American Journal of Drug and Alcohol Abuse*, *20*, 87-100.

Royce, D. (2008). Research methods in social work (5th edition). Belmont, CA: Thomson Brooks/Cole.

Scott, C. K. (2004). A replicable model for achieving over 90% follow-up rates in longitudinal studies of substance abusers. *Drug and Alcohol Dependence*, *74*, 21-36.

Sobell, L. C., Brown, J., Leo, G. I., & Sobell, M. B. (1996). The reliability of the Alcohol Timeline Followback when administered by telephone and by computer. *Drug and Alcohol Dependence*, *42*, 49-54.

Sobell, L. C., & Sobell, M. B. (1992). Timeline follow-back. In *Measuring alcohol consumption* (pp. 41-72). New York, NY: Humana Press.

Sullivan, C. M., Rumptz, M. H., Campbell, R., Eby, K. K., & Davidson, W. S. (1996). Retaining Participants in Longitudinal Community Research: A Comprehensive Protocol. *The Journal of Applied Behavioral Science*, *32*(3), 262-276.

Teplin, L. A., Elkington, K. S., McClelland, G. M., Abram, K. M., Mericle, A. A., & Washburn, J. J. (2005). Major mental disorders, substance use disorders, comorbidity, and HIV-AIDS risk behaviors in juvenile detainees. *Psychiatric Services (Washington, DC)*, *56*(7), 823.

Twitchell, G. R., Hertzog, C. A., Klein, J. L., & Schuckit, M. A. (1992). The anatomy of a follow-up. *British Journal of Addiction*, *87*(9), 1327-1333.

Walton, M. A., Ramanathan, C. S., & Reischl, T. M. (1998). Tracking substance abusers in longitudinal research: Understanding follow-up contact difficulty. *American Journal of Community Psychology*, *26*(2), 233-253.

Watson, N., & Wooden, M. (2009). Identifying factors affecting longitudinal survey response. *Methodology of Longitudinal Surveys*, *1*, 157-182.

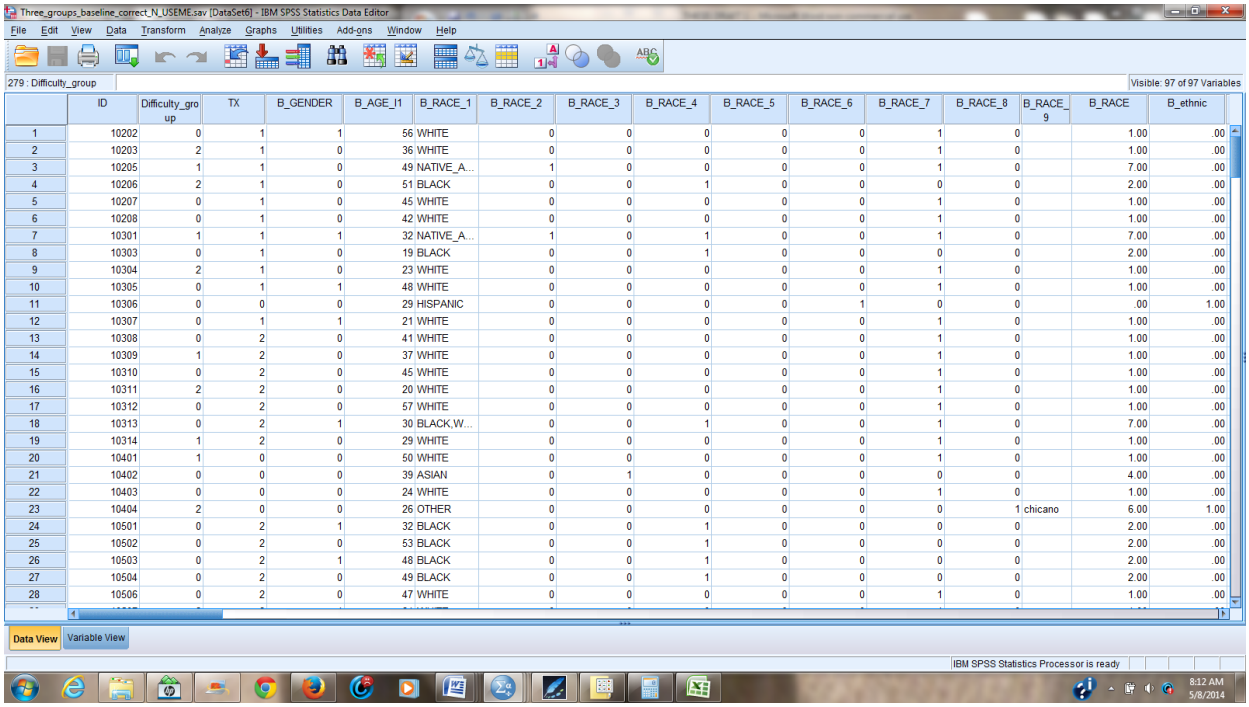
Weisberg, H. F. (2009). *The total survey error approach: A guide to the new science of survey research*. Chicago, IL: University of Chicago Press.

Wright, J. D., Allen, T. L., & Devine, J. A. (1995). Tracking non-traditional populations in longitudinal studies. *Evaluation and Program Planning*, *18*(3), 267-277.

Wutzke, E. K., Conigrave, M., Kogler, B. E., Saunders, J. B., Hall, W. D. S. (2000). Longitudinal research: Methods for maximizing subject follow-up. *Drug and Alcohol Review*, *19*(2), 159-163.

Ziek, K., Beardsley, M., Deren, S., & Tortu, S. (1996). Predictors of follow-up in a sample of urban crack users. *Evaluation and Program Planning*, *19*(3), 219-224.

APPENDIX A. Sample data collection spreadsheet



Three_groups_baseline_correct_N_USEME.sav [DataSet6] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help

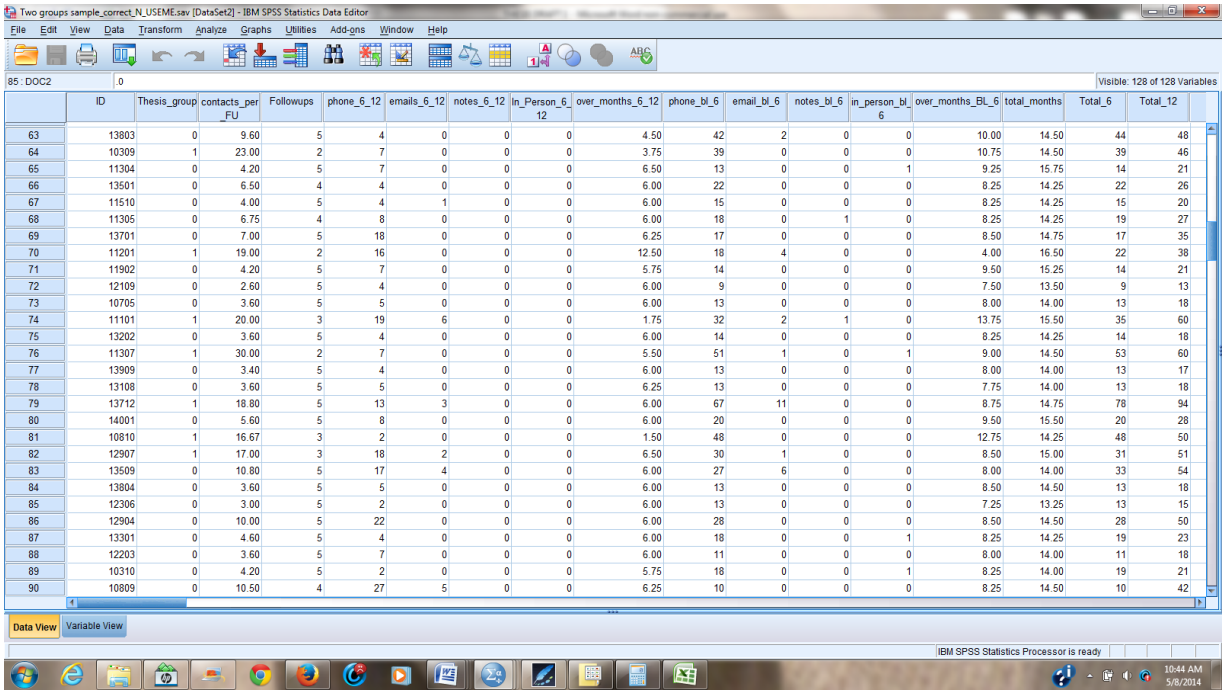
279 : Difficulty_group — Visible: 97 of 97 Variables

	ID	Difficulty_group	TX	B_GENDER	B_AGE_I1	B_RACE_1	B_RACE_2	B_RACE_3	B_RACE_4	B_RACE_5	B_RACE_6	B_RACE_7	B_RACE_8	B_RACE_9	B_RACE	B_ethnic
1	10202	0	1	1	56	WHITE	0	0	0	0	0	1	0		1.00	.00
2	10203	2	1	0	36	WHITE	0	0	0	0	0	1	0		1.00	.00
3	10205	1	1	0	49	NATIVE_A...	1	0	0	0	0	1	0		7.00	.00
4	10206	2	1	0	51	BLACK	0	0	1	0	0	0	0		2.00	.00
5	10207	0	1	0	45	WHITE	0	0	0	0	0	1	0		1.00	.00
6	10208	0	1	0	42	WHITE	0	0	0	0	0	1	0		1.00	.00
7	10301	1	1	1	32	NATIVE_A...	1	0	1	0	0	1	0		7.00	.00
8	10303	0	1	0	19	BLACK	0	0	1	0	0	0	0		2.00	.00
9	10304	2	1	0	23	WHITE	0	0	0	0	0	1	0		1.00	.00
10	10305	0	1	1	48	WHITE	0	0	0	0	0	1	0		1.00	.00
11	10306	0	0	0	29	HISPANIC	0	0	0	0	1	0	0		.00	1.00
12	10307	0	1	1	21	WHITE	0	0	0	0	0	1	0		1.00	.00
13	10308	0	2	0	41	WHITE	0	0	0	0	0	1	0		1.00	.00
14	10309	1	2	0	37	WHITE	0	0	0	0	0	1	0		1.00	.00
15	10310	0	2	0	45	WHITE	0	0	0	0	0	1	0		1.00	.00
16	10311	2	2	0	20	WHITE	0	0	0	0	0	1	0		1.00	.00
17	10312	0	2	0	57	WHITE	0	0	0	0	0	1	0		1.00	.00
18	10313	0	2	1	30	BLACK,W...	0	0	1	0	0	1	0		7.00	.00
19	10314	1	2	0	29	WHITE	0	0	0	0	0	1	0		1.00	.00
20	10401	1	0	0	50	WHITE	0	0	0	0	0	1	0		1.00	.00
21	10402	0	0	0	39	ASIAN	0	1	0	0	0	0	0		4.00	.00
22	10403	0	0	0	24	WHITE	0	0	0	0	0	1	0		1.00	.00
23	10404	2	0	0	26	OTHER	0	0	0	0	0	0	1	chicano	6.00	1.00
24	10501	0	2	1	32	BLACK	0	0	1	0	0	0	0		2.00	.00
25	10502	0	2	0	53	BLACK	0	0	1	0	0	0	0		2.00	.00
26	10503	0	2	1	48	BLACK	0	0	1	0	0	0	0		2.00	.00
27	10504	0	2	0	49	BLACK	0	0	1	0	0	0	0		2.00	.00
28	10506	0	2	0	47	WHITE	0	0	0	0	0	1	0		1.00	.00

Data View | Variable View

IBM SPSS Statistics Processor is ready

8:12 AM 5/8/2014



Two groups sample_correct_N_USEME.sav [DataSet2] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help

85 : DOC2 — Visible: 128 of 128 Variables

	ID	Thesis_group	contacts_per_FU	Followups	phone_6_12	emails_6_12	notes_6_12	In_Person_6_12	over_months_6_12	phone_bl_6	email_bl_6	notes_bl_6	in_person_bl_6	over_months_BL_6	total_months	Total_6	Total_12
63	13803	0	9.60	5	4	0	0	0	4.50	42	2	0	0	10.00	14.50	44	48
64	10309	1	23.00	2	7	0	0	0	3.75	39	0	0	0	10.75	14.50	39	46
65	11304	0	4.20	5	7	0	0	0	6.50	13	0	0	1	9.25	15.75	14	21
66	13501	0	6.50	4	4	0	0	0	6.00	22	0	0	0	8.25	14.25	22	26
67	11510	0	4.00	5	4	1	0	0	6.00	15	0	0	0	8.25	14.25	15	20
68	11305	0	6.75	4	8	0	0	0	6.00	18	0	1	0	8.25	14.25	19	27
69	13701	0	7.00	5	18	0	0	0	6.25	17	0	0	0	8.50	14.75	17	35
70	11201	1	19.00	2	16	0	0	0	12.50	18	4	0	0	4.00	16.50	22	38
71	11902	0	4.20	5	7	0	0	0	5.75	14	0	0	0	9.50	15.25	14	21
72	12109	0	2.60	5	4	0	0	0	6.00	9	0	0	0	7.50	13.50	9	13
73	10705	0	3.60	5	5	0	0	0	6.00	13	0	0	0	8.00	14.00	13	18
74	11101	1	20.00	3	19	6	0	0	1.75	32	2	1	0	13.75	15.50	35	60
75	13202	0	3.60	5	4	0	0	0	6.00	14	0	0	0	8.25	14.25	14	18
76	11307	1	30.00	2	7	0	0	0	5.50	51	1	0	1	9.00	14.50	53	60
77	13909	0	3.40	5	4	0	0	0	6.00	13	0	0	0	8.00	14.00	13	17
78	13108	0	3.60	5	5	0	0	0	6.25	13	0	0	0	7.75	14.00	13	18
79	13712	1	18.80	5	13	3	0	0	6.00	67	11	0	0	8.75	14.75	78	94
80	14001	0	5.60	5	8	0	0	0	6.00	20	0	0	0	9.50	15.50	20	28
81	10810	1	16.67	3	2	0	0	0	1.50	48	0	0	0	12.75	14.25	48	50
82	12907	1	17.00	3	18	2	0	0	6.50	30	1	0	0	8.50	15.00	31	51
83	13509	0	10.80	5	17	4	0	0	6.00	27	6	0	0	8.00	14.00	33	54
84	13804	0	3.60	5	5	0	0	0	6.00	13	0	0	0	8.50	14.50	13	18
85	12306	0	3.00	5	2	0	0	0	6.00	13	0	0	0	7.25	13.25	13	15
86	12904	0	10.00	5	22	0	0	0	6.00	28	0	0	0	8.50	14.50	28	50
87	13301	0	4.60	5	4	0	0	0	6.00	18	0	0	1	8.25	14.25	19	23
88	12203	0	3.60	5	7	0	0	0	6.00	11	0	0	0	8.00	14.00	11	18
89	10310	0	4.20	5	2	0	0	0	5.75	18	0	0	1	8.25	14.00	19	21
90	10809	0	10.50	4	27	5	0	0	6.25	10	0	0	0	8.25	14.50	10	42

Data View | Variable View

IBM SPSS Statistics Processor is ready

10:44 AM 5/8/2014