Adolescent Recovery Capital and Application of Exploratory Methods

By

Emily A. Hennessy

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Community Research and Action

May, 2017

Nashville, Tennessee

Approved:

Emily E. Tanner-Smith, Ph.D.

Andrew J. Finch, Ph.D.

Craig Anne Heflinger, Ph.D.

Kevin J. Grimm, Ph.D.

ACKNOWLEDGEMENTS

The Institute for Social Research and the James Morgan Fund for New Directions in the Analysis of Complex Interactions supported the research reported in Chapters III and IV. I would like to thank James N. Morgan for pioneering the SEARCH method and for this generous funding. I am also grateful to Peter Solenberger and Patrick Shields for their assistance in navigating the fellowship and the SEARCH program. The research reported in Chapter III and IV was made possible by Grant Number R01DA029785-01A1 from the National Institute on Drug Abuse (NIDA). Its contents are solely the responsibility of the author and do not necessarily represent the official views of NIDA or the NIH. The National Survey on Drug Use and Health provided the data for Chapter II; I am grateful for their generosity in making the data publicly available.

I would like to thank my dissertation committee. Emily, I am so grateful that you took the time to provide your analytical, academic, professional, and personal support and mentorship throughout my time at Vanderbilt. Andy, thank you for introducing me to RHSs – and for your thoughtful insights as I wrangled with the recovery capital model. Craig Anne, thank you for your detailed feedback and the many resources you thoughtfully sent my way over the past five years. Kevin, thank you for introducing me to these data mining methods through the ICPSR-SEARCH course and for providing technical and analytical guidance in these methods.

Finally, many thanks are due to my friends and family – I am especially grateful to Joseph, Holly, Robert, and Amie for your support in writing, job searching, and for your friendships. To my parents, who have constantly supported my educational goals with love and prayers. Gary – words cannot express how thankful I am that you decided to take this journey to Nashville (and beyond) with me. You have made this journey more fun and laughter-filled than I ever could have imagined and I look forward to our next chapter(s) together.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Substance use disorders (SUDs) among adolescents are a major public health problem in the United States: according to the 2013 National Survey on Drug Use and Health (SAMHSA, 2014), 5.2% of 12-17 year olds were either abusing or dependent on alcohol or illicit drugs. And in 2010, 7% of SUD treatment admissions in the United States, that is over 130,000 admissions, were for 12-17 year olds (SAMHSA, 2012). However, SUDs are often considered chronic conditions, requiring multiple treatment episodes and continuing care supports posttreatment (Brown, D'Amico, McCarthy, & Tapert, 2001; Ramo, Prince, Roesch, & Brown, 2012; White et al., 2004). Indeed, research has demonstrated that youth seeking SUD treatment do not always successfully complete that treatment (Kaminer, Burleson, Burke, & Litt, 2014; Pugatch, Knight, McGuiness, Sherritt, & Levy, 2014; Winters, Stinchfield, Latimer, & Lee, 2007), and among those that do, more than 45% return to rates of previous use within months of treatment discharge (Anderson, Ramo, Schulte, Cummins, & Brown, 2007; Brown, et al., 2001; Ramo et al., 2012; White et al., 2004). As a result, an SUD requires sustained and multi-pronged intervention and follow-up support and there are ongoing efforts to better understand the multi-layered contexts in which addiction is situated to effectively reach adolescents with an SUD (Gonzales, Anglin, Beattie, Ong, & Glik, 2012). One such perspective that may be useful in this regard is the recovery capital model (Granfield & Cloud, 1999; White & Cloud, 2008).

Thus, to better understand adolescent recovery processes, this dissertation will address the adolescent recovery process by exploring the adolescent recovery capital model, a model that shows promise but has not yet been examined against the adolescent recovery experience (Hennessy, 2017). In addition, I will apply exploratory data methods to demonstrate their utility in addressing adolescent recovery and the adolescent recovery capital model. In this chapter, I first introduce adolescent substance use and problems stemming from heavy use. Next, I present an overview of the adolescent treatment and recovery process as it relates to the recovery capital model. Finally, I provide a brief introduction to each of the three empirical dissertation papers.

Adolescent Substance Use

Excessive use of substances during adolescence can have severe effects on future life outcomes, including altering the developing brain and damaging cognitive functioning (Brown & Tapert, 2004; Hanson, Medina, Padula, Tapert, & Brown, 2011; Lisdahl, Wright, Kirchner-Medina, Maple, & Shollenbarger, 2014). For example, research has demonstrated that heavy alcohol and drug use can result in significantly diminished memory capabilities and executive functioning among adolescents, with effects persisting many years later (Brown & Tapert, 2004; Hanson et al., 2011). A review of cannabis research also demonstrated similar negative effects among adolescents who used marijuana, including poorer attention and verbal memory and a reduction in IQ scores compared to non- or non-regular users (Lisdahl et al., 2014).

Among adolescent heavy substance users, detrimental effects on achievement and social outcomes often extend into young adulthood. For example, heavy marijuana use during adolescence has been linked to a reduced likelihood of attending postsecondary education (Homel, Thompson, & Leadbeater, 2014). And, among adolescents who completed SUD treatment, those who had abstained from use four years later had better educational attainment and employment status than individuals who returned to heavy use (Brown et al., 2001). Early marijuana or alcohol use has also been linked to a reduced likelihood of adult marriage or an increased likelihood of divorce (Menasco & Blair, 2014).

Adolescent Treatment

Given the scope of adolescent substance use, a number of treatment facilities have been introduced into the public health system in the United States to address SUDs among adolescents. Although originally designed to treat adults, these facilities now increasingly serve adolescents and offer developmentally appropriate services for youth or treatment options designed solely for adolescents (Black & Chung, 2014; White, Dennis, & Tims, 2002). There are a variety of treatment options for adolescents, and the exact trajectory of treatment service use depends on the individual and the severity of his/her substance use problem as well as on the resources of the family. Adolescent SUD treatment consists of five broad levels of treatment varying in intensity: intensive inpatient, inpatient, intensive outpatient, outpatient, and early intervention, such as screening and brief intervention (American Society of Addiction Medicine [ASAM], 2013). A referral to a particular level of care depends on many factors, including the adolescent's incoming substance use severity, comorbidities, pressures from legal and school systems, and potential for relapse. Completion of one level of care leads to a reassessment and potential placement in a stepped-down level of care or continuing care services.

Many adolescents who need SUD treatment do not enroll in or complete treatment. For example, studies have shown that approximately 10-20% of participants fail to fully engage in treatment (Pugatch et al., 2014; Rohde, Waldron, Turner, Brody, & Jorgensen, 2014). A variety of factors affect adolescent treatment engagement and effectiveness including mood disorders (Pugatch et al., 2014), age of substance use initiation (Kennedy & Minami, 1993), family estrangement (Winters, Tanner-Smith, Bresani, & Meyers, 2014), history of deviant behavior

(Winters et al., 2014), peer drug use environment (Winters et al., 2014), severity of comorbid psychopathology (Kennedy & Minami, 1993), and broader contextual factors such as community-level characteristics (Jones, Heflinger, & Saunders, 2007).

Although some adolescents remain abstinent following discharge from SUD treatment, research has demonstrated that around one-half of adolescents released from treatment relapse within three or six months (Brown, Vik, & Creamer, 1989; Cornelius et al., 2003; Kennedy & Minami, 1993; Spear, Ciesla, & Skala, 1999). Longitudinal studies have also demonstrated that a majority of adolescents treated for an SUD return to some level of use (Brown et al., 2001; Stanger, Ryan, Scherer, Norton, & Budney, 2015), even with continuing care supports in place (Burleson, Kaminer, & Burke, 2012).

Adolescent Continuing Care

Given high rates of relapse after formal SUD treatment, continuing care supports such as the 12-Step programs Alcoholics Anonymous (AA) or Narcotics Anonymous (NA) are available in the community. Twelve-Step programs were originally designed for adults, however, the model was eventually used in adolescent treatment programs and practitioners encourage adolescents to visit AA/NA meetings as part of their continuing care posttreatment (Kelly & Myers, 2007; Sussman, 2010). The 12-Step model is focused on abstinence and involves group meetings and having a recovery sponsor, that is, an individual who can support another's recovery process through phone check-ins or meetings. The meetings and sponsorship are free and thus, if present in the community, can be useful continuing care supports. Yet, 12-Step meetings are typically geared toward adults and are often comprised of a small proportion of youth (Alcoholics Anonymous, 2014; Sussman, 2010). Twelve-Step meetings may not be an effective continuing care tool for many youth, as youth may not align with the "abstinence for life" perspective, the focus on spirituality, or may have difficulty empathizing with the experience of older attendees (Kelly, Myers, & Brown, 2005; Kelly, Pagano, Stout, & Johnson, 2011; Sussman, 2010). Indeed, small percentages of adolescents attend 12-Step meetings following treatment, and among that group, attendance tends to decrease over time (Kennedy & Minami, 1993; Kelly et al., 2000; Kelly, Brown, Abrantes, Kahler, & Myers, 2008); however, among adolescents that remain in 12-Step programs, their attendance has a significant, positive relationship with maintaining abstinence (Hennessy & Fisher, 2015).

Another continuing care support for youth with SUDs is the recovery high school (RHS), which addresses academic advancement and recovery maintenance among adolescents who have completed treatment for an SUD (Finch & Frieden, 2014). These institutions have strict enrollment criteria that involve abstinence or a desire for abstinence and often require that adolescents have completed some formal SUD treatment. The primary focus in RHSs are academics, but the schools also incorporate recovery-specific elements into the day, such as a daily group check-in, community service, and individual counseling sessions (Moberg & Finch, 2007). Depending on the location and policies of the educational system, the schools may be free of charge for students (Finch, Karakos, & Hennessy, 2016). Descriptive studies have provided support for their effectiveness as a continuing care support and in enabling students to successfully complete high school and engage in college or the workforce (Kochanek, 2008; Moberg & Finch, 2007).

Recovery and a Theoretical Framework

Recovery from an SUD is a cyclical process, and often involves adolescents returning to treatment and continuing care services multiple times. Thus, SUD treatment and continuing care are often considered two separate stages in the recovery process, with those in treatment initially exhibiting more problematic substance use behaviors and problems, and those in continuing care having more control over their substance use. For the purposes of this dissertation, recovery will be discussed as a process and not as an outcome. This is important to note as the definition of recovery is under debate (Arndt & Taylor, 2007; Betty Ford Institute Consensus Panel, 2007; Hser & Anglin, 2011; White, 2007), and among adolescents this definition is even more ambiguous. Indeed, it may be inappropriate to label an adolescent as "recovered" from an SUD, given their early stage in the development process. Thus, one definition given by William White is useful in this regard:

> Recovery is the experience (a process and a sustained status) through which individuals, families, and communities impacted by severe alcohol and other drug (AOD) problems utilize internal and external resources to voluntarily resolve these problems, heal the wounds inflicted by AOD-related problems, actively manage their continued vulnerability to such problems, and develop a healthy, productive, and meaningful life. (White, 2007, p. 236)

Similarly, the Betty Ford Institute Consensus Panel defined recovery from substance dependence as "a voluntarily maintained lifestyle characterized by sobriety, personal health, and citizenship" (2007, p. 222). Thus, recovery as a process involves an adolescent voluntarily working through their substance use issues and related problems using structured supports and relationships.

Previous research has addressed adolescent recovery using different theoretical frameworks because numerous types of factors influence the recovery process and these factors interact differently for different individuals (e.g., see Hersh, Curry, & Kaminer, 2014; Mason, Malott, & Knoper, 2009). Indeed, from a social-ecological perspective (Bronfenbrenner 1977; 1994), multiple and diverse factors at the individual level (e.g., motivation, self-efficacy: Kelly, Myers, & Brown, 2000; 2002) and within various microsystems (e.g., family, peer networks: Mason, Mennis, Linker, Bares, & Zaharakis, 2014) as well as the mesosystem or larger community context (e.g., rurality/urbanicity, availability of treatment services, county-level educational attainment and criminal activity: Heflinger & Christens, 2006; Jones et al., 2007) have been linked to addiction and recovery behaviors and treatment service utilization. Attention has also been given to understanding the interaction between microsystems by studying the relative influence of family and peers on substance use behavior (Mrug & McCay, 2013) as well as the broader macrosystem in relation to resource availability (Finch et al., 2016).

In light of these complexities, *recovery capital*, the accumulation of all potential resources available for an individual to use in recovery, has been proposed as a broad framework for exploring the full range of resources that enable recovery from an SUD (Granfield & Cloud, 1999). Although there are multiple frameworks that could be used to study adolescents in recovery, previous research has ignored different salient elements of the recovery experience or focused on specific pathways instead of the broader amount of resources that can be used in recovery. The recovery capital model is thus a recovery-specific risk and protective factor model that highlights resources beneficial for recovery within an ecological framework that attends to individual, interindividual, and community factors.

Adolescent recovery capital. The recovery capital model has been studied in depth with adult samples, (e.g., see Hennessy, 2017); however, this dissertation uses an adapted adolescent recovery capital model, generated from previous recovery capital models (Granfield & Cloud, 1999; Hewitt, 2007; White & Cloud, 2008) to fit the adolescent experience and includes four primary domains (See Figure 1): *human, financial, social, and community recovery capital*. There is also the possibility of *growth capital*, which represents the synergy among capital in these four domains resulting in exponential growth in recovery. Growth capital is outside the scope of the proposed studies so will not be discussed further (see Hewitt, 2007 for further discussion of this concept).

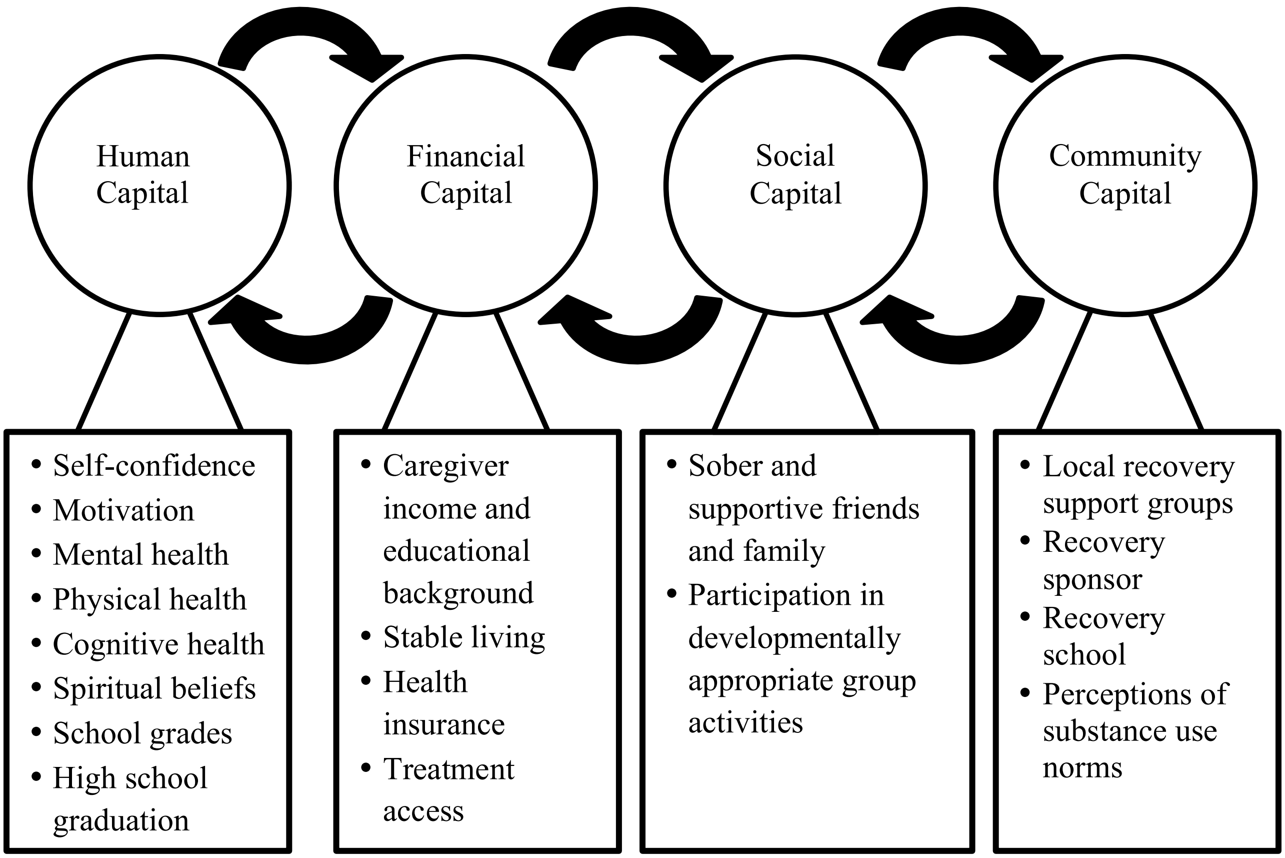

Figure 1. Adolescent Model of Recovery Capital
Note. The black arrows represent *growth capital*, the growth generated through synergy of the other capital dimensions.

Human recovery capital. The first domain, *human recovery capital,* is any personal characteristic that one can use to achieve personal goals. Multiple individual-level factors have been studied as relevant to adolescent treatment outcomes, and thus could be considered relevant to understanding human capital. These include factors such as mental health (Hersh et al., 2014; Rohde et al., 2014; Winters et al., 2007; Yu, Buka, Fitzmaurice, & McCormick, 2006), abstinence motivation (Kelly et al., 2000; 2002), and spirituality (Chi, Kaskutas, Sterling, Campbell, & Weisner, 2009; Ritt-Olson, Milam, Unger, Trinidad, Teran, Dent, & Sussman, 2004).

Success and engagement at school are also internal resources supporting adolescent recovery efforts. Among adults, level of education and employment skills and opportunities were considered vital to human capital (Best, Gow, Knox, Taylor, Groshkova, & White, 2012; Granfield & Cloud, 1999; Skogens & von Greiff, 2014) and among young adults, postsecondary education or learning life skills generated and supported recovery capital (Keane, 2011; Terrion, 2012; 2014): these resources can lead to the creation of additional capital, connect individuals with others, and offer hope and confidence in one's potential outside of substance use. In the case of adolescents who may have limited experience on the job market, school attendance, engagement, and meeting academic goals might be more relevant than employment experience. These factors could motivate positive change as well as enable the adolescent to replace substance use activity with positive growth activities and engage with prosocial peers.

Financial recovery capital. *Financial recovery capital* refers to material resources that could be used toward successful recovery (Granfield & Cloud, 1999). Adolescent financial capital consists of factors such as family socioeconomic status (SES; measured through caregiver education levels, employment, income), health insurance, and access to treatment.

The exact nature of the relationship between parental income and adolescent treatment and recovery has not been studied in great detail, but is likely complex. For example, SES has had inconsistent relationships with alcohol use and associated problems depending on the age of the adolescent (Kendler, Gardner, Hickman, Heron, Macleod, Lewis, & Dick, 2014) — whereas low SES has been linked to cognitive impairments among cannabis dependent youth (Vo, Schacht, Mintzer, & Fishman, 2014), high SES has been associated with a greater likelihood of abstinence during recovery among adolescents (Brown, Myers, Mott, & Vik, 1994).

Adolescents with previous treatment experiences may also be more likely to seek out and engage in continuing care supports. This could be due to increased community capital because of an increase in financial capital, that is, access to treatment produces more knowledge of and access to posttreatment supports. However, it could also be because of greater addiction severity and the resulting need for more intensive follow-up supports (Chi, Campbell, Sterling, & Weisner, 2011; Kelly et al., 2008; Kelly, Dow, Yeterian, & Kahler, 2010).

Social recovery capital. *Social recovery capital* enables one to effectively bond with family, peers, and community institutions and consists of resources available to an individual through these relationships (Granfield & Cloud, 1999). It is also the presence of family and peers that support recovery efforts (White & Cloud, 2008). For adolescents, social capital consists of sober and supportive friends and family and participation in positive group activities that attempt to build social networks, such as through sports or faith-based organizations.

As research has demonstrated, interactions with family and friends can have positive and/or negative influences on substance use outcomes (Mason et al., 2014); thus, social influences impact the process of recovery in different ways across individuals. For example, adolescents in an antisocial peer group have a higher risk of substance use than those not associating with such a group (Lamont, Woodlief, & Malone, 2014). Additionally, social factors and situations have been identified as highly predictive of initial and subsequent relapses for adolescents in recovery (Ramo & Brown, 2008; Ramo et al., 2012). Indeed, adolescent abstinence behaviors are associated with higher family functioning scores (Brown et al., 1994), higher levels of social support (Brown et al., 2001), and the number of friends and family actively supporting recovery efforts (Chi et al., 2009).

Studies of SUD treatment where the family system is the point of intervention are also useful in highlighting how important family is to youth substance use outcomes (Tanner-Smith, Wilson, & Lipsey, 2013). For example, family therapy, compared to peer group therapy, has been found to reduce youth substance use, substance use frequency, substance use problems, delinquency, and internalized distress largely through changing parenting practices (Henderson, Rowe, Dakof, Hawes, & Liddle, 2009; Liddle, Rowe, Dakof, Henderson, & Greenbaum, 2009). Additionally, intervention groups with home and office-based treatment involving participants' parents fared significantly better than office-based treatment alone (Stanger et al., 2015). Thus, family influences are important to consider in the adolescent recovery capital model.

Community recovery capital. *Community recovery capital* includes all community-level resources that are related to addiction and recovery (White & Cloud, 2008). Although social capital enables an adolescent to bond and engage with others, community capital is the availability of recovery resources in the community in which an adolescent can engage. Thus, for adolescents, community capital consists of continuing care supports including availability of and attendance at self-help support groups (e.g., 12-Step programs like AA or alternative peer groups) and RHSs.

Community capital also encompasses *cultural recovery capital*. Cultural recovery capital has been defined in a few different ways, and for the purposes of this dissertation, a combination of definitions will be used, following Burns and Marks (2013). In the original description of community capital, cultural capital was incorporated as access to culturally appropriate forms of treatment, such as tailored programming for individuals in specific ethnic communities (White & Cloud, 2008). Cultural capital has also been described as individual values and behavioral patterns generated from membership within a certain cultural group that foster recovery (Burns

& Marks, 2013; Cloud & Granfield, 2008). Thus, among adolescents, perceived substance use and norms of appropriate substance use among peers and family are an indication of cultural capital. For example, perceived substance use among peers has been linked to a tendency to use substances (Mason et al., 2014), especially among younger adolescents (Wambeam, Canen, Linkenbach, & Otto, 2014). Cultural capital thus constitutes access to culturally appropriate forms of treatment and microsystem norms around substance use behaviors.

Aim and Brief Introduction to Three Papers

A recent systematic review of the literature revealed that recovery capital had been studied in depth with adult populations (Hennessy, 2017). However, although recovery capital has been used to explore and explain adult recovery experiences (Best et al., 2012; Duffy & Baldwin, 2013; Neale, Nettleton, & Pickering, 2014), it has not yet been explored with adolescents, who have markedly different recovery patterns. In addition, studies have not systematically applied a singular recovery capital model, using anywhere from three to five dimensions. Thus, this dissertation will explore the proposed adolescent recovery capital framework to aid others in systematically applying it to adolescents and to identifying additional mechanisms of change in adolescent recovery processes to explore in future research. Additionally, this dissertation will highlight exploratory methods that can be used when studying adolescent recovery, and demonstrate how these methods can be used in quasi-experimental designs.

In the first empirical paper (Chapter II), an exploratory latent variable mixture model was used to assess whether there are different classes of recovery capital among adolescents identified as "needing [SUD] treatment" in the previous year, and if so, what characteristics lead to different categorizations of recovery capital. This paper analyzes a subset of data for the non-institutionalized adolescent population, from the 2012 National Survey on Drug Use and Health.

The second empirical paper (Chapter III) explored predictors of access to one type of community recovery capital, recovery high schools (RHS), using four different methods: logistic regression, and three data mining approaches (SEARCH, classification trees, and the ensemble method of random forest). Using these methods, variables relevant to each recovery capital domain, and relevant to overall recovery capital, were used to predict access to RHSs among students that attended or did not attend the schools. This paper analyzed data from an ongoing observational study of adolescents in recovery, "Effectiveness of Recovery High Schools as Continuing Care."

The third and final empirical paper (Chapter IV) expanded the results from the second paper by developing sets of propensity scores to balance the non-RHS and RHS groups so that treatment effects could be estimated. The focus in this paper is on using exploratory methods to select covariates for propensity score estimation models. Propensity scores can be used to address potential bias in treatment effects estimates, but the correct covariates (covariates that would influence selection into treatment as well as the outcome of interest) must be collected and included in the estimation of the propensity score. Three sets of estimated propensity scores, one from logistic regressions, classification trees, and random forests, were developed[1]. Using the propensity score, participants were stratified and included covariates were then assessed for balance. A multilevel analysis accounting for school clusters then used these estimated propensity scores to examine how each method performed in predicting substance use at follow-

[1] Unlike empirical paper 2, the SEARCH method could not be used in paper 3 because it did not identify enough covariates for propensity score estimation.

up measurements for RHS versus non-RHS groups. This paper again analyzed data collected from the study, “Effectiveness of Recovery High Schools as Continuing Care”.

Finally, implications from these three papers will be discussed (Chapter V). These three papers contribute to the broader literature on adolescent recovery and provide a deeper understanding of the recovery process in context, albeit for different samples of youth, within different recovery stages. These papers will therefore provide a way forward for researchers wishing to study recovery capital among adolescents in a systematic and replicable way. The papers will also demonstrate the usefulness of exploratory methods to this complex social and public health issue.

References

Alcoholics Anonymous. (2014). *2014 membership survey*. New York, NY: Alcoholics Anonymous World Services, Inc. Retrieved from http://www.aa.org/assets/en_US/p-48_membershipsurvey.pdf

American Society of Addiction Medicine (ASAM). (2013). *The ASAM Criteria. Treatment criteria addictive, substance-related, and co-occurring conditions*. The Change Companies: Nevada.

Arndt, S., & Taylor, P. (2007). Commentary on "defining and measuring 'recovery'". *Journal of Substance Abuse Treatment*, *33*(3), 275 - 276. doi:10.1016/j.jsat.2007.07.008

Best, D., Gow, J., Knox, T., Taylor, A., Groshkova, T., & White, W. (2012). Mapping the recovery stories of drinkers and drug users in Glasgow: Quality of life and its associations with measures of recovery capital. *Drug and Alcohol Review*, *31*(3), 334-41. doi:10.1111/j.1465-3362.2011.00321.x

Betty Ford Institute Consensus Panel. (2007). What is recovery? A working definition from the Betty Ford Institute. *Journal of Substance Abuse Treatment*, *33*(3), 221 - 228. doi:10.1016/j.jsat.2007.06.001

Black, J. J., & Chung, T. (2014). Mechanisms of change in adolescent substance use treatment: How does treatment work? *Substance Abuse*, *35*(4), 344-51. doi:10.1080/08897077.2014.925029

Bronfenbrenner, U. (1977). Towards an experimental ecology of human development. *American Psychologist,* 518-531.

Bronfenbrenner, U. (1994). Ecological models of human development. In *International Encyclopedia of Education, Vol. 3, 2nd Ed.* Oxford: Elsevier. Reprinted in: M. Gauvain & M. Cole (Eds.), *Readings on the Development of Children, 2nd Ed.* (1993, pp. 37-43). NY: Freeman.

Brown, S. A., D'Amico, E. J., McCarthy, D. M., & Tapert, S. F. (2001). Four-year outcomes from adolescent alcohol and drug treatment. *Journal of Studies on Alcohol*, *62*(3), 381-388.

Brown, S. A., Myers, M. G., Mott, M. A., & Vik, P. W. (1994). Correlates of success following treatment for adolescent substance abuse. *Applied & Preventive Psychology, 3,* 61-73.

Brown, S. A., & Tapert, S. F. (2004). Adolescence and the trajectory of alcohol use: Basic to clinical studies. *Annals of the New York Academy of Sciences*, *1021*, 234-44. doi:10.1196/annals.1308.028

Brown, S., Vik, P., & Creamer, V. (1989). Characteristics of relapse following adolescent substance abuse treatment. *Addictive Behaviors, 14*, 43-52.

Burleson, J. A., Kaminer, Y., & Burke, R. H. (2012). Twelve-month follow-up of aftercare for adolescents with alcohol use disorders. *Journal of substance abuse treatment*, *42*(1), 78-86.

Burns, J., & Marks, D. (2013). Can recovery capital predict addiction problem severity? *Alcoholism Treatment Quarterly, 31*(3), 303-320.

Chi, F. W., Campbell, C. I., Sterling, S., & Weisner, C. (2011). Twelve-Step attendance trajectories over 7 years among adolescents entering substance use treatment in an integrated health plan. *Addiction*, *107*(5), 933-942. doi:10.1111/j.1360-0443.2011.03758.x

Chi, F. W., Kaskutas, L. A., Sterling, S., Campbell, C. I., & Weisner, C. (2009). Twelve-Step affiliation and 3-year substance use outcomes among adolescents: Social support and religious service attendance as potential mediators. *Addiction, 104*(6), 927-939. doi:doi:10.1111/j.1360-0443.2009.02524.x

Cloud, W., & Granfield, R. (2008). Conceptualizing recovery capital: Expansion of a theoretical construct. *Substance Use & Misuse, 43*(12-13), 1971-86. doi:10.1080/10826080802289762

Cornelius, J. R., Maisto, S. A., Pollock, N. K., Martin, C. S., Salloum, I. M., Lynch, K. G., & Clark, D. B. (2003). Rapid relapse generally follows treatment for substance use disorders among adolescents. *Addictive Behaviors*, *28*(2), 381-6.

Duffy, P., & Baldwin, H. (2013). Recovery post treatment: Plans, barriers and motivators. *Substance Abuse Treatment, Prevention, & Policy*, *8*, 6. doi:10.1186/1747-597X-8-6

Finch, A. J., & Frieden, G. (2014). The ecological and developmental role of recovery high schools. *Peabody Journal of Education*, *89*(2), 271-287. doi:10.1080/0161956X.2014.897106

Finch, A., Karakos, H., & Hennessy, E. A. (2016). Exploring the policy context around Recovery High Schools: A case study of schools in Minnesota and Massachusetts. In T. Reid (Ed.), *Substance Abuse: Influences, Treatment Options and Health Effects* (pp. 109-138). New York: Nova Science Publishers.

Gonzales, R., Anglin, M. D., Beattie, R., Ong, C. A., & Glik, D. C. (2012). Understanding recovery barriers: Youth perceptions about substance use relapse. *American Journal of Health Behavior*, *36*(5), 602-14. doi:10.5993/AJHB.36.5.3

Granfield, R., & Cloud, W. (1999). *Coming clean: Overcoming addiction without treatment*. New York: New York University Press.

Hanson, K. L., Medina, K. L., Padula, C. B., Tapert, S. F., & Brown, S. A. (2011). Impact of adolescent alcohol and drug use on neuropsychological functioning in young adulthood: 10-year outcomes. *Journal of Child & Adolescent Substance Abuse*, *20*(2), 135-154. doi:10.1080/1067828X.2011.555272

Heflinger, C. A., & Christens, B. (2006). Rural behavioral health services for children and adolescents: An ecological and community psychology analysis. *Journal of Community Psychology*, *34*(4), 379-400. doi:10.1002/jcop.20105

Henderson, C. E., Rowe, C. L., Dakof, G. A., Hawes, S. W., & Liddle, H. A. (2009). Parenting practices as mediators of treatment effects in an early-intervention trial of multidimensional family therapy. *The American Journal of Drug and Alcohol Abuse*, *35*(4), 220-226.

Hennessy, E. A. (2017). Recovery Capital: A systematic review of the literature. *Addiction Research and Theory*, in press.

Hennessy, E. A., & Fisher, B. W. (2015). A meta-analysis exploring the relationship between 12-step attendance and adolescent substance use relapse. *Journal of Groups in Addiction & Recovery*, *10*(1), 79-96. doi:10.1080/1556035X.2015.999621

Hersh, J., Curry, J. F., & Kaminer, Y. (2014). What is the impact of comorbid depression on adolescent substance abuse treatment? *Substance Abuse*, *35*(4), 364-75. doi:10.1080/08897077.2014.956164

Homel, J., Thompson, K., & Leadbeater, B. (2014). Trajectories of marijuana use in youth ages 15-25: Implications for postsecondary education experiences. *Journal of Studies on Alcohol and Drugs*, *75*(4), 674-83.

Hser, Y. -I., & Anglin, M. D. (2011). Addiction treatment and recovery careers. In W. L. White & J. F. Kelly (Eds.), *Addiction recovery management: Theory, research, and practice* (pp. 9-29). Totowa, NJ: Humana Press. doi:10.1007/978-1-60327-960-4_2

Jones, D. L., Heflinger, C. A., & Saunders, R. C. (2007). The ecology of adolescent substance abuse service utilization. *American Journal of Community Psychology*, *40*(3-4), 345-58. doi:10.1007/s10464-007-9138-8

Kaminer, Y., Burleson, J. A., Burke, R., & Litt, M. D. (2014). The efficacy of contingency management for adolescent cannabis use disorder: A controlled study. *Substance Abuse*, *35*(4), 391-8. doi:10.1080/08897077.2014.933724

Keane, M. (2011). *The role of education in developing recovery capital in recovery from substance addiction*. Dublin: Soilse Drug Rehabilitation Project. Archived by WebCite® at http://www.webcitation.org/6UZxzeZpo

Kelly, J. F., Brown, S. A., Abrantes, A., Kahler, C. W., & Myers, M. (2008). Social recovery model: An 8-year investigation of adolescent 12-step group involvement following inpatient treatment. *Alcoholism, Clinical and Experimental Research*, *32*(8), 1468-78. doi:10.1111/j.1530-0277.2008.00712.x

Kelly, J. F., Dow, S. J., Yeterian, J. D., & Kahler, C. W. (2010). Can 12-step group participation strengthen and extend the benefits of adolescent addiction treatment? A prospective analysis. *Drug and Alcohol Dependence*, *110*(1–2), 117 - 125. doi:10.1016/j.drugalcdep.2010.02.019

Kelly, J. F., & Myers, M. G. (2007). Adolescents' participation in alcoholics anonymous and narcotics anonymous: Review, implications and future directions. *Journal of Psychoactive Drugs*, *39*(3), 259-269.

Kelly, J. F., Myers, M. G., & Brown, S. A. (2000). A multivariate process model of adolescent 12-step attendance and substance use outcome following inpatient treatment. *Psychology of Addictive Behaviors*, *14*(4), 376-389.

Kelly, J. F., Myers, M. G., & Brown, S. A. (2002). Do adolescents affiliate with 12-step groups? A multivariate process model of effects. *Journal of Studies on Alcohol*, *63*(3), 293-304.

Kelly, J. F., Pagano, M. E., Stout, R. L., & Johnson, S. M. (2011). Influence of religiosity on 12-step participation and treatment response among substance-dependent adolescents. *Journal of Studies on Alcohol and Drugs*, *72*(6), 1000-1011.

Kendler, K. S., Gardner, C. O., Hickman, M., Heron, J., Macleod, J., Lewis, G., & Dick, D. M. (2014). Socioeconomic status and alcohol-related behaviors in mid-to-late adolescence in the Avon Longitudinal study of parents and children. *Journal of Studies on Alcohol and Drugs*, *75*(4), 541-545.

Kennedy, B. P., & Minami, M. (1993). The Beech Hill hospital/outward bound adolescent chemical dependency treatment program. *Journal of Substance Abuse Treatment, 10*, 395-406.

Kochanek, T. T. (2008). *Recovery high schools in Massachusetts: A promising, comprehensive model for adolescent substance abuse and dependence*. Retrieved from http://massrecoveryhs.org/documents/RecoveryHighSchooloverview.pdf

Lamont, A., Woodlief, D., & Malone, P. (2014). Predicting high-risk versus higher-risk substance use during late adolescence from early adolescent risk factors using latent class analysis. *Addiction Research and Theory*, *22*(1), 78-89.

Liddle, H. A., Rowe, C. L., Dakof, G. A., Henderson, C. E., & Greenbaum, P. E. (2009). Multidimensional family therapy for young adolescent substance abuse: twelve-month outcomes of a randomized controlled trial. *Journal of Consulting and Clinical Psychology*, *77*(1), 12.

Lisdahl, K. M., Wright, N. E., Kirchner-Medina, C., Maple, K. E., & Shollenbarger, S. (2014). Considering cannabis: The effects of regular cannabis use on neurocognition in adolescents and young adults. *Current Addiction Reports*, *1*(2), 144-156. doi:10.1007/s40429-014-0019-6

Mason, M. J., Malott, K., & Knoper, T. (2009). Urban adolescents' reflections on brief substance use treatment, social networks, and self-narratives. *Addiction Research & Theory*, *17*(5), 453-468.

Mason, M. J., Mennis, J., Linker, J., Bares, C., & Zaharakis, N. (2014). Peer attitudes effects on adolescent substance use: The moderating role of race and gender. *Prevention Science*, *15*(1), 56-64. doi:10.1007/s11121-012-0353-7

Menasco, M. A., & Blair, S. L. (2014). Adolescent substance use and marital status in adulthood. *Journal of Divorce & Remarriage*, *55*(3), 216-238. doi:10.1080/10502556.2014.887382

Moberg, D. P., & Finch, A. J. (2007). Recovery high schools: A descriptive study of school programs and students. *Journal of Groups in Addiction & Recovery*, *2*, 128-161. doi:10.1080/15560350802081314

Mrug, S., & McCay, R. (2013). Parental and peer disapproval of alcohol use and its relationship to adolescent drinking: Age, gender, and racial differences. *Psychology of Addictive Behaviors*, *27*(3), 604-14. doi:10.1037/a0031064

Neale, J., Nettleton, S., & Pickering, L. (2014). Gender sameness and difference in recovery from heroin dependence: A qualitative exploration. *International Journal of Drug Policy*, *25*(1), 3-12. doi:10.1016/j.drugpo.2013.08.002

Pugatch, M., Knight, J. R., McGuiness, P., Sherritt, L., & Levy, S. (2014). A group therapy program for opioid-dependent adolescents and their parents. *Substance Abuse*, *35*(4), 435-41. doi:10.1080/08897077.2014.958208

Ramo, D. E., & Brown, S. A. (2008). Classes of substance abuse relapse situations: A comparison of adolescents and adults. *Psychology of Addictive Behaviors*, *22*(3), 372-9. doi:10.1037/0893-164X.22.3.372

Ramo, D. E., Prince, M. A., Roesch, S. C., & Brown, S. A. (2012). Variation in substance use relapse episodes among adolescents: A longitudinal investigation. *Journal of Substance Abuse Treatment*, *43*(1), 44-52. doi:10.1016/j.jsat.2011.10.003

Ritt-Olson, A., Milam, J., Unger, J. B., Trinidad, D., Teran, L., Dent, C. W., & Sussman, S. (2004). The protective influence of spirituality and "health-as-a-value" against monthly substance use among adolescents varying in risk. *The Journal of Adolescent Health*, *34*(3), 192-199. doi:10.1016/j.jadohealth.2003.07.009

Rohde, P., Waldron, H. B., Turner, C. W., Brody, J., & Jorgensen, J. (2014). Sequenced versus coordinated treatment for adolescents with comorbid depressive and substance use disorders. *Journal of Consulting and Clinical Psychology*, *82*(2), 342-8. doi:10.1037/a003580

Skogens, L. & von Greiff, N. (2014). Recovery capital in the process of change-differences and similarities between groups of clients treated for alcohol or drug problems. *European Journal of Social Work, 17*(1), 58-73.

Spear, S. F., Ciesla, J. R., & Skala, S. Y. (1999). Relapse patterns among adolescents treated for chemical dependency. *Substance Use & Misuse*, *34*(13), 1795-1815

Stanger, C., Ryan, S. R., Scherer, E. A., Norton, G. E., & Budney, A. J. (2015). Clinic- and home-based contingency management plus parent training for adolescent cannabis use disorders. *Journal of the American Academy of Child and Adolescent Psychiatry*, *54*(6), 445-453.e2. doi:10.1016/j.jaac.2015.02.00

Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality. (2012). *Treatment Episode Data Set (TEDS): 2000-2010. National Admissions to Substance Abuse Treatment Services.* DASIS Series S-61, HHS Publication No. (SMA) 12-4701. Rockville, MD: Substance Abuse and Mental Health Services Administration.

Substance Abuse and Mental Health Services Administration (2014). *Results from the 2013 National Survey on Drug Use and Health: Summary of National Findings*, NSDUH Series H-48, HHS Publication No. (SMA) 14-4863. Rockville, MD: Substance Abuse and Mental Health Services Administration.

Sussman, S. (2010). A review of alcoholics anonymous/ narcotics anonymous programs for teens. *Evaluation & the Health Professions*, *33*(1), 26-55. doi:10.1177/0163278709356186

Tanner-Smith, E. E., Wilson, S. J., & Lipsey, M. W. (2013). The comparative effectiveness of outpatient treatment for adolescent substance abuse: A meta-analysis. *Journal of Substance Abuse Treatment*, *44*(2), 145-58. doi:10.1016/j.jsat.2012.05.006

Terrion, J. L. (2012). The experience of post-secondary education for students in recovery from addiction to drugs or alcohol: Relationships and recovery capital. *Journal of Social and Personal Relationships,* 1-21. doi:10.1177/0265407512448276

Terrion, J. L. (2014). Building recovery capital in recovering substance abusers through a service-learning partnership: A qualitative evaluation of a communication skills training program. *Journal of Service-Learning in Higher Education*, *3*, 47-63.

Vo, H. T., Schacht, R., Mintzer, M., & Fishman, M. (2014). Working memory impairment in cannabis- and opioid-dependent adolescents. *Substance Abuse*, *35*(4), 387-90.

Wambeam, R. A., Canen, E. L., Linkenbach, J., & Otto, J. (2014). Youth misperceptions of peer substance use norms: A hidden risk factor in state and community prevention. *Prevention Science, 15*(1), 75-84. doi:10.1007/s11121-013-0384-8

White, W. L. (2007). Addiction recovery: Its definition and conceptual boundaries. *Journal of Substance Abuse Treatment*, *33*(3), 229 - 241. doi:10.1016/j.jsat.2007.04.015

White, W., & Cloud, W. (2008). Recovery capital: A primer for addictions professionals. *Counselor, 9*(5), 22-27.

White, W., Dennis, M., & Tims, F. (2002). Adolescent treatment: Its history and current renaissance. *Counselor Magazine*, *3*(2), 20-25.

White, A. M., Jordan, J. D., Schroeder, K. M., Acheson, S. K., Georgi, B. D., Sauls, G., . . . Swartzwelder, H. S. (2004). Predictors of relapse during treatment and treatment completion among marijuana-dependent adolescents in an intensive outpatient substance abuse program. *Substance Abuse*, *25*(1), 53-9. doi:10.1300/J465v25n01_08

Winters, K. C., Stinchfield, R., Latimer, W. W., & Lee, S. (2007). Long-term outcome of substance-dependent youth following 12-step treatment. *Journal of Substance Abuse Treatment*, *33*(1), 61-9. doi:10.1016/j.jsat.2006.12.003

Winters, K. C., Tanner-Smith, E. E., Bresani, E., & Meyers, K. (2014). Current advances in the treatment of adolescent drug use. *Adolescent Health, Medicine and Therapeutics, 5*, 199-210. doi:10.2147/AHMT.S48053

Yu, J. W., Buka, S. L., Fitzmaurice, G. M., & McCormick, M. C. (2006). Treatment outcomes for substance abuse among adolescents with learning disorders. *The Journal of Behavioral Health Services & Research*, *33*(3), 275-286. doi:10.1007/s11414-006-9023-5

CHAPTER II

A LATENT CLASS EXPLORATION OF ADOLESCENT RECOVERY CAPITAL

Introduction

Adolescent substance use is a major public health issue and can result in substance abuse or dependence leading to lifelong problems. For adolescents diagnosed with a substance use disorder (SUD), there is a high rate of relapse even after formal treatment has been completed (Cornelius et al., 2003; Ramo, Prince, Roesch, & Brown, 2012; Spear, Ciesla, & Skala, 1999). Recovery from problematic substance use has been characterized as “a voluntarily maintained lifestyle characterized by sobriety, personal health, and citizenship” (Betty Ford Institute Consensus Panel, 2007, p. 222) and is often described as a cyclical process requiring multiple supports. As a result, attention has increasingly focused on recovery-related resources for supporting patients after discharge from substance use treatment (White, 2012). Risk and protective factor models have explored a variety of supports and barriers to successful recovery (e.g., Galea, Nandi, & Vlahov, 2004; Moon, Jackson, & Hecht, 2000). Indeed, many factors support adolescent recovery from problematic substance use including being motivated for abstinence (Kelly, Myers, & Brown, 2000; 2002), returning to school (Anderson, Ramo, Cummins, & Brown, 2010), having supportive relationships with family and friends (Godley, Kahn, Dennis, Godley, & Funk, 2005; Hoffman & Su, 1998; Lanham & Tirado, 2011), being in substance-free environments with sober peers (Mason, Mennis, Linker, Bares, & Zaharakis, 2014), and engaging in supportive aftercare services (Hennessy & Fisher, 2015; Kaminer, Burleson, & Burke, 2008). Thus, from an ecological perspective (Bronfenbrenner, 1977; 1994), factors affecting recovery span individual, interpersonal, and community levels, which interact to produce different outcomes for different individuals.

Recovery Capital

Recovery capital has been proposed as one ecological construct to model the resources leading to successful recovery among adults (Granfield & Cloud, 1999; White & Cloud, 2008) and although it has not yet been empirically tested among adolescents (Hennessy, 2017), it has recently been adapted to address the adolescent recovery experience. The recovery capital model is comprised of four primary domains of resources necessary to initiate and sustain recovery: financial, human, social, and community recovery capital. Financial and human recovery capital refers to primarily individual-level factors while social recovery capital addresses interindividual (microsystem) factors and community recovery capital focuses on the broader context and interactions between microsystem levels (mesosystem). All four domains will be briefly described below.

Financial recovery capital refers to material resources that could be used towards recovery (Granfield & Cloud, 1999). Given adolescents' positions as minors under the care of others, adolescent financial recovery capital consists of factors such as caregiver income, health insurance, and access to treatment (Brown, Myers, Mott, & Vik, 1994; Kendler et al., 2014). Financial recovery capital also includes factors such as being in a stable living situation and having basic needs met, such as having enough to eat.

Human recovery capital is comprised of any internal characteristic that an adolescent could use to achieve personal goals in the recovery process. Human recovery capital includes characteristics such as motivation for abstinence (Kelly et al., 2000; 2002), education and employment skills and opportunities (Best, Gow, Knox, Taylor, Groshkova, & White, 2012;

Granfield & Cloud, 1999; Skogens & von Greiff, 2014), mental health (Rohde, Waldron, Turner, Brody, & Jorgensen, 2014; Winters, Stinchfield, Latimer, & Lee, 2007; Yu, Buka, Fitzmaurice, & McCormick, 2006), and religion or spirituality (Chi, Kaskutas, Sterling, Campbell, & Weisner, 2009; Kelly, Pagano, Stout, & Johnson, 2011; Rew & Wong, 2006; Ritt-Olson, Milam, Unger, Trinidad, Teran, Dent, & Sussman, 2004). Human recovery capital resources are theorized to support the recovery process by providing emotional strength and motivation for recovery, promoting alternative activities to replace substance use and generate new forms of fulfillment, and developing the skills to successfully navigate home, school, and/or neighborhood environments external to substance use treatment environments.

Social recovery capital enables an adolescent to effectively bond with others and consists of recovery-supportive relationships and resources made available through these relationships (Granfield & Cloud, 1999; White & Cloud, 2008). Among adolescents in recovery, social factors and situations have been identified as highly predictive of initial and subsequent relapses (Brown, D'Amico, McCarthy, & Tapert, 2001; Brown et al., 1994; Chi et al., 2009; Mason, Malott, & Knoper, 2009; Ramo & Brown, 2008; Ramo et al., 2012). Social recovery capital factors might include associations with sober friends, social interactions in substance-free settings, and positive family dynamics and relationships with parents (Henderson, Rowe, Dakof, Hawes, & Liddle, 2009; Stanger, Ryan, Scherer, Norton, & Budney, 2015; Tanner-Smith, Wilson, & Lipsey, 2013). In contrast, substance-approving attitudes and behaviors of friends and family (e.g., parental supply of alcohol to adolescents) are harmful to the recovery process and thus signify a lack of social recovery capital (Allen, Donohue, Griffin, Ryan, & Turner, 2003; Mason et al., 2014; Mattick et al., 2014; Mrug & McCay, 2013).

Community recovery capital includes all community-level resources related to addiction and recovery (White & Cloud, 2008), such as self-help support groups, recovery sponsors, and a local recovery high school. Community recovery capital also includes individual values and behavioral patterns generated from membership within a cultural group that support abstinence or reduction of problematic use of substances (Burns & Marks, 2013; Cloud & Granfield, 2008; White & Cloud, 2008). For example, among adolescents, perceived substance use among fellow school students increases the tendency to use substances (Wambeam, Canen, Linkenbach, & Otto, 2013). Thus, perceived behaviors around substance use in the local community are relevant cultural indicators of community recovery capital.

Exploring Recovery Capital and the Recovery Process

According to the recovery capital model, resources available to an individual interact within their particular environment and can both be accumulated and depleted (Granfield & Cloud, 1999; 2001). Additionally, capital is often cumulative—i.e., having some recovery capital can lead to the generation of more capital and this process can occur across different ecological levels. For example, by having some social recovery capital, such as a network of sober and recovery-supportive friends, an adolescent also has access to the resources of those friends, thereby generating more possibilities within the other recovery capital domains (e.g., access to financial or other material goods which bolsters their financial recovery capital). Among adults, qualitative studies have supported the contention that the presence of recovery capital is linked to better recovery outcomes (Best, Gow, Taylor, Knox, & White, 2011; Granfield & Cloud, 1999; Terrion, 2012), longer time in recovery (van Melick, McCartney, & Best, 2013), and psychological well-being and high quality of life ratings (Best, Honor, Karpusheff, Loudon, Hall, Groshkova, & White, 2012). Although a growing body of research has studied diverse factors related to adolescent recovery outcomes (Anderson et al., 2010; Godley et al., 2005; Hennessy & Fisher, 2015; Hoffman & Su, 1998; Kaminer et al., 2008; Lanham & Tirado, 2011; Mason et al., 2014; Tanner-Smith et al., 2013), the recovery capital model has not yet been empirically addressed among adolescents (Hennessy, 2017). Thus, this paper will address this gap through an exploration of the four recovery capital domains among a national sample of youth in need of treatment for substance use and will identify whether adolescent characteristics are associated with different recovery capital patterns.

Quantitative Approach to Studying Recovery Capital among Adolescents

A useful quantitative method for exploring patterns of recovery capital among adolescents is latent-variable mixture modeling. This paper applies one type of latent-variable mixture model: latent class analysis, a person-centered approach that can identify qualitatively distinct classes of individuals based on multiple observed characteristics (Collins & Lanza, 2010). Previous studies using mixture models have classified trajectories of alcohol use beginning in adolescence and identified predictors of being in a heavy use trajectory (van der Zwaluw, Otten, Kleinjan, & Engels, 2013), distinguished different classes of adults by their sources of alcohol treatment usage and explored the relationship to later alcohol use outcomes (Mowbray, Glass, & Grinnell-Davis, 2015), and classified adolescents into trajectories of substance use posttreatment to predict outcomes during young adulthood (Anderson et al., 2010). Although these studies have identified classes of individuals related to substance use or trajectories of substance use, there has not yet been an exploration of whether measures of recovery capital resources could distinguish between classes of recovery capital among adolescents in need of substance use treatment.

Study aims. Given that previous studies have demonstrated a diversity of recovery capital experiences among adults (e.g., Best et al., 2012; 2011; Skogens & von Greiff, 2014), and that adolescent recovery experiences differs from adult experiences (Deas, Riggs, Langenbucher, Goldman, & Brown, 2000), this study aimed to explore whether there are different classes of recovery capital among adolescents in need of substance use treatment. Based on previous literature, there are likely at least three distinct classes of recovery capital, that is, individuals with (1) low, (2) medium, and (3) high levels of recovery capital; however, it is also possible that there are other qualitatively distinct recovery capital classes among adolescents. In addition to exploring whether there are different classes of recovery capital, this study also explores whether adolescent characteristics predict membership in a certain recovery capital class, namely adolescents' sex, race/ethnicity, age, and prior receipt of substance use treatment (Becker, Stein, Curry, & Hersh, 2012; Guerrero, Marsh, Duan, Oh, Perron, & Lee, 2013; Lamont, Woodlief, & Malone, 2014; Stevens, Estrada, Murphy, McKnight, & Tims, 2004; Wellman, Contreras, Dugas, O'Loughlin, & O'Loughlin, 2014). Exploring first whether there are differing classes of recovery capital among adolescents based on the four recovery capital domains will further our understanding of using the recovery capital construct among this population. In addition, if there are different patterns of recovery capital among this age group, identifying key predictors of recovery capital patterns may facilitate tailored treatment services and aftercare supports. See Figure 1 for the proposed path diagram.

Methods

Data

This study used cross-sectional data from the 2012 National Survey on Drug Use and Health (NSDUH, publicly available for download; United States Department of Health and

Human Services, 2012).[2] This is a national survey in the United States, conducted to assess the prevalence and correlates of substance use (e.g., alcohol, illicit drugs, and tobacco) among the noninstitutionalized civilian population. A multistage probability sampling technique was used to identify and survey U.S. individuals aged 12 and older. Details on the sampling and data collection procedures are described in further detail in other publications (Substance Abuse and Mental Health Administration [SAMHSA], 2012).

A total of 68,309 computer-assisted interviews were completed for the 2012 survey. The current study restricted analyses to adolescents (12-17 years of age) identified as being in need of treatment for alcohol or illicit drug use in the past year (N = 1,171). "Needing treatment" was a variable with a value of one (Yes = 1, No = 0) if a respondent answered "yes" to any of the following: 1) dependent on any illicit drug/alcohol in past year; 2) abused illicit drugs/alcohol in past year; or 3) received treatment for illicit drug/alcohol use at a specialty facility in past year.

Outcome Variables

The outcome of interest is the latent construct of recovery capital, which is measured using 18 variables (binary = 18, continuous = 1)[3], capturing the four domains of recovery capital: financial, human, social, and community recovery capital. All binary variables were coded such that 1 = positive outcome (i.e., evidence of recovery capital) and 0 = negative outcome (i.e., lack of recovery capital). For the one continuously measured variable, overall health status, higher values indicate a higher level of capital (range 0 – 3). Overall, there was a minimal amount of missing data across the outcome variables, with the greatest amount (6%) from the variable

[2] http://www.icpsr.umich.edu/icpsrweb/SAMHDA/download

[3] Although the NSDUH employs questions with multiple response options resulting in ordered categorical and continuous outcomes (e.g., Likert style responses) the majority of these variables were recoded by the survey authors to binary variables. The NSDUH codebook instructs users to use the recoded (primarily binary) variables and thus this recommendation was followed for all analyses.

"grades in school." See Table 1 for additional information on the sample, including sample proportions or means and standard deviations for each variable entered in the model.

Financial recovery capital. Measures of *financial recovery capital* (F1-F2 in Figure 1) included two variables: health insurance coverage and family financial standing relative to the poverty level. Respondents were asked a series of questions about types of health insurance coverage (private, CAIDCHIP, etc.) and a variable measuring whether the respondent reported any form of health insurance from a combination of these responses was created. Financial standing was calculated based on an individual's poverty threshold across a number of variables (as described by SAMHSA, 2012, p. 635) and was dichotomized to 0 = at/below poverty level and 1 = above the poverty threshold.

Human recovery capital. Measures of *human recovery capital* (H1-H5 in Figure 1) included five variables. Respondents were asked (H1) "What were your grades for the last semester or grading period you completed?" where responses of D or lower indicate no recovery capital and A, B, or C indicate evidence of recovery capital. Respondents answered questions about mental health and depressive episodes and a variable indicating the occurrence of any major depressive episodes (MDE) in the past year was created (H2) where Yes MDE = 0 (lack of recovery capital) and No MDE = 1 (evidence of recovery capital). Regarding physical health, respondents were asked: (H3) "Would you say your health in general is excellent, very good, good, fair, or poor?" Overall health status was available on a continuum with fair/poor = 0, good = 1, very good = 2, and excellent = 3. Finally, respondents were asked to rate the following statements: (H4) "Your religious beliefs are a very important part of your life" and (H5) "Your religious beliefs influence how you make decisions in your life." Responses to these questions were used as a proxy of religious (vs. secular) belief orientation with religious belief orientation suggestive of human recovery capital: strongly disagree/disagree = 0 indicates a lack of recovery capital and strongly agree/agree = 1 indicates evidence of recovery capital.

Social recovery capital. Measures of *social recovery capital* (S1-S8 in Figure 1) included eight variables. Two variables measured adolescent-reports of parents' attitudes toward drug use: (S1) "How do you think your parents would feel about you using marijuana or hashish once a month or more?" and (S2) "How do you think your parents would feel about you having one or two drinks of an alcohol beverage nearly every day?" Response categories indicate that neither approve or disapprove/somewhat disapprove = 0 (no recovery capital) and strongly disapprove = 1 (evidence of recovery capital). Participants were also asked to rate their close friends' perceptions: (S3) "How do you think your close friends would feel about you using marijuana or hashish once a month or more?" and (S4) "How do you think your close friends would feel about you having one or two drinks of an alcohol beverage nearly every day?" Response categories indicate that neither approve or disapprove = 0 (no recovery capital) and somewhat/strongly disapprove = 1 (evidence of recovery capital). Respondents were also asked (S5) if they had someone to talk to if they had a serious problem (0 = no, 1 = yes).

As a proxy for capturing involvement in positive social situations leading to potential social recovery capital, an additional three items were included: (S6) In the past 12 months, did respondents attend any type of school for any time (school attendance = 1, evidence of recovery capital). Respondents were also asked (S7) about their extracurricular activities through a number of questions addressing school-, community-, and faith-based activities. The activities variable combines all such opportunities and categorizes them between 1/none = 0 (no recovery capital) and two/more = 1 (evidence of recovery capital) activities. Finally, participants were asked: (S8) "During the past 12 months, how many times did you attend religious services?

Please do not include special occasions such as weddings, funerals, or other special events in your answers." A response of 0 indicates attending 25 or less religious services (no recovery capital) and a response of 1 indicates attending 25 or more times (evidence of recovery capital).

Community recovery capital. *Community recovery capital* (C1- C4 in Figure 1) was comprised of four variables. Respondents were asked whether they had participated in a program or meeting, such as Alcoholics Anonymous or Alateen to help with drug or alcohol use by themselves or by a family member (C1). This variable was thus used as a proxy for an indication of local community recovery capital. Three additional variables were used to measure aspects of cultural recovery capital within community capital. Respondents were asked "How many of the students in your grade at school would you say" (C2) "use marijuana or hashish?" or (C3) "drink alcohol?" or (C4) "get drunk weekly?" In this analysis, most of them/all of them = 0, indicates no recovery capital and none of them/a few of them = 1, indicates evidence of recovery capital.

Predictors

Four proposed predictors of the likelihood of being in a certain recovery capital class were: sex, race/ethnicity, age, and whether or not treatment had previously been received. In the original dataset the race/ethnicity variable was categorized into seven unique responses, but given that over half of the sample was White (57%, N = 671), to aid in estimation and interpretation, in this analysis the race/ethnicity variable was dichotomized: White = 1, non-White = 0. Age was included as a continuous predictor because it had a skewed distribution towards older adolescents in the subsample as well as a significant relationship with previous treatment receipt (44% of 17 year olds compared to 0% of 12-year olds reported having ever received treatment: $\chi^2_{(1)}$= 6.18, p = 0.013). The age variable was standardized prior to inclusion in the model to aid in interpretation (as recommended in Collins & Lanza, 2010). The final predictor is whether or not the adolescent reported ever receiving treatment in his/her lifetime (y = 1; N = 191). None of the predictor variables were missing data in this sample.

Analysis

Given that this model is theoretically based, yet exploratory, 1-, 2-, 3-, 4-, and 5-class conditional models were fit using Mplus, version 7.4 (Muthén & Muthén, 2012) to explore whether (1) there are different classes of recovery capital among this sample and if so, whether (2) certain key characteristics predict class membership. Once the number of classes was chosen, the three-step automated approach to including covariates was used (Asparouhov & Muthén, 2014; Vermunt, 2010).

Due to the complex sampling procedures used in the NSDUH (Asparouhov, 2005; Vermunt & Magidson, 2007), person-level sampling weights were incorporated in all analyses (SAMHSA, 2012). The models were estimated with maximum likelihood estimation using the expectation-maximization algorithm that includes the design-based adjustment for unequal selection probabilities (Dempster, Laird, & Rubin, 1977) and any missing data were handled using full information maximum likelihood. To avoid reaching a local maximum instead of the global maximum for the model (Titterington, Smith, & Makov, 1985), for each set of 1000 sets of random starting values, the 200 highest log-likelihood values from these possibilities were iterated until convergence was achieved.

The Bayesian information criterion (BIC; Schwartz, 1978) and Akaike information criterion (AIC; Akaike, 1987) fit indices were used to choose the best fitting model. The BIC and AIC fit statistics are based on model fit and parsimony and assess the relative fit of a series of nested models between each other: among nested models, the model with the minimum BIC or AIC value is considered a better-fitting model as it represents the optimum tradeoff between these two criteria (Collins & Lanza, 2010). The entropy statistic was used to assess the degree of class separation. The entropy statistic ranges from 0 to 1, with values closer to 1 indicating a greater degree of separation between classes and therefore greater classification accuracy.

Results

Sample Characteristics

Within this population sample of 12-17 year olds in need of substance use treatment, the average age was 15.8 (SD = 1.3) and 49% were males. Over half of the sample was White (57%, N = 671), 23% (N = 265) were Hispanic, 9% (N = 110) were Black, and the remaining 11% (N = 125) identified as other races. The majority of the sample had some form of health insurance (94%) and had family incomes above the poverty line (74%). Only 16% (N = 187) had ever received treatment for substance use. See Table 1 for additional information on the sample, including sample means and standard deviations for each variable included in the model.

Model Selection

All models converged and the BIC and AIC values supported the five-class model, as this model had the lowest BIC and AIC values of all competing models (23215.69 and 22709.12, respectively). See Table 2 for models of each tested class and their relative fit indices. As will be discussed in further detail below, the five-class model was also conceptually meaningful. This is especially noticeable when compared to the four-class model. The four-class model had classes of almost equal proportions (ranging from 17%-32%) with each class producing similar variations and only minor qualitative differences across the four domains of recovery capital: no class stood out as having substantially low or high recovery capital. In contrast, the five-class model identified five conceptually distinct classes, including a much smaller class (2.4%) that reflects a high-risk profile of recovery capital. Thus, a five-class model of recovery capital was chosen as the final and best-fitting model. Additionally, the entropy statistic was highest in the five-class model (0.821), indicating there was a great degree of separation between classes in this model.

Class Profiles

Overall, the five-class model had good categorization into class accuracy, with average latent class probabilities ranging from .85 to .99. In the five-class model, there was clearly a (1) *Religious, resource-poor* recovery capital class (class 3), (2) *Resource-wealthy* recovery capital class (class 1), (3) *Strong social, weak community* recovery capital class (class 2), (4) *Strong community* recovery capital class (class 4), and (5) *Secular, weak community* recovery capital class (class 5). Classes were assigned these labels based on the item-probability features that distinguished them from each other. Figure 2, a graph of the item probability endorsement for each class, is a visual representation of these differences. In this figure when an item probability is close to 1.0 for a specific item (e.g., "Religious beliefs influence life decisions", class 3), conditional on class, the response to this item can be determined with near certainty to be endorsed. Alternatively, when an item probability is close to 0 for an item (e.g., Health Insurance, class 3), we can determine with near-certainty that this item will not be endorsed (again, conditional on class).

Within the five-class model, the smallest class was the *religious, resource-poor* recovery capital class (2.4%, class three). Although it was a very small class compared to the rest of the sample, it was retained because it identified a potentially vulnerable class: that is, adolescents in this class were the least likely to positively endorse items in each of the four recovery capital domains, indicating a substantial lack of recovery capital, despite strong human capital in the form of religiosity. This class had the lowest mean score for overall health (0.95, SE = 0.31), over half a standard deviation lower than the class with the closest score (the strong community class, M = 1.47). In addition they were more likely to report financial instability and a lack of health insurance (low financial capital), increased perception of other students' substance use (low community capital), to have friends that were not likely to strongly disapprove of marijuana use or drinking, and to be less involved in youth activities, including religious services (low social capital). The other student and close friend items indicate that their views around the norms of substance use may align with viewing risky substance use as more normative and pervasive than other adolescents view such behavior. This finding suggests an increased risk for engaging in substance use as youth in this class may view substance use as part of typical peer behavior. This class was, however, the only class where every member positively endorsed parental views that strongly disapproved of substance use, indicating some social recovery capital from their families. They were also the only class where every member positively endorsed religious beliefs being important and influencing their lives suggesting some human capital in the form of spiritual beliefs.

In comparison, adolescents in the *resource-wealthy* recovery capital class (17.10%, class one) were the most likely to endorse almost every item across the four recovery capital domains indicating high recovery capital. This class had significantly higher odds of being above the poverty level than the religious, resource-poor class (OR = 21.04, 95% CI [3.12, 141.85]). This class also positively endorsed items across the social and human recovery capital domains, including school attendance, activity participation, parental and peer disapproval of substance use, good grades, no major depressive episode in the past year, and religious beliefs being important and influential. Although the likelihood of regular church attendance was low across all classes, adolescents in the resource-wealthy class were significantly more likely to endorse regular church attendance in the last year compared to all other classes: versus the strong social, weak community class (OR = 2.41, 95% CI [1.13, 5.17]), the religious, resource-poor class (OR = 14.80, 95% CI [1.90, 201.06]), the secular, weak community class (OR = 16.70, 95% CI [6.83, 40.82]), and the strong community class (no one in this class endorsed this item). They also had one of the highest overall health averages (1.81, SE = 0.12). Adolescents in this class had slightly reduced community recovery capital in that they were not likely to report that they attended a self-help group for themselves or another family member's substance use; however, this item was not positively endorsed in most classes. In addition, just under half were likely to report thinking that other students in their grade used marijuana/hashish monthly and drank alcohol. Yet, three other recovery capital classes were at significantly greater odds than this resource-wealthy class to endorse reporting that their peers regularly used marijuana: the strong social, weak community class was at 6.99 greater odds (95% CI [2.53, 19.32]) and the secular, weak community class was at 10.65 greater odds (95% CI [4.45, 25.51]) than the resource-wealthy class to endorse peer regular use of marijuana, and everyone in the religious, resource-poor class endorsed this item. Adolescents in the resource-wealthy class were also not likely to endorse the statement that other students in their grade get drunk weekly, so it appears this group of youth do not consider getting drunk regularly to be normative behavior.

The largest class was the *strong social, weak community* recovery capital class (32.39%, class two). Although adolescents in this class had low community capital, they were otherwise considered fairly high in the other recovery capital domains, especially social recovery capital. Similar to the resource-wealthy class, this class was not likely to endorse attending a self-help group for themselves or another family member's substance use. Comparing youth in this class to youth in the resource-wealthy class, we see that those in the resource-wealthy class were at

6.99 higher odds to report that other students did not use marijuana (95% CI [2.53, 19.32]), at 26.65 higher odds to report that other students did not drink (95% CI [5.34, 133.00]), and at 30.81 higher odds to report that other students did not get drunk (95% CI [3.77, 251.52]). Compared to the strong community and the secular, weak community classes, this strong social, weak community class had higher proportions to endorse thinking that other students in their grade used marijuana/hashish monthly and drank alcohol, and about half of the class was also likely to endorse the statement that other students in their grade get drunk weekly (although these differences were not statistically significant). Regarding social capital, youth in the strong social, weak community class were significantly more likely than the strong community class to endorse parental disapproval of marijuana use (OR = 2.35, 95% CI [1.25, 4.38]) and significantly more likely than the secular, weak community class to report peer disapproval of marijuana use (OR = 1.84, 95% CI [1.03, 3.28]). They were also more likely than the strong community class to report having someone to talk to (OR = 2.60, 95% CI [1.19, 5.67]). They were at significantly higher odds than the religious, resource-poor class and the strong community class to report school attendance (OR = 4.24, 95% CI [1.36, 113.24] and OR = 6.63, 95% CI [1.77, 24.85], respectively). Finally, they were significantly more likely to report participating in youth activities than the religious, resource-poor class (OR = 11.78, 95% CI [1.85, 74.91]), than the strong community class (OR = 3.06, 95% CI [1.73, 5.41]), and than the secular, weak community class (OR = 2.36, 95% CI [1.39, 4.03]).

Adolescents in the *strong community* recovery capital class (18.08%, class four) had consistently high levels of community recovery capital; adolescents in this class were significantly more likely than the secular, weak community class to positively endorse participating in a substance use program (OR = 3.09, 95% CI [1.41, 6.78]) and positively endorsed other community capital items. Adolescents in this class had a mix of financial capital: they were not likely to be below the poverty line, but had smaller proportions than the resource-wealthy class, the strong social, weak community class, and the secular, weak community class to endorse having health insurance, although the difference was non-significant. In addition, although adolescents in this strong community class positively endorsed most human capital items, the average of their overall health rating was 1.47 (SE = 0.10): this rating is above the health average of the religious, resource-poor class, but lower than the other three classes. Adolescents in this class were also less likely than the resource-wealthy class and the strong social, weak community class to endorse the religious belief items. They were less likely than the resource-wealthy class and strong social, weak community class to endorse some items on the social capital domain including parents' and close friends' disapproving views of their use of substances; however, they were significantly more likely than the secular, weak community class to report that their friends disapproved of marijuana use (OR = 2.01, 95% CI [1.12, 3.60]). Finally, adolescents in this strong community class were significantly less likely than the resource-wealthy class and strong social, weak community class to endorse engaging in two or more activities and in religious service attendance.

Adolescents in the *secular, weak community* recovery capital class (30.03%, class five) had low levels of community capital and low levels of religiosity. Adolescents in this class were likely to be above the poverty line and to have health insurance, indicating high financial recovery capital. Similar to other recovery capital classes, this class was not likely to endorse attending a self-help group for themselves or another family member's substance use. Adolescents in this class were at significantly greater odds than the resource-wealthy and the strong community classes to endorse thinking that other students in their grade used marijuana/hashish monthly (OR = 10.46, 95% CI [4.45, 25.51] and OR = 14.70, 95% CI [7.24, 25.90], respectively). The entire class endorsed thinking that other students in their grade drank and this secular, weak community class had significantly higher odds when compared to the resource-wealthy and the strong community classes to endorse the statement that other students in their grade get drunk weekly (OR = 28.10, 95% CI [3.21, 246.13] and OR = 81.86, 95% CI [8.54, 784.25], respectively). This class also had a mix of human capital. They were likely to endorse having good grades and had the highest average score for overall health ratings (M = 1.88, SE = 0.09). These adolescents were not likely, however, to endorse either of the religion items and were the class most likely to endorse experiencing a major depressive episode in the last year, although the estimated proportion was less than half and the odds of endorsement compared to other classes was not significant. Finally, regarding social recovery capital, adolescents in this class endorsed having someone to talk to, attending school, and participation in youth activities, yet were not likely to endorse attending religious services. They were also less likely than other classes to endorse that their parents and friends disapproved of marijuana use and drinking. For example, the resource-wealthy class, the strong social, weak community class, and the strong community class all had significantly greater odds of endorsing that their friends disapproved of marijuana use than this secular, weak community class (OR = 6.70, 95% CI [3.61, 12.44], OR = 1.84, 95% CI [1.03, 3.28], and OR = 2.01, 95% CI [1.12, 3.60], respectively). The resource-wealthy class was at significantly greater odds of endorsing that their friends disapproved of drinking (OR = 3.97, 95% CI [2.03, 7.80]), and that that their parents disapproved of drinking (OR = 58.82, 95% CI [4.69, 737.96]) and marijuana use (OR = 42.38, 95% CI [2.43, 739.91]) compared to this secular, weak community class.

Predictors of Class Membership

Table 4 presents the results of multinomial regressions for each predictor, first comparing the religious, resource-poor class to every other class, then comparing the resource-wealthy class to the remaining three classes. Finally, results are presented for the few class comparisons not included in the above comparisons. Given their relative position at both ends of a recovery capital spectrum, the two primary class comparisons, between the religious, resource-poor class and other classes and between the resource-wealthy class and other classes are discussed in the text below. Additionally, as treatment is important to the recovery process, the results of the analysis related to treatment receipt and recovery capital class membership will also be discussed.

Religious, resource-poor class. Holding other predictors constant, adolescents in the religious, resource-poor class were more likely to belong to a racial/ethnic minority group than individuals in all other classes ranging from a factor of OR= 23.27 (resource-wealthy class) to 47.80 (secular, weak community class). Adolescents in this religious, resource-poor class were also significantly more likely to be older than individuals in the resource-wealthy class (OR = 0.17, 95% CI [0.04, 0.66]) and the strong community class (OR = 0.16, 95% CI [0.04, 0.63]).

Resource-wealthy class. Adolescents in the resource-wealthy class had significantly higher odds of being female than those in the strong community class (OR= 2.67, 95% CI [1.34, 5.32]). Adolescents in the resource-wealthy class also had significantly higher odds of being younger than individuals in three of the four other classes (i.e., the religious, resource-poor class, the strong social, weak community class, and the secular, weak community class). Finally, adolescents in the resource-wealthy class had significantly lower odds of being White when compared to those in the secular, weak community class (OR= 2.05, 95% CI [1.11, 3.79]).

Treatment receipt and other recovery capital classes. Treatment receipt was only a significant predictor in the relationship between being a member of the strong community class and the secular, weak community class. Adolescents in the strong community class had higher odds of having received treatment by a factor of OR = 0.48 (95% CI [0.26, 0.87]) compared to those in the secular, weak community class. In addition, adolescents in the strong community class had higher odds of being male (OR = 0.44, 95% CI [0.20, 0.96]), and being younger than those in the secular, weak community class (OR = 2.31, 95% CI [1.72, 3.10]). Thus, in this sample, younger males were more likely to receive treatment and to be members of the strong community class.

Discussion

The purpose of this study was to identify whether, and if so, what were the distinct patterns of recovery capital classes in an adolescent population identified as needing substance use treatment. Latent class analysis revealed five distinct and meaningful classes among this sample that varied within a range of religious, resource-poor to resource-wealthy recovery capital. These results demonstrated that a linear model of recovery capital is likely too simplistic, particularly among adolescents. That is, although the analysis identified both a highly vulnerable and a highly resourced recovery capital class, all youth had some areas of capital resulting in categorically unique classes of recovery capital, which included a strong social, weak community class, a strong community class, and a secular, weak community class. There was no absolutely lacking or robust class: even adolescents in the religious, resource-poor class had capital in the form of parental disapproval of substance use and in personal religious beliefs, and adolescents in the resource-wealthy class lacked some community recovery capital in the form of perceptions of peer use of substances. These results suggest multiple areas for future intervention within a recovery capital framework: by (1) addressing areas of typical resource gaps for all adolescents, (2) filling specific gaps in resources for vulnerable populations, and (3) identifying and building upon existing recovery capital.

Using the recovery capital model, we see that overall for this population of adolescents in need of substance use treatment, risk factors were evident at the microsystem level. For example, many of the included adolescents perceived that peers in their surrounding environment were regularly using substances, including binge drinking, which demonstrates low community recovery capital in the form of low cultural recovery capital. Previous research suggests that adolescents often overestimate peer use of substances (Lewis & Mobley, 2010; Neighbors, Dillard, Lewis, Bergstrom, & Neil, 2006); thus, correcting potential misperceptions of local norms and changing individual youth behavior in spite of these perceived norms, remains an area for treatment and aftercare intervention. A large majority of the sample also reported that close friends did not disapprove of marijuana use or drinking, indicating barriers to social recovery capital through social networks comprised of members who may not support their friends' sobriety. This finding indicates that practitioners must continue to address adolescent social networks by helping youth to locate sources of sober peer supports in their communities so they can build new social recovery capital via community recovery capital supports. This could involve encouraging youth involvement in local positive youth activities based on their personal skills and interests, but could also involve more targeted support to encourage and enable youth to engage in sober peer environments such as alternative peer groups (Collier, Hilliker, & Onwuegbuzie, 2014) or recovery high schools (Finch & Frieden, 2014; Finch, Moberg, & Krupp, 2014). These youth-specific supports may be necessary for this population given the findings that community substance use program attendance, such as at a 12-Step program, was markedly low in the sample. Although youth 12-Step program attendance has been shown to positively correlate with a reduction in substance use after treatment (Hennessy & Fisher, 2015), 12-Step meetings may only typically be attended by youth who have had previous treatment and feel comfortable with this particular program model. Thus, in this sample of youth who had largely not had previous treatment, other sources of youth-specific sober supports could be beneficial.

Researchers and practitioners could also use the recovery capital model to identify areas of vulnerability for particular groups of adolescents with few recovery capital resources and seek to fill these gaps by building additional resource supports in these areas (Cloud & Granfield, 2004; White & Cloud, 2008) and engaging in further research on the best ways to do so. For example, youth in the religious, resource-poor class demonstrated the fewest recovery capital resources and were also more likely to be members of a racial/ethnic minority group and somewhat older than other classes. This finding suggests two areas for building additional resource support, especially at the microsystem level for this vulnerable population: there is a need to find appropriate resources for treatment-seeking or treatment-completing racial/ethnic minority adolescents and for older adolescents, both of whom may have fewer recovery capital resources. Older adolescents may have had a substance use problem for longer than younger adolescents, thereby further eroding their recovery capital at different ecological levels. For example, they may have missed large amounts of academic education or be delayed in life skills development due to prolonged use of substances and enrollment in treatment programs (reduced human recovery capital) or distanced previously recovery-supportive friends and family by their substance using behaviors or negative consequences from these behaviors (reduced social recovery capital). Thus, supports that may directly target getting these older adolescents up-tospeed on developmentally appropriate academic and life skills progress and that may rebuild their social recovery capital with recovery-supportive friends and family are necessary.

In support of culturally-appropriate outreach for substance involved youth, recent meta-analytic findings have found evidence that culturally-adapted substance use treatment programs are more effective in reducing substance use than generic programming for racial/ethnic minority youth, although the evidence for each specific ethnic/racial minority group in the United States is limited (Steinka-Fry, Tanner-Smith, Dakof, & Henderson, 2017). Additionally, research on access to treatment and treatment completion have demonstrated a disparity between White and racial/ethnic minority individuals and service usage, with minority individuals reporting less access to healthcare (Alegria, Carson, Goncalves, & Keefe, 2011; Guerrero et al., 2013). Given that the religious, resource-poor recovery capital class was also likely to be below the poverty level and without health insurance, this implies that practitioners should consider potential individual-level barriers to treatment access that could be addressed via larger structural changes at the microsystem or mesosystem, such as providing transportation, identifying and addressing comorbid health issues through better linkages with the primary healthcare system, or reducing cultural and language barriers to engage youth and their families in treatment services.

However, adolescents in this religious, resource-poor class also reported some capital, and therefore, part of their existing human and social recovery capital at the individual-level and microsystem could be garnered to enhance capital. For example, adolescents in this class were likely to endorse parental views that strongly disapproved of substance use and all endorsed having religious beliefs suggesting some potential for support through religion and spirituality. This finding suggests that further involving parents in treatment and aftercare supports may be one way to bolster this class' recovery capital, an implication also supported by empirical research (Tanner-Smith et al., 2013). Introducing spiritual recovery supports, such as through faith-based institutions, may also be relevant for these adolescents.

Alternatively, individuals in the resource-wealthy class were the most likely to endorse almost every item across the four recovery capital domains indicating they had substantial recovery capital in each domain. Simply because youth have many recovery capital resources, however, does not guarantee positive recovery outcomes; thus, their existing resources should be identified and built upon as assets by practitioners in the recovery support process (Cloud & Granfield, 2004). This group was also more likely to be female than the strong community class and somewhat younger than most of the other classes, suggesting that various recovery capital profiles of adolescents may require different developmentally- and gender-appropriate services.

Limitations

This study used a relatively new conceptualization of recovery capital among adolescents using available cross-sectional survey data. It is possible some key factors of each form of recovery capital were omitted due to limitations with the data or with this conceptualization of recovery capital. In addition, this analysis used primarily binary indicators of recovery capital, and future research should utilize variables on a continuum as these may produce more sensitive categorizations of recovery capital classes. Additionally, religion and spirituality are popular with adults in recovery, yet less so with youth, although these components are still often included in the 12-Step treatment model in which youth participate (Kelly, Hoeppner, Stout, & Pagano, 2012). Thus, whether, and if so, how, religious beliefs among adolescents should be studied merits further attention: the items used in this analysis were relatively simplistic for this purpose. Similarly, the items used to measure community recovery capital were somewhat limited to individual-level behaviors and beliefs and future research with data that allows for identification of actual recovery resources available in communities (e.g., number and proximity of 12-Step meetings a week) would prove useful for a more accurate measurement of community recovery capital.

Because recovery is a process, the cross-sectional nature of this study's data is also a limitation in considering recovery capital as part of the overall recovery process. For example, the relationship between treatment receipt and recovery capital is still unclear and needs research attention: previous treatment could have either eroded some forms of recovery capital (e.g., ties to local community) or bolstered other forms (e.g., ties to sober and supportive community). Additionally, receipt of treatment could be seen as a resource indicative of higher capital (e.g., financial resources which allow better treatment) but could also be due to more severe substance use problems leading to a higher likelihood of receiving treatment. Future research with recovering adolescents should model recovery capital and the recovery process longitudinally, assess how episodes of treatment interact with recovery capital, seek to determine how recovery capital built/diminished at one ecological level influences recovery capital at another ecological level, and explore distal outcomes such as return to substance use.

Conclusions

Despite these limitations, the study is strengthened by a large, nationally-representative sample of adolescents in the United States, is among the first to quantitatively explore recovery capital among adolescents, and demonstrates five distinct classes of recovery capital. These findings suggest the need for future research on the relationship between adolescent recovery capital and the recovery process, which should further delineate recovery capital factors and interactions between the different recovery capital domains. This research should seek a better understanding of factors at the microsystem and mesosystem levels, e.g., via further study of community recovery capital factors, and how they interact with individual-level factors, including asking broader questions about knowledge of available community recovery resources and measuring engagement in these resources in addition to attendance. Attention to previous treatment and aftercare receipt and severity of substance use history should also be taken into account when studying adolescent recovery capital and its relationship to the recovery process. Overall, recovery capital is a useful model to understanding the many intersecting adolescent recovery factors that facilitate progress and potential areas of vulnerability in recovery. Further study and use of the recovery capital model in practice will provide additional information to understand the nuances of adolescent recovery and may improve tailoring of treatment and aftercare supports for adolescents, especially for youth with particular vulnerabilities arising from structural barriers to care.

References

Aike, H. (1987). Factor analysis and AIC. *Psychometrika, 53*, 317-332. http://dx.doi.org/10.1007/BF02294359

Alegria, M., Carson, N. J., Goncalves, M., & Keefe, K. (2011). Disparities in treatment for substance use disorders and co-occurring disorders for ethnic/racial minority youth. *Journal of the American Academy of Child and Adolescent Psychiatry*, *50*(1), 22-31. doi:10.1016/j.jaac.2010.10.005

Allen, M., Donohue, W. A., Griffin, A., Ryan, D., & Turner, M. M. M. (2003). Comparing the influence of parents and peers on the choice to use drugs: A meta-analytic summary of the literature. *Criminal Justice and Behavior*, *30*(2), 163-186. doi:10.1177/0093854802251002

Anderson, K. G., Ramo, D. E., Cummins, K. M., & Brown, S. A. (2010). Alcohol and drug involvement after adolescent treatment and functioning during emerging adulthood. *Drug and Alcohol Dependence*, *107*(2-3), 171-81. doi:10.1016/j.drugalcdep.2009.10.005

Asparouhov, T. (2005). Sampling weights in latent variable modeling. *Structural Equation Modeling*, 12(3), 411-434.

Asparouhov, T., & Muthén, B. (2014). Auxiliary variables in mixture modeling: Three-step approaches using M plus. *Structural Equation Modeling: A Multidisciplinary Journal*, *21*(3), 329-341.

Becker, S. J., Stein, G. L., Curry, J. F., & Hersh, J. (2012). Ethnic differences among substance-abusing adolescents in a treatment dissemination project. *Journal of Substance Abuse Treatment*, *42*(3), 328-36. doi:10.1016/j.jsat.2011.08.007

Best, D., Gow, J., Knox, T., Taylor, A., Groshkova, T., & White, W. (2012). Mapping the recovery stories of drinkers and drug users in Glasgow: Quality of life and its associations with measures of recovery capital. *Drug and Alcohol Review, 31*(3), 334-41. doi:10.1111/j.1465-3362.2011.00321.x

Best, D., Gow, J., Taylor, A., Knox, A., & White, W. (2011). Recovery from heroin or alcohol dependence: A qualitative account of the recovery experience in Glasgow. *Journal of Drug Issues, 41*(3).

Best, D., Honor, S., Karpusheff, J., Loudon, L., Hall, R., Groshkova, T., & White, W. (2012). Well-Being and recovery functioning among substance users engaged in posttreatment recovery support groups. *Alcoholism Treatment Quarterly*, *30*(4), 397-406.

Betty Ford Institute Consensus Panel. (2007). What is recovery? A working definition from the Betty Ford Institute. *Journal of Substance Abuse Treatment*, *33*(3), 221 - 228. doi:10.1016/j.jsat.2007.06.001

Bronfenbrenner, U. (1977). Towards an experimental ecology of human development. *American Psychologist,* 518-531.

Bronfenbrenner, U. (1994). Ecological models of human development. In *International Encyclopedia of Education, Vol. 3, 2nd Ed.* Oxford: Elsevier. Reprinted in: M. Gauvain & M. Cole (Eds.), *Readings on the Development of Children, 2nd Ed.* (1993, pp. 37-43). NY: Freeman.

Brown, S. A., D'Amico, E. J., McCarthy, D. M., & Tapert, S. F. (2001). Four-year outcomes from adolescent alcohol and drug treatment. *Journal of Studies on Alcohol*, *62*(3), 381-388.

Brown, S. A., Myers, M. G., Mott, M. A., & Vik, P. W. (1994). Correlates of success following treatment for adolescent substance abuse. *Applied & Preventive Psychology, 3,* 61-73.

Burns, J., & Marks, D. (2013). Can recovery capital predict addiction problem severity? *Alcoholism Treatment Quarterly, 31*(3), 303-320.

Chi, F. W., Kaskutas, L. A., Sterling, S., Campbell, C. I., & Weisner, C. (2009). Twelve-step affiliation and 3-year substance use outcomes among adolescents: Social support and religious service attendance as potential mediators. *Addiction, 104*(6), 927-939. doi:doi:10.1111/j.1360-0443.2009.02524.x

Cloud, W., & Granfield, R. (2004). A life course perspective on exiting addiction: The relevance of recovery capital in treatment. *NAD Publication*, *44*, 185-202.

Cloud, W., & Granfield, R. (2008). Conceptualizing recovery capital: Expansion of a theoretical construct. *Substance Use & Misuse*, *43*(12-13), 1971-86. doi:10.1080/10826080802289762

Collier, C., Hilliker, R., & Onwuegbuzie, A. (2014). Alternative peer group: A model for youth recovery. *Journal of Groups in Addiction & Recovery*, *9*(1), 40-53. http://dx.doi.org/10.1080/1556035X.2013.836899

Collins, L. M., & Lanza, S. T. (2010). *Latent Class and Latent Transition Analysis*. John Wiley & Sons, Inc.

Cornelius, J. R., Maisto, S. A., Pollock, N. K., Martin, C. S., Salloum, I. M., Lynch, K. G., & Clark, D. B. (2003). Rapid relapse generally follows treatment for substance use disorders among adolescents. *Addictive Behaviors*, *28*(2), 381-6.

Deas, D., Riggs, P., Langenbucher, J., Goldman, M., & Brown, S. (2000). Adolescents are not adults: Developmental considerations in alcohol users. *Alcoholism: Clinical and Experimental Research*, *24*(2), 232-237.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM Algorithm. *Journal of the Royal Statistical Society. Series B (Methodological), 39*, 1-38.

Finch, A. J., & Frieden, G. (2014). The ecological and developmental role of recovery high schools. *Peabody Journal of Education*, *89*(2), 271-287. doi:10.1080/0161956X.2014.897106

Finch, A. J., Moberg, D. P., & Krupp, A. L. (2014). Continuing care in high schools: A descriptive study of recovery high school programs. *Journal of Child & Adolescent Substance Abuse*, *23*(2), 116-129. doi:10.1080/1067828X.2012.751269

Galea, S., Nandi, A., & Vlahov, D. (2004). The social epidemiology of substance use. *Epidemiologic Reviews*, *26*(1), 36-52. doi:10.1093/epirev/mxh007

Godley, M. D., Kahn, J. H., Dennis, M. L., Godley, S. H., & Funk, R. R. (2005). The stability and impact of environmental factors on substance use and problems after adolescent outpatient treatment for cannabis abuse or dependence. *Psychology of Addictive Behaviors*, *19*(1), 62-70. doi:10.1037/0893-164X.19.1.62

Granfield, R., & Cloud, W. (1999). *Coming clean: Overcoming addiction without treatment.* New York: New York University Press.

Granfield, R., & Cloud, W. (2001). Social context and natural recovery: The role of social capital in the resolution of drug-associated problems. *Substance Use & Misuse*, *36*(11), 1543-1570.

Guerrero, E. G., Marsh, J. C., Duan, L., Oh, C., Perron, B., & Lee, B. (2013). Disparities in completion of substance abuse treatment between and within racial and ethnic groups. *Health Services Research*, *48,* 1450-67. doi:10.1111/1475-6773.12031

Henderson, C. E., Rowe, C. L., Dakof, G. A., Hawes, S. W., & Liddle, H. A. (2009). Parenting practices as mediators of treatment effects in an early-intervention trial of multidimensional family therapy. *The American Journal of Drug and Alcohol Abuse*, *35*(4), 220-226.

Hennessy, E. A. (2017). Recovery Capital: A systematic review of the literature. *Addiction Research and Theory*, in press.

Hennessy, E. A., & Fisher, B. W. (2015). A meta-analysis exploring the relationship between 12-Step attendance and adolescent substance use relapse. *Journal of Groups in Addiction and Recovery, 10*, 79-96.

Hoffmann, J. P., & Su, S. S. (1998). Parental substance use disorder, mediating variables and adolescent drug use: A non-recursive model. *Addiction*, *93*(9), 1351-64.

Kaminer, Y., Burleson, J. A., & Burke, R. H. (2008). Efficacy of outpatient aftercare for adolescents with alcohol use disorders: A randomized controlled study. *Journal of the American Academy of Child and Adolescent Psychiatry*, *47*(12), 1405-12. doi:10.1097/CHI.0b013e318189147c

Kelly, J. F., Hoeppner, B., Stout, R. L., & Pagano, M. (2012). Determining the relative importance of the mechanisms of behavior change within alcoholics anonymous: A multiple mediator analysis. *Addiction*, *107*(2), 289-99. doi:10.1111/j.1360-0443.2011.03593.x

Kelly, J. F., Myers, M. G., & Brown, S. A. (2000). A multivariate process model of adolescent 12-step attendance and substance use outcome following inpatient treatment. *Psychology of Addictive Behaviors*, *14*(4), 376-389.

Kelly, J. F., Myers, M. G., & Brown, S. A. (2002). Do adolescents affiliate with 12-step groups? A multivariate process model of effects. *Journal of Studies on Alcohol*, *63*(3), 293-304.

Kelly, J. F., Pagano, M. E., Stout, R. L., & Johnson, S. M. (2011). Influence of religiosity on 12-step participation and treatment response among substance-dependent adolescents. *Journal of Studies on Alcohol and Drugs*, *72*(6), 1000-1011.

Kendler, K. S., Gardner, C. O., Hickman, M., Heron, J., Macleod, J., Lewis, G., & Dick, D. M. (2014). Socioeconomic status and alcohol-related behaviors in mid-to-late adolescence in the Avon Longitudinal study of parents and children. *Journal of Studies on Alcohol and Drugs*, *75*(4), 541-545.

Lamont, A., Woodlief, D., & Malone, P. (2014). Predicting high-risk versus higher-risk substance use during late adolescence from early adolescent risk factors using latent class analysis. *Addiction Research and Theory*, *22*(1), 78-89.

Lanham, C. C., & Tirado, J. A. (2011). Lessons in sobriety: An exploratory study of graduate outcomes at a recovery high school. *Journal of Groups in Addiction & Recovery*, *6*(3), 245-263. doi:10.1080/1556035X.2011.597197

Lewis, T. F., & Mobley, A. K. (2010). Substance abuse and dependency risk: The role of peer perceptions, marijuana involvement, and attitudes toward substance use among college students. *Journal of Drug Education*, *40*(3), 299-314.

Mason, M. J., Malott, K., & Knoper, T. (2009). Urban adolescents' reflections on brief substance use treatment, social networks, and self-narratives. *Addiction Research & Theory*, *17*(5), 453-468.

Mason, M. J., Mennis, J., Linker, J., Bares, C., & Zaharakis, N. (2014). Peer attitudes effects on adolescent substance use: The moderating role of race and gender. *Prevention Science*, *15*(1), 56-64. doi:10.1007/s11121-012-0353-7

Mattick, Wadolowski, Aiken, Kypri, Slade, Hutchinson, . . . McBride. (2014, September 8). Early parental supply of alcohol: Association with drinking in mid-adolescence? (Presented at the National Drug and Alcohol Research Centre Annual Symposium). Retrieved from http://ndarc.med.unsw.edu.au/sites/default/files/newsevents/events/Parental%20supply%20presentation%20slides.pdf

Moon, D. G., Jackson, K. M., & Hecht, M. L. (2000). Family risk and resiliency factors, substance use, and the drug resistance process in adolescence. *Journal of Drug Education*, *30*(4), 373-98.

Mowbray, O., Glass, J. E., & Grinnell-Davis, C. L. (2015). Latent class analysis of alcohol treatment utilization patterns and 3-year alcohol related outcomes. *Journal of Substance Abuse Treatment*, *54*, 21-8. doi:10.1016/j.jsat.2015.01.012

Mrug, S., & McCay, R. (2013). Parental and peer disapproval of alcohol use and its relationship to adolescent drinking: Age, gender, and racial differences. *Psychology of Addictive Behaviors*, *27*(3), 604-14. doi:10.1037/a0031064

Muthén, L. K., & Muthén, B. O. (1998-2012). *Mplus User's Guide*. Seventh Edition. Los Angeles, CA: Muthén & Muthén.

Neighbors, C., Dillard, A. J., Lewis, M. A., Bergstrom, R. L., & Neil, T. A. (2006). Normative misperceptions and temporal precedence of perceived norms and drinking. *Journal of Studies on Alcohol*, *67*(2), 290-9.

Ramo, D. E., & Brown, S. A. (2008). Classes of substance abuse relapse situations: A comparison of adolescents and adults. *Psychology of Addictive Behaviors*, *22*(3), 372-9. doi:10.1037/0893-164X.22.3.372

Ramo, D. E., Prince, M. A., Roesch, S. C., & Brown, S. A. (2012). Variation in substance use relapse episodes among adolescents: A longitudinal investigation. *Journal of Substance Abuse Treatment*, *43*(1), 44-52. doi:10.1016/j.jsat.2011.10.003

Rew, L., & Wong, Y. J. (2006). A systematic review of associations among religiosity/spirituality and adolescent health attitudes and behaviors. *The Journal of Adolescent Health*, *38*(4), 433-442. doi:10.1016/j.jadohealth.2005.02.004

Ritt-Olson, A., Milam, J., Unger, J. B., Trinidad, D., Teran, L., Dent, C. W., & Sussman, S. (2004). The protective influence of spirituality and “health-as-a-value” against monthly substance use among adolescents varying in risk. *The Journal of Adolescent Health*, *34*(3), 192-199. doi:10.1016/j.jadohealth.2003.07.009

Rohde, P., Waldron, H. B., Turner, C. W., Brody, J., & Jorgensen, J. (2014). Sequenced versus coordinated treatment for adolescents with comorbid depressive and substance use disorders. *Journal of Consulting and Clinical Psychology*, *82*(2), 342-8. doi:10.1037/a003580

Schwartz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics, 6*, 461-464.

Skogens, L. & von Greiff, N. (2014). Recovery capital in the process of change-differences and similarities between groups of clients treated for alcohol or drug problems. *European Journal of Social Work, 17*(1), 58-73.

Spear, S. F., Ciesla, J. R., & Skala, S. Y. (1999). Relapse patterns among adolescents treated for chemical dependency. *Substance Use & Misuse*, *34*(13), 1795-1815.

Stanger, C., Ryan, S. R., Scherer, E. A., Norton, G. E., & Budney, A. J. (2015). Clinic- and home-based contingency management plus parent training for adolescent cannabis use disorders. *Journal of the American Academy of Child and Adolescent Psychiatry*, *54*(6), 445-453.e2. doi:10.1016/j.jaac.2015.02.00

Stevens, S. J., Estrada, B., Murphy, B. S., McKnight, K. M., & Tims, F. (2004). Gender difference in substance use, mental health, and criminal justice involvement of adolescents at treatment entry and at three, six, 12 and 30 month follow-up. *Journal of Psychoactive Drugs*, *36*(1), 13-25.

Substance Abuse and Mental Health Services Administration (SAMHSA). (2012). *Results from the 2010 National Survey on Drug Use and Health: Summary of national findings*. Rockville, MD: Substance Abuse and Mental Health Services Administration.

Tanner-Smith, E. E., Wilson, S. J., & Lipsey, M. W. (2013). The comparative effectiveness of outpatient treatment for adolescent substance abuse: A meta-analysis. *Journal of Substance Abuse Treatment*, *44*(2), 145-58. doi:10.1016/j.jsat.2012.05.006

Terrion, J. L. (2012). The experience of post-secondary education for students in recovery from addiction to drugs or alcohol: Relationships and recovery capital. *Journal of Social and Personal Relationships*, 1-21. doi:10.1177/0265407512448276

Titterington, D. M., Smith, A. F. M., & Makov, U. E. (1985). *Statistical Analysis of Finite Mixture Distributions*. John Wiley: New York.

United States Department of Health and Human Services. (2014). National Survey on Drug Use and Health, 2012. Substance Abuse and Mental Health Services Administration. Center for Behavioral Health Statistics and Quality. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor]. http://doi.org/10.3886/ICPSR34933.v2

van der Zwaluw, C. S., Otten, R., Kleinjan, M., & Engels, R. C. (2013). Different trajectories of adolescent alcohol use: Testing gene-environment interactions. *Alcoholism, Clinical and Experimental Research, 38,* 704-12. doi:10.1111/acer.12291 van Melick, M., McCartney, D., & Best, D. (2013). Ongoing recovery support and peer networks: A preliminary investigation of recovery champions and their peers. *Journal of Groups in Addiction & Recovery*, *8*(3), 185-199.

Vermunt, J. K. (2010). Latent class modeling with covariates: Two improved three-step approaches. *Political analysis*, *18*(4), 450-469.

Vermunt, J. K. & Magidson, J. (2007). Latent class analysis with sampling weights: A maximum likelihood approach. *Sociological Methods & Research, 36*, 87-111.

Wambeam, R. A., Canen, E. L., Linkenbach, J., & Otto, J. (2013). Youth misperceptions of peer substance use norms: A hidden risk factor in state and community prevention. *Prevention Science*, *15*(1), 75-84. doi:10.1007/s11121-013-0384-8

Wellman, R. J., Contreras, G. A., Dugas, E. N., O'Loughlin, E. K., & O'Loughlin, J. L. (2014). Determinants of sustained binge drinking in young adults. *Alcoholism, Clinical and Experimental Research*, *38*(5), 1409-15. doi:10.1111/acer.12365

White, W. L. (2012). *Recovery/remission from substance use disorders. An analysis of reported outcomes in 415 scientific reports, 1868-2011.* Retrieved from http://nattc.org/learn/topics/rosc/docs/RecoveryRemissionWW.pdf

White, W., & Cloud, W. (2008). Recovery capital: A primer for addictions professionals. *Counselor, 9*(5), 22-27.

Winters, K. C., Stinchfield, R., Latimer, W. W., & Lee, S. (2007). Long-term outcome of substance-dependent youth following 12-step treatment. *Journal of Substance Abuse Treatment*, *33*(1), 61-9. doi:10.1016/j.jsat.2006.12.003

Yu, J. W., Buka, S. L., Fitzmaurice, G. M., & McCormick, M. C. (2006). Treatment outcomes for substance abuse among adolescents with learning disorders. *The Journal of Behavioral Health Services & Research*, *33*(3), 275-286. doi:10.1007/s11414-006-9023-5

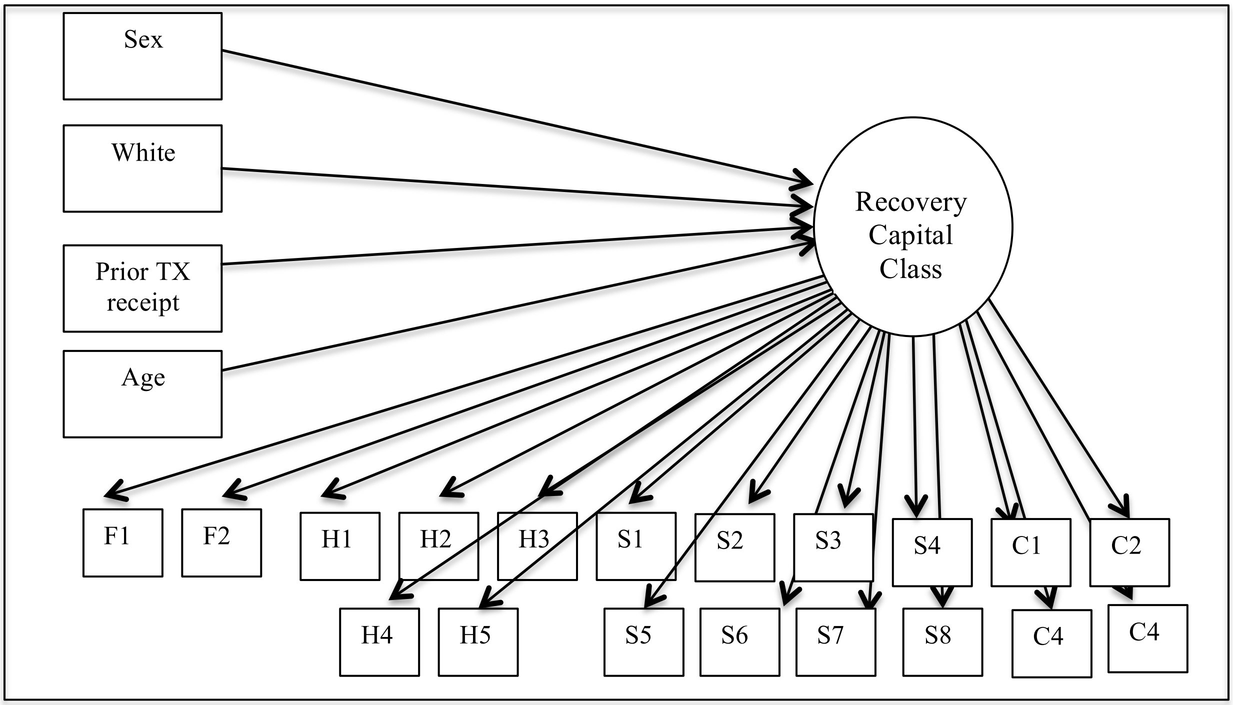

Figure 1. Recovery Capital Path Diagram
Note. For visual simplicity none of the error terms are included; however, there is measurement error associated with each item. TX = Treatment. F = *Financial Recovery Capital*. H = *Human Recovery Capital*. S = *Social Recovery Capital*. C = *Community Recovery Capital*.

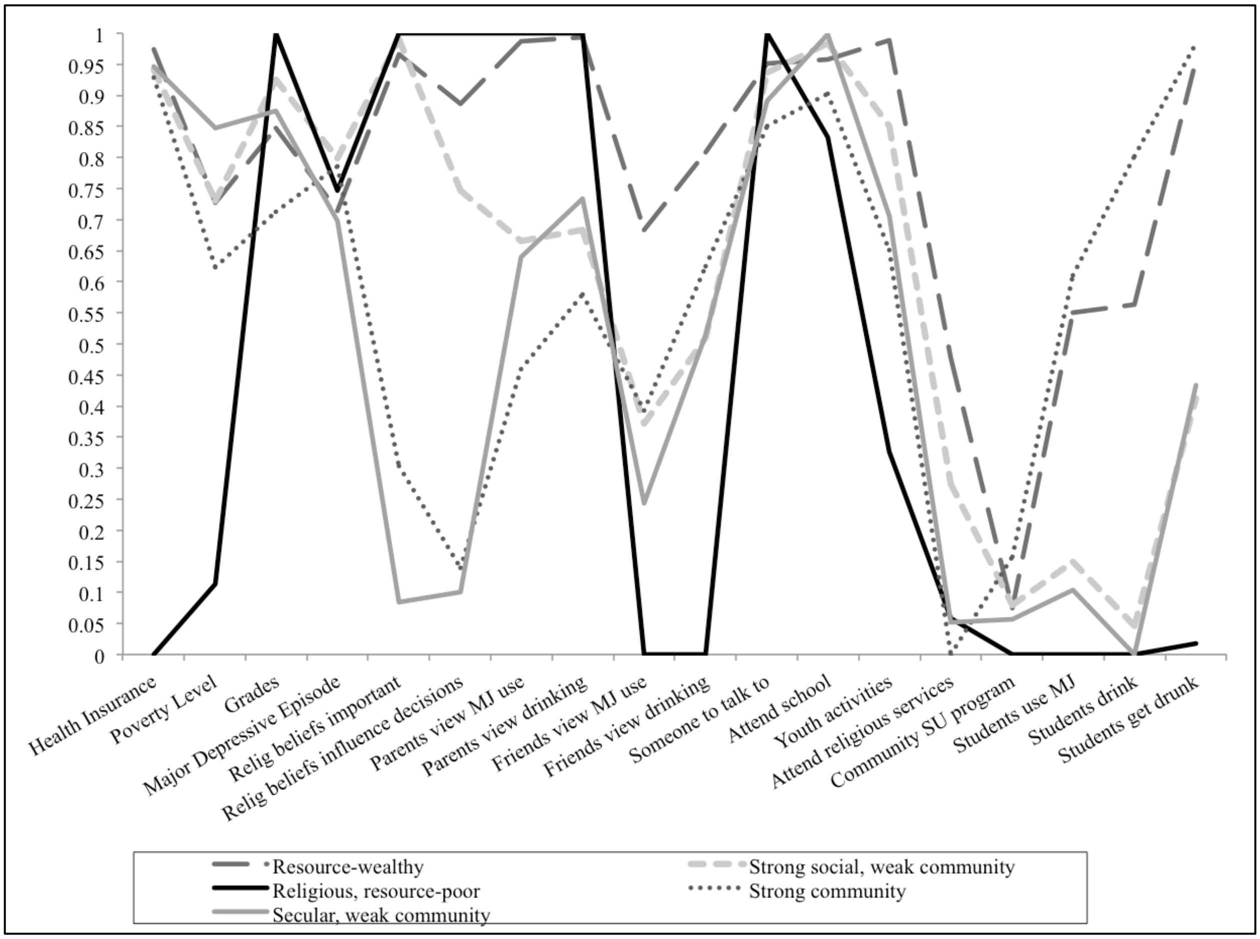

Figure 2. Item Response Probabilities for Recovery Capital Class Profiles, 5-class Model

Table 1. Descriptive statistics

	n	M (SD) / % (n)	Range
Financial RC indicators			
F1. Health insurance	1,171	94 (1101)	0-1
F2. Poverty level (below, ≥2x poverty level)	1,171	74 (867)	0-1
Human RC indicators			
H1. Grades (D or lower, A/B/C)	1,095	86 (942)	0-1
H2. Major depressive episode, past year	1,149	75 (862)	0-1
H3. Overall health (fair/poor, very good, good, excellent)	1,171	1.75 (0.87)	0-3
H4. Religious beliefs important	1,144	56 (641)	0-1
H5. Religious beliefs influence life decisions	1,144	44 (503)	0-1
Social RC indicators			
S1. Parents view MJ/hash monthly use (strongly disapprove, somewhat disapprove/neither)	1,151	67 (771)	0-1
S2. Parents view daily drinking (strongly disapprove, somewhat disapprove/neither)	1,155	74 (855)	0-1
S3. Friends view MJ/hash monthly use (strongly/somewhat disapprove, neither)	1,155	39 (450)	0-1
S4. Friends view daily drinking (strongly/somewhat disapprove, neither)	1,156	58 (670)	0-1
S5. Someone to talk to about serious problems	1,134	91 (1032)	0-1
S6. Attend school	1,164	96 (1117)	0-1
S7. Youth activity participation (none or 1, 2 or more)	1,160	78 (905)	0-1
S8. Religious service attendance, past year (<25, 25 or more)	1,156	18 (208)	0-1
Community RC indicators			
C1. Community substance use program participation	1,157	9 (104)	0-1
C2. Students use MJ/hash monthly (none/few, most/all)	1,113	30 (334)	0-1
C3. Students drink alcohol (none/few, most/all)	1,112	25 (278)	0-1
C4. Students get drunk weekly (none/few, most/all)	1,101	62 (683)	0-1
Predictor variables			
Male	1,171	49 (574)	0-1
White	1,171	57 (667)	0-1
Age	1,171	15.75 (1.27)	12-17
Ever received treatment	1,171	16 (187)	0-1

Note. All binary variables coded so that 1 = positive outcome (i.e., evidence of recovery capital) and 0 = negative outcome (i.e., lack of recovery capital). RC = *Recovery Capital*.

Table 2. Recovery Capital Model Comparisons

No. of Classes	AIC	BIC	Entropy
1	24303.87	24405.18	
2	23525.52	23728.14	0.774
3	23184.21	23488.14	0.744
4	22879.74	23284.99	0.766
5	22709.12	23215.69	0.821

Table 3. Recovery Capital Class Proportions: 5-class Model

Class	Class counts	Class proportions	Average latent class probabilities	Recovery Capital Class Profile
1	200.24	17.10%	0.853	Resource-wealthy
2	379.27	32.39%	0.856	Strong social, weak community
3	28.10	2.40%	0.987	Religious, resource-poor
4	211.70	18.08%	0.906	Strong community
5	351.69	30.03%	0.929	Secular, weak community

Table 4. Predictors of Recovery Capital Class Membership: Comparisons Across 5 Classes

	Gender			Treatment			Age			White		
		95% CI			95% CI			95% CI			95% CI	
	OR	LCL	UCL	OR	LCL	UCL	OR	LCL	UCL	OR	LCL	UCL
Religious, resource-poor versus:												
Resource-wealthy	0.66	0.10	4.55	0.47	0.06	3.93	0.17 *	0.04	0.66	23.27 *	1.79	302.67
Strong social, weak community	0.83	0.12	5.61	0.28	0.03	2.22	0.29	0.08	1.13	22.85 *	1.81	288.66
Strong community	1.76	0.26	11.91	0.55	0.07	4.43	0.16 **	0.04	0.63	27.99 **	2.21	355.01
Secular, weak community	0.84	0.13	5.58	0.24	0.03	1.90	0.37	0.10	1.45	47.80 **	3.79	602.62
Resource-wealthy versus:												
Strong social, weak community	1.26	0.64	2.47	0.60	0.24	1.49	1.72 ***	1.24	2.39	0.98	0.51	1.90
Strong community	2.67 **	1.34	5.32	1.17	0.50	2.76	0.95	0.69	1.31	1.20	0.63	2.32
Secular, weak community	1.27	0.70	2.32	0.51	0.22	1.18	2.19 ***	1.59	3.01	2.05 *	1.11	3.79
Strong social, weak community versus:												
Strong community	2.12 *	1.16	3.88	1.97	0.91	4.25	0.55 ***	0.42	0.72	1.23	0.37	6.24
Secular, weak community	1.01	0.62	1.64	0.86	0.42	1.75	1.27	0.98	1.65	2.09 **	1.26	3.47
Strong community versus:												
Secular, weak community	0.48 *	0.26	0.87	0.44 *	0.20	0.96	2.31 ***	1.72	3.10	1.71	0.96	3.03

Note. Odds ratios represent the estimated odds ratio of class membership relative to odds of reference group: each model includes all four predictors (gender, prior treatment receipt, age, white). OR = Odds ratio. LCL = Lower confidence interval. UCL = Upper confidence interval. *** p<.001, ** p < .01, * p < .05

CHAPTER III

ADOLESCENT RECOVERY CAPITAL AND RECOVERY HIGH SCHOOL ATTENDANCE: AN EXPLORATORY DATA MINING APPROACH

Introduction

Adolescents report access to alcohol or drugs in many places, including at home, through friends, and at school (Centers for Disease Control and Prevention, 2012; Spear & Skala, 1995; Substance Abuse and Mental Health Services Administration [SAMHSA], 2013). Although not every adolescent who experiments with alcohol or illicit drugs is diagnosed with a substance use disorder (SUD), according to the Treatment Episode Data Set (TEDS), in 2014, approximately 78,000 adolescents aged 12-17 were admitted to publicly-funded substance abuse treatment in the United States (SAMHSA, 2016). Many adolescents who leave treatment often return to the same community, whether it be home, school, and/or neighborhoods, where they were previously using substances, and these "recovery hostile" environments may increase the risk of relapse (Brown et al., 2001; Ramo, Prince, Roesch, & Brown, 2012; White et al., 2004). Indeed, given the availability of substances in environments adolescents frequent as well as how highly influential social pressures are in instances of relapse (Ramo et al., 2012), continuing care services should focus on settings of potential risk and address gaps in community supports by promoting environments that foster sobriety among youth (Gonzales et al., 2012; White, 2009). To address this need, recovery high schools (RHSs) were created in the 1970's to ensure academic advancement and recovery maintenance among adolescents in recovery (Finch & Frieden, 2014). Since then over 70 RHSs across the United States have been in operation (Finch et al., 2016) and federal offices recently highlighted RHSs as viable youth recovery supports (National Institutes of Drug Abuse, 2014; Office of National Drug Control Policy, 2014).

RHS Model and Recovery Capital

From an ecological perspective, RHSs could be considered one form of community recovery capital. Recovery capital is an ecological model that attends to individuals, their social networks, and their communities and encompasses all resources that an individual has to use towards SUD recovery (Granfield & Cloud, 1999; White & Cloud, 2008). This model has recently been adapted from previous adult-focused recovery capital models to fit the adolescent experience (Hennessy, 2017). It is comprised of four primary domains: (1) *financial recovery capital*: material resources such as caregiver income, health insurance, and access to treatment (Granfield & Cloud, 1999); (2) *human recovery capital*: personal characteristics that one can use to achieve personal goals (Granfield & Cloud, 1999); (3) *social recovery capital*: resources that enable an adolescent to effectively bond with family, peers, and community institutions and the resources available to an individual through these relationships (White & Cloud, 2008); (4) *community recovery capital:* including community-level, culturally-appropriate resources related to addiction and recovery (White & Cloud, 2008) and cultural capital such as individual values and behavioral patterns generated from cultural group membership(s) that support sobriety (Burns & Marks, 2013; White & Cloud, 2008).

The primary focus in an RHS is academic advancement, however, RHSs require a student to be sober or to have a desire to remain abstinent and work toward an abstinent lifestyle. In addition, RHSs provide recovery-focused services such as daily group check-ins and individual counseling for students. These elements are designed to create a recovery-supportive culture where both students and staff attend to academic development and recovery maintenance

(Karakos, 2014; Moberg & Finch, 2007). As a recovery-specific educational institution in the community, RHSs provide a tangible form of community recovery capital for students and could generate new or build upon existing recovery capital. For example, school attendance, engagement, and academic achievement could motivate adolescent personal growth as well as replace substance use activity, peers, and environments with alternative positive activities and prosocial, sober peers. Thus, RHSs could foster human recovery capital, by teaching adolescents academic and other skills for postsecondary education or later employment, as well as foster social recovery capital, by surrounding an adolescent with other supportive peers who are dedicated to recovery (Kochanek, 2008; Moberg & Finch, 2007).

Prior research suggests that individuals with some recovery capital have a better chance of developing and accessing more recovery capital (Granfield & Cloud, 2001). For example, among adults in recovery, healthy social relationships can result in access to financial recovery capital including money or loaned possessions (Neale & Stevenson, 2014), and the development of recovery capital has been described as an ongoing process where human and social recovery capital evolve and support the growth of each other (Best et al., 2012; 2011). Thus, in the case of RHSs, adolescents with a certain degree of financial stability may attend a treatment center that offers continuing care supports and is knowledgeable about community resources, whereas adolescents with less financial stability might not have access to such resources. Alternatively, adolescents with supportive, connected parents or friends in recovery may be more likely to hear about an RHS, or through peer affiliation, adolescents might be more motivated to attend a school focused on sobriety, such as an RHS.

However, each of the 19 current RHSs in the United States only enroll between 2 and 115 students each year (Association of Recovery Schools [ARS], 2016). Thus, far fewer adolescents in the United States attend an RHS relative to the number of adolescents admitted to treatment centers each year (SAMHSA, 2016). As such, despite the recovery support an RHS offers, many adolescents do not attend an RHS posttreatment. Prior research that has explored RHS initiation or closure suggests that a lack of students, possibly due to lack of transportation, might be a reason for low attendance rates (Finch et al., 2016). Preliminary research has also explored potential disparities in access to RHSs, given, for example, the requirement of most RHSs that youth have previously attended formal SUD treatment, which may not be feasible for low-income students or youth involved with the juvenile justice system (Oser et al., 2016). However, research to date has not empirically addressed potential predictors of attending an RHS. In addition, although there is an extensive body of research on adult recovery capital, few studies have applied the model to adolescents (Hennessy, 2017; e.g., see Best et al., 2012; Duffy & Baldwin, 2013; Neale et al., 2014), a population with treatment and recovery patterns different from adults (Brown & Ramo, 2006; Winters, Tanner-Smith, Bresani, & Meyers, 2014).

Study Aims

Given our lack of understanding about the characteristics predictive of RHS attendance, this exploratory study aimed to examine factors predicting RHS attendance. As RHSs are a form of community recovery capital specific to adolescents, factors from all recovery capital domains (i.e., financial, human, social, and community) will be used to predict RHS attendance among a sample of youth in recovery. Conceptualized this way, recovery capital suggests that adolescents attending an RHS, which is indicative of community recovery capital, would have higher levels of other recovery capital resources, compared to adolescents who do not attend an RHS. However, it is possible that some factors, such as length of time in treatment, indicate different degrees of severity of substance use problems. Indeed, it is likely that factors interact, with some factors producing a greater likelihood of attendance while others diminish the possibility of attendance.

Given this complexity and the relative lack of empirical research on the recovery capital model for adolescents, using exploratory data analysis methods may be more useful than confirmatory methods of analysis. Thus, this study employs three exploratory statistical techniques including (1) SEARCH, a binary segmentation procedure (Sonquist et al., 1974; Morgan, 2005); (2) classification trees (McArdle, 2014) and (3) an ensemble classification tree method, random forest (Breiman, 1996; 2001). These methods were developed in part as a result of the interactions between the many interacting variables important to social science issues (McArdle, 2014; Morgan & Sonquist, 1963). Each method sequentially tests available predictors and attempts to categorize participants into subgroups using covariate(s) that best predict the outcome of interest (Morgan & Sonquist, 1963; Sonquist et al., 1974). The resulting subgroups are thus indications of interactions between included variables. These identified interactions can be important for identifying what characteristics, in combination, predict a particular outcome—a useful technique for under-researched questions where there are few hypothesized theoretical relationships (Blankers et al., 2013; Dierker et al., 2004; Doyle & Donovan, 2014; McArdle, 2014; Morgan, 2005; Sonquist & Morgan, 1963; Scott et al., 2014).

Methods

Data

This project used data collected through a multi-site, longitudinal, observational study on the effectiveness of RHSs, which received Institutional Review Board approval from the University of Minnesota. In this study, adolescents and their caregiver(s) in Minnesota, Wisconsin, and Texas were recruited at the end of adolescents' substance use treatment or at the initiation of school attendance posttreatment, and interviewed at baseline, 3-, 6- and 12-month follow-ups. Trained interviewers collected study data through youth and caregiver interviews lasting 60-90 minutes. Prior to interviews, study staff met with a potential adolescent participant and his/her caregiver to review the consent/assent forms. Adolescents younger than 18 years were given assent forms and their respective caregiver signed consent for themselves and their adolescent to participate.

The interview schedule was compiled from a set of previously validated measures and for the current study, a subset of variables collected from youth during baseline and 3-month interviews were selected and categorized by recovery capital domain (see Appendix A, Table 1 for detailed information on scale properties). Control variables, including demographic information and substance use history and diagnoses, were collected using the Adolescent Diagnostic Interview (ADI and ADI-Parent: Winters & Henley, 1993; Winters & Stinchfield, 2000), MINI-Structured Clinical Interview (MINI-SCID: Sheehan, Janavs, Baker, Harnett-Sheehan, Knapp, & Sheehan, 1999), and Timeline Followback (TLFB: Sobell & Sobell, 1992). Financial recovery capital variables included family income, health insurance, homelessness history (GAIN: Dennis, 2010), and parent education and occupation, which were combined into a single social position score for the family using the Social Position Index (Haug & Sussman, 1970). Human recovery capital variables included the following: substance use expectancies (Personal Experience Inventory [PEI]: Winters & Henley, 1989); school attitudes (Behavior Assessment System for Children [BASC]: Reynolds & Kamphuas, 1992); life satisfaction, school problems, stress, and physical health (Global Appraisal of Individual Needs [GAIN]: Dennis, 2010); a five-factor measure of problem solving orientation and skills (Social Problem Solving Inventory: D'Zurilla & Nezu, 1990; D'Zurilla, Nezu, & Maydeu-Olivares, 2002); and mental health diagnoses (MINI-SCID: Sheehan et al., 1999). Social recovery capital variables included social and spiritual social support (GAIN: Dennis, 2010); substance approving peer attitudes (PEI: Winters & Henley, 1989); social competence (High School Questionnaire [HSQ]: Moberg & Finch, 2008); neighborhood social connections (Profiles of Student Life-Attitudes and Beliefs [PSL-Adapted]: Leffert et al., 1998); and youth-parent relationships (Youth Happiness with Parent Scale: DeCato, Donohue, Azrin, & Teichner, 2001). Community recovery capital variables included perceived availability of drugs (modified Monitoring the Future survey: Johnston, O'Malley, Bachman, & Schulenberg, 2011); youth knowledge of RHS, receipt of AOD/SUD counseling outside school, and 12-Step meeting attendance frequency (HSQ: Moberg & Finch, 2008). For additional information on the larger parent study and details on the rationale for and development of the interview schedule, see Botzet, McIlvaine, Winters, Fahnhorst, and Dittel (2014).

In all analyses, baseline data from the 260 adolescent participants retained at 3 months (88.7% retention rate) reflecting the four recovery capital domains were used to predict the 3-month outcome; whether or not an adolescent enrolled at an RHS (1 = yes; n = 120) or did not, including enrolling in another school, such as a public school (0 = no; n = 140). For each of the methods, five separate analyses were conducted: for each of the four recovery capital domains (either *financial*, *human, social*, or *community*), variables specific to that domain were included as covariates in separate analyses to predict RHS attendance (See Table 1). Then the fifth analysis included important variables from each domain to predict RHS attendance from an assessment of total recovery capital (*financial* + *human* + *social* + *community* domains).

An additional set of 11 variables was included in each analysis as these variables were expected to interact with recovery capital. For example, age, race[4], and sex were included because research has demonstrated their importance to recovery outcomes (e.g., Becker et al., 2012; Sterling et al., 2009; Stevens et al., 2004; Wellman et al., 2014). Previous SUD service receipt and baseline substance use (alcohol, marijuana, other drugs) were also included as these variables are likely to affect recovery capital (Becker et al., 2012; Stevens et al., 2004).

Analysis

This study compares four primary statistical methods: (1) logistic regression; (2) SEARCH (Sonquist et al., 1974; Morgan, 2005); (3) classification trees (McArdle, 2014); (4) random forest (Breiman, 1996; 2001). Generally, each of the exploratory methods (2-4, above) use the available data to sequentially split it into mutually exclusive subgroups by the covariate that best predicts the dependent variable; that is, it seeks and chooses the covariate that will account for more of the variance than another predictor variable (Morgan & Sonquist, 1963; Sonquist et al., 1974; Therneau, Atkinson, & Ripley, 2015). Due to the sequential nature of these exploratory methods, they will be compared to a traditional logistic regression approach because unlike traditional regression methods where interactions must be specified a priori, exploratory methods are likely to identify important interactions between variables if they exist (Blankers et al., 2013; Dierker et al., 2004; Doyle & Donovan, 2014; McArdle, 2014; Morgan, 2005; Sonquist & Morgan, 1963; Scott et al., 2014). These exploratory methods are not limited by variable collinearity or missing data. This study also compares multiple exploratory methods because although previous research has demonstrated similarities in accuracy between various data mining approaches (Lim, Loh, & Shih, 2000), it is possible that the different methods would

[4] Given the proportion of White youth compared to youth of other race/ethnicities in the sample, one binary variable indicating White/non-White was included to represent race. The baseline demographics in Table 1 provide the percentages for all races.

produce divergent outcomes due to the use of different software and algorithms. Thus, comparing predictors of RHS attendance using these exploratory methods is useful to find if and where they agree.

Logistic regression. McArdle (2014) and others (e.g., see Cattell, 1966; Tukey 1962) have recommended first conducting confirmatory analysis and then any exploratory analyses. Thus, this study initially conducted separate confirmatory logistic regressions in Stata SE (14.2; StataCorp, 2015), one for each unique recovery capital domain. Significant predictors ($p < .05$) in each recovery capital domain from these four regressions were then included in one final regression equation of overall recovery capital.

SEARCH. Following the logistic regressions, analyses were conducted using the SEARCH program in Stata SE (14.2; StataCorp, 2015) using the same variables entered in the logistic regression models for each recovery capital domain; however, to aid in estimation, continuous variables used in this analysis were first transformed to categorical variables, with the exception of age and age of first alcohol or other drug (AOD) treatment which remained continuous. Due to the primary focus on recovery capital, variables that directly related to recovery capital factors were given first priority in determining each split in the models (rank = 1), whereas the control variables were given second priority (rank = 2), meaning that if a recovery capital predictor satisfied the model selection criteria it was selected before a control variable. Given the small sample, all models employed selection criteria of each split needing to explain at least 40% of the variance in the dependent variable and resulting in a minimum group size of at least five individuals in each group created from the split.[5]

[5] Predictors that emerged as important from these four SEARCH analyses were originally planned for inclusion in one final SEARCH analysis of overall recovery capital; however each domain had almost identical results given the high importance of one control variable, so the final model of overall recovery capital used all available predictors from each of the recovery capital domains and the control variables.

Classification trees. Next, five classification trees were created in the RStudio statistical environment (0.99.486) utilizing the Recursive PARTitioning (`rpart`) package with the same variables entered in previous analyses for each recovery capital domain (Therneau, Atkinson, & Ripley, 2015). Often when using this method the sample is split into a training and a test data set, but given the small sample size in the current study, 10-fold cross validation was instead used where the sample is randomly divided into 10 equal subsamples (Hastie, Tibshairani, & Friedman, 2008). For each iteration a tree is grown from nine of the subsamples and the 10th is used as a pseudo-test sample so that the sum of squared errors for the tree can be calculated. This process is repeated 10 times, allowing each sub-sample to serve as a pseudo-test and training sample. Trees were allowed to grow with no stopping rules, and the complexity parameter, a measure of which splitting a variable node improves the relative error (Therneau et al., 2015) was used to prune each tree after growth. This helps to ensure a parsimonious model without spurious variables that capitalized on chance (e.g., see Breiman et al., 1994; Torgo, 1999). In the final model of overall recovery capital, only those variables that appeared in the pruned tree for each recovery capital domain were included in the analysis.

Random forest. To further protect against including spurious predictors, an ensemble method utilizing bootstrap aggregations (Breiman, 1996; 2001; Scott et al., 2014) in the R package, `randomForest` (Liaw & Weiner, 2015), was used to create random forests for each recovery capital domain and overall recovery capital. The forest method is similar to the sequential variable selection used in the classification trees but operates using boosted aggregations of the sample to generate X number of trees and randomly chooses a set of variables (k) to attempt at each split. Based on previous recommendations, 1000 trees were generated and k was set to three (e.g., see Hastie et al., 2009; Liaw & Wiener, 2002; Strobl et al., 2008). The results across all trees in the forest are used to identify consistently important predictors. Because of the random selection of variables in each tree and the number of trees generated, correlation between the trees is reduced and no pruning due to overfitting is necessary (Hastie et al., 2008). For each tree in the forest, some of the sample is left out and dropped through the resulting tree: these results are aggregated across trees resulting in an out-of-bag (OOB) error rate that demonstrates how well the model predicts the outcome of interest. However, given the number of trees created, extracting and relying on a single visual tree is not appropriate with this method. Instead, the variable importance results, a measure of the prediction strength of each variable based on all trees (Hastie et al., 2008), are presented and can be used to assess how well the results might generalize to other samples.

Random forests were created for each recovery capital domain and for overall recovery capital. All variables included across the individual recovery capital domain models were given a variable importance value based on the Gini Impurity Index where higher values indicate greater variable importance. Because of their inclusion in each unique recovery capital domain model, the control variables were often given multiple variable importance values, depending on which model was explored. For the final model of overall recovery capital, the variable importance results from the previous four models of individual domains of recovery capital were reordered with the highest variable importance value chosen for those variables with multiple variable importance measures. These values were then visualized to assess where large drops in variable importance occurred (see Appendix A, Figure 1) and the top 22 variables (variable importance ranged from 4.12 to 11.29) were chosen for inclusion in the final random forest of overall recovery capital to predict RHS attendance[6].

Results

This sample of recovering youth was primarily male (53%) and White (84%) and was around 16 years of age (SD = 1.03; range 13-19 years). On average, the sample had first received treatment for AOD use around the age of 15 years (SD = 1.20; range 12-18 years): at baseline, 63% were diagnosed with an alcohol use disorder and 95% were diagnosed with a drug use disorder. Results from each unique domain will be presented first, followed by the results from the overall model of recovery capital. To aid in interpretation while reading the results, readers are directed to Table 2 which is a summary table providing the results of the logistic regressions, SEARCH, and classification trees for each variable included in each domain model and whether that method identified each as a significant predictor of RHS attendance in the model (indicated by a "+").

Financial Recovery Capital

Within the financial recovery capital domain, logistic regressions indicated that only age was a significant predictor of RHS attendance, with the odds of attending an RHS being 1.48 higher for older adolescents (95% CI [1.02, 2.17]). The model explained approximately 9% of the variance in RHS attendance (χ^2 = 28.63, p = 0.02). See Table 3 for results of all included variables.

The SEARCH analysis contained one root node with two terminal nodes where participants were only distinguished by the number of times of previous SUD treatment receipt,

[6] Each time the random forest model is estimated new variable importance values are generated based on the included variables for the model. Although these 22 variables were chosen for inclusion in the final model of overall recovery capital based on their higher variable importance, once in this model, they were given new variable importance values, which are presented in the column for overall recovery capital (Table 3).

explaining 4% of the variance (Figure 1). Adolescents with 1-4 instances of SUD treatment were more likely to be in the non-RHS group, and adolescents with more than four instances of SUD treatment were more likely to be in the RHS group.

Similarly, the classification tree (Figure 2) contained one root node with two terminal nodes and correctly identified approximately 18% (n = 21) of RHS attendees and 95% (n = 133) of non-RHS attendees (pseudo-R^2 = 12%). The tree indicated that only the number of times of previous SUD treatment receipt was predictive of RHS attendance (ranked 90% important), with adolescents having fewer than 4.5 instances of SUD treatment more likely to be in the non-RHS group and adolescents with 4.5 or more instances of SUD treatment more likely to be in the RHS group. The age of first AOD treatment was ranked at 10% importance but was not included in the final pruned tree given its relatively minor predictive utility.

Results from the random forest indicated multiple important variables (see column 1, Table 4), one of which was the SUD treatment receipt variable identified in the classification tree as well as age, which was identified as significant in the logistic regression. However, several variables were ranked as more important than both of these variables, indicating the need to be cautious in interpreting results from single trees: this included two of the financial recovery capital factors (parental social position and parental income) as well as five of the control variables (use of other drugs, use of alcohol or marijuana, number of times received mental health services, and age of first AOD treatment). The out-of-bag error rate was 45%: 66% (n = 93) of the non-RHS students and 42% (n = 51) of the RHS students were correctly classified.

Human Recovery Capital

Within the human recovery capital domain, logistic regressions indicated that the odds of attending an RHS were higher for youth with negative attitudes toward school (OR = 5.60, 95%

CI [1.04, 30.90]), a diagnosis of major depression (OR = 3.19, 95% CI [1.15, 8.88]), higher life satisfaction (OR = 2.77, 95% CI [1.32, 5.83]), higher reported stress (OR = 1.35, 95% CI [1.03, 1.78]), and more days of school attendance in the past 12 months (OR = 1.02, 95% CI [1.00, 1.03]). In addition, the odds of RHS attendance were 0.53 lower for youth with higher levels of rational problem solving (95% CI [0.33, 0.87]) and 1.74 higher for youth with higher levels of impulsive/careless problem solving (95% CI [1.02, 2.98]). The model explained approximately 28% of the variance in RHS attendance (χ^2 = 71.99, p = 0.00).

The SEARCH analysis contained two branches with three terminal nodes, explaining 7.1% of the variance (see Figure 3). Adolescents with the highest levels of rational problem solving factors scores (above the average rational problem solving factor score) were more likely to attend a non-RHS. Alternatively, for adolescents around or below the average rational problem solving scores, those who had over four instances of previous SUD treatment were more likely to attend an RHS, while those with 1-4 previous instances of SUD treatment more likely to attend a non-RHS.

The classification tree was comprised of 10 pairs of branches and 11 terminal nodes (see Figure 4). This model correctly identified approximately 76% (n = 91) of RHS attendees and 79% (n = 111) of non-RHS attendees (pseudo-R^2 = 52%). Similar to both the SEARCH and the logistic regression results, the classification tree identified rational problem solving scores as highly predictive of RHS attendance. In addition, the classification tree and the logistic regression both identified days of school attendance in the past 12 months and a diagnosis of major depression as factors predictive of RHS attendance. However, the classification tree also identified additional interacting factors as predictive of RHS attendance, resulting in multiple pathways to school type classifications.

The tree's primary split was on the rational problem solving factor score and results for youth with lower rational problem solving factor scores (right side of the tree) will be discussed first. For these youth, a diagnosis of major depression and a lower negative problem orientation factor score was predictive of RHS attendance while a diagnosis of major depression and higher negative problem orientation factor scores resulted in non-RHS attendance. Alternatively, youth with lower rational problem solving scores but no diagnosis of major depression and a low score on social benefits of substance use expectancies were more likely to attend a non-RHS while those with a high score on social benefits of substance use expectancies were more likely to attend an RHS.

Moving to the left side of the tree, youth with higher rational problem solving scores, higher positive problem orientation and greater number of days of school attendance were all more likely to attend an RHS, while youth with similar trajectories but fewer days of school attendance were likely to attend a non-RHS. Youth with higher rational problem solving scores, lower positive problem orientation scores, and no diagnosis of PTSD were more likely to attend a non-RHS, while youth with a similar trajectory but a dual-diagnosis of PTSD and Mania were also more likely to attend a non-RHS. Youth with higher rational problem solving scores, lower positive problem orientation scores, a diagnosis of PTSD but not mania and a higher physical health rating were more likely to attend an RHS. Alternatively, youth with higher rational problem solving scores, lower positive problem orientation scores, a diagnosis of PTSD but not mania, a lower physical health rating, and fewer than 4.5 instances of MH services were more likely to attend an RHS, while students with a similar trajectory but with 4.5 or more instances of MH services were more likely to attend a non-RHS.

Results from the random forest indicate multiple important variables (see column 2, Table 4), and agreed with the single tree that the rational problem solving factor score was the most important indicator of RHS attendance. In addition, the random forest demonstrated that, as identified in the single tree, days in school during the past 12 months and positive and negative problem orientation were important factors in predicting RHS school attendance. However, there may be other important variables that the single tree missed such as impulsivity/carelessness style of problem solving, life satisfaction, an avoidance style of problem solving, and days of other drug use in the previous three months. Thus, although all variables identified as important in the classification tree were rated as important in the random forest, the aggregations across the forest indicate that some of the variables identified in the tree may be due to random variations in the data. The out-of-bag error rate was 37%: 74% (n = 104) of non-RHS students and 50% (n = 60) of RHS students were correctly classified.

Social Recovery Capital

Within the social recovery capital domain, logistic regressions found that only sex was a significant predictor of RHS attendance, with the odds of attending an RHS being 0.50 lower for males (95% CI [0.28, 0.89]). Similar to results from the financial recovery capital model, this model explained approximately 9% of the variance in RHS attendance (χ^2 = 30.44, p = 0.05).

The SEARCH analysis for this dimension matched previous results from the financial recovery capital domain (see Figure 1). The results again contained only one root and two terminal nodes where participants were distinguished by the number of times of previous SUD treatment receipt: adolescents with 1-4 instances of SUD treatment more likely to attend a non-RHS, and adolescents more than four instances of SUD treatment more likely attend an RHS.

The classification tree contained three branches with four terminal nodes (see Figure 5). This model correctly identified approximately 53% (n = 64) of RHS attendees and 76% (n = 106) of non-RHS attendees (pseudo-R^2 = 25%). The classification tree and the SEARCH analysis again agree that number of previous SUD treatment receipt instances was important in predicting RHS attendance. Unlike the logistic regression results, the classification tree did not identify sex as important, but instead identified number of previous SUD treatment receipt instances, substance-approving peers, and neighborhood social connection as predictive of RHS attendance. Youth with more than 4.5 instances of prior treatment receipt were more likely to attend an RHS. For youth with less than 4.5 instances of prior treatment, higher ratings of substance-approving peers and lower neighborhood social connection predicted RHS attendance for about half the remaining youth who attended RHSs, while the other half reported lower substance-approving peer attitudes.

Results from the random forests confirmed the importance of the three variables identified as predictive of RHS attendance in the classification trees, with substance-approving peers identified as the most important, followed by neighborhood social connection and prior SUD treatment receipt (see column 3, Table 4). However, there may be other important variables that the single tree missed such as youth-parent relationship ratings, days of use of alcohol, marijuana, or other drugs, number of times received MH services, and social competence. Thus, although all variables identified as important in the classification tree were rated as important in the random forest, the aggregations across the forest indicate that some of the variables identified in the tree may not be as important as originally rated. The out-of-bag error rate was 45%: 65% (n = 91) of non-RHS students and 42% (n = 51) of RHS students were correctly classified.

Community Recovery Capital

Within the community recovery capital domain, the odds of attending an RHS were higher for adolescents who frequently attended self-help groups (OR = 1.39, 95% CI [1.11, 1.74]) and were lower for youth who knew about RHSs prior to treatment (OR = 0.37, 95% CI [0.19, 0.73]) and who were males (OR = 0.43, 95% CI [0.22, 0.84]). This model explained approximately 13% of the variance in RHS attendance (χ^2 = 37.78, p = 0.00).

The SEARCH analysis matched previous results from financial and social recovery capital domains (see Figure 1): it again contained one root node with two terminal nodes where participants were only distinguished by the number of times of previous SUD treatment receipt (1-4 versus more than four previous instances of SUD treatment).

The classification tree contained 8 pairs of branches and 9 terminal nodes (see Figure 6). This model correctly identified approximately 63% (n = 75) of RHS attendees and 79% (n = 111) of non-RHS attendees (pseudo-R^2 = 34%). The classification tree and the SEARCH results matched for at least the initial split on number of previous SUD treatment experiences: youth with more than 4 (SEARCH) or more than 4.5 (classification tree) instances of prior SUD treatment receipt were more likely to attend an RHS. Additionally, similar to the logistic regressions, the classification tree identified knowledge of RHSs prior to treatment and self-help group attendance as predictive of RHS attendance. However, unlike the SEARCH and the logistic regression results, the classification tree also identified a number of other factors as predictive of RHS attendance.

For youth with less than 4.5 instances of prior treatment, there were a few different pathways predicting school choice. First, youth with less previous treatment receipt and those that attended 12-step meetings never/less than once a month were more likely to attend a non-RHS. Second, those with less previous treatment receipt but who attended 12-step meetings at least once a month or more, and had more than 79 days of other drug use in the past 90 were more likely to be an RHS student. Third, those with less previous treatment receipt but who attended 12-step meetings at least once a month or more, used other drugs for less than 79 days in the past 90, and knew about the RHS prior to treatment, were more likely to attend a non-RHS if they received AOD/MH counseling outside school and, alternatively, to attend an RHS if they did not receive AOD/MH counseling outside school. Fourth, those with fewer instances of previous treatment receipt but who attended 12-step meetings at least once a month or more, used other drugs for less than 79 days in the past 90, did not know about the RHS prior to treatment and perceived less availability of access to drugs, were more likely to attend a non-RHS. Finally, those with fewer instances of previous treatment but who attended 12-step meetings at least once a month or more, used other drugs for less than 79 days in the past 90, did not know about the RHS prior to treatment and perceived greater access to drugs, but less than the top scoring group were likely to attend an RHS, while those with a similar trajectory but in the top scoring group of perceived access to drugs were more likely to attend an RHS if they had received MH treatment services less than 1.5 times.

Results from the random forests confirmed the predictive importance of five of the seven variables identified in the classification trees (see column 4, Table 4), including two community recovery capital factors (perceived availability of drugs and frequency of 12-Step meeting attendance) and three control factors (days of other drug use, number of previous MH services and previous AOD treatment). Although the random forest identified knowledge of RHS prior to treatment and current AOD/MH treatment as important, which were also important in the single tree, these were ranked as less important than other variables and thus should be interpreted with caution. For example, days of alcohol and marijuana use, age first treated for AOD, and age were all ranked as more important in the random forest results than prior knowledge of RHS and current AOD/MH treatment. The out-of-bag error rate for the community recovery capital random forest was 43%: 65% of non-RHS students (n = 91) and 48% (n = 58) of RHS students were correctly classified.

Overall Recovery Capital

Based on the logistic regression analyses for each unique recovery capital domain, age, sex, seven human recovery capital variables, and two community recovery capital variables were included in the final regression model of overall recovery capital. This model explained approximately 20% of the variance in RHS attendance (χ^2 = 63.19, p = 0.00). The odds of attending an RHS were significantly higher for youth with more negative school attitudes (OR = 8.27, 95% CI [2.24, 30.53]), higher life satisfaction (OR = 2.21, 95% CI [1.28, 3.81]), a major depression diagnosis (OR = 2.50, 95% CI [1.11, 5.68]), and more frequent self-help group attendance (OR = 1.29, 95% CI [1.06, 1.58]). The odds of attending an RHS were significantly lower for youth with higher levels of rational problem solving (OR = 0.65, 95% CI [0.44, 0.96]) and those who had knowledge of RHSs prior to treatment (OR = 0.40, 95% CI [0.21, 0.76]).

The SEARCH analysis for overall recovery capital exactly matched the results for human recovery capital and again contained two branches with three terminal nodes (see Figure 4). Adolescents with the highest levels of rational problem solving factor scores compared to the mean of the entire sample were more likely to attend a non-RHS. For adolescents around or below the mean of rational problem solving scores, those who had over four instances of previous SUD treatment were more likely to attend an RHS, while those with 1-4 previous instances of SUD treatment more likely to attend a non-RHS.

All variables included in the pruned trees from each unique recovery capital domain were included in the classification tree of overall recovery capital (k = 18). The resulting classification tree contained seven pairs of branches and eight terminal nodes with variables from the human recovery capital and community recovery capital domains only (see Figure 7). This model correctly identified approximately 68% (n = 81) of RHS attendees and 82% (n = 115) of non-RHS attendees (pseudo-R^2 = 47%). Again, similar to the SEARCH results, the classification tree identified rational problem solving factor scores as important, and similar to the logistic regressions, the classification tree identified both rational problem solving factor scores and frequency of self-help group attendance as predictive of RHS attendance. However, the tree again demonstrated divergent pathways to school attendance based on a number of additional variables.

First, students with a lower rational problem solving factor score were more likely to be non-RHS students if they attended 12-Step programs never/less than once a month and more likely to be RHS students if they attended 12-Step programs more than once a month. The remaining pathways were marked by higher factor scores on rational problem solving. For these students, if they also had higher positive problem orientation, they were more likely to be RHS students if they attended school more than 144 days in the past year and were more likely to be non-RHS students if they attended for fewer than 144 days. Alternatively, for students with lower positive problem orientation scores, the different pathways were marked by mental health diagnoses and current treatment receipt. For example, non-RHS youth were more likely to (a) have lower positive problem orientation and no diagnosis of PTSD or (b) have lower positive problem orientation, a diagnosis of PTSD and a diagnosis of a manic episode. Finally, RHS youth with a lower positive problem orientation, and a diagnosis of PTSD but not a manic episode were more likely than non-RHS youth to report receiving AOD or MH counseling outside the school.

Results from the random forests using 22 factors identified as important across the unique recovery capital domains confirmed the importance of four of the seven variables identified as predictive of RHS attendance in the overall recovery capital classification tree (see column 5, Table 4), including three human recovery capital factors (positive problem orientation, days in school during the previous 12 months, and rational problem solving) and one community recovery capital factor (12-Step meeting frequency attendance). Again, rational problem solving was ranked as most important in the random forest and was also the first split in the classification tree. However, although the single classification tree identified manic episodes or PTSD and current receipt of AOD/MH services outside school as important, these were not identified as important in the random forests and should be interpreted with caution from the single classification tree. In addition, other factors were not identified as important in the single tree but were rated as important in the random forests. This included four human recovery capital factors (impulsivity/carelessness style problem solving, negative problem orientation, avoidance style problem solving, and life satisfaction), three social recovery capital factors (substance approving peers, youth-parent relationship, neighborhood social connection), one financial recovery capital factor (parent social position), and three control variables (days of other drug, alcohol, and marijuana use), as well as other factors with a variable importance rating below three. The out-of-bag error rate for the overall recovery capital random forest was 37%: 75% of non-RHS students (n = 105) and 50% (n = 60) of RHS students were correctly classified.

Discussion

Overall, there appear to be multiple and diverse pathways to attending an RHS. Contrary to expectations, youth attending RHSs did not simply have higher recovery capital than youth attending other types of schools. RHS students, like other students in recovery, have a mix of strengths and vulnerabilities. According to most of the methods used in this study, the best predictors of RHS attendance were primarily human recovery capital factors, including problem solving orientation and skills, days of school attendance, and life satisfaction.

Indeed, positive problem orientation and rational problem solving, measures of positive and adaptive problem solving cognition and applied skills, were two consistently important initial factors for predicting RHS attendance (Maydeu-Olivares & D'Zurilla, 1996); however, not necessarily in the direction that would be expected. Students were more likely to be in an RHS than a non-RHS if they had a combination of lower rational problem solving style factor scores and frequently attended 12-Step meetings. Coping skills such as problem solving have been linked to engagement in substance use behaviors (Jaffee & D'Zurilla, 2009; Wills, Sandy, Yaeger, Cleary, & Shinar, 2001), especially among youth who heavily use substances (Waldron & Kaminer, 2004). As a result, treatment and continuing care supports, such as RHSs, emphasize building life skills such as problem solving to help youth better manage the daily challenges of being an adolescent and in recovery. Thus it is possible that some youth and/or their families are aware of their problem solving deficits and view RHSs and 12-Step meetings as environments to learn and practice these skills and so engage in both. Conversely, some youth with low rational problem solving skills may not see these as skills they need to learn/practice and thus may not engage in these supports.

In addition, community recovery capital in the form of 12-Step meeting attendance, financial recovery capital in the form of parental social position, and social recovery capital, in the form of higher substance-approving peers and lower neighborhood social connection, were also strong predictors of RHS attendance. Finally, individual-level factors such as days of alcohol, marijuana and other drug use were predictive of school choice, although perhaps only important in predicting RHS attendance when interactions with other more important variables (e.g., problem solving) are included.

Risk factors at the individual or community level that predicted RHS school attendance, such as higher levels of substance use or lower neighborhood social connection, indicated that some youth may choose to attend an RHS if they believe these schools may address specific recovery risk factors in their lives. RHSs may indeed be one community recovery capital support that arose to target the needs of youth with certain areas of low recovery capital. Supporting this hypothesis is the fact that higher 12-Step meeting attendance was predictive in the final model of overall recovery capital of RHS attendance, indicating either greater severity of dependence and/or recognition that multiple supports were needed to successfully maintain recovery. Additionally, there was a small but consistent group of youth in the individual domain models (i.e., financial, social, and community recovery capital) for both the SEARCH and classification tree analyses who had more than four instances of previous treatment and were more likely to be RHS students. Some adolescents may be more likely to engage in multiple community-level recovery supports given previous experiences in these or similar settings, while others may engage in fewer recovery-specific supports yet still be engaging in other positive social settings. For example, research suggests that although 12-Step attendance is beneficial to adolescents (Chi et al., 2009; Hennessy & Fisher, 2015; Kelly, Myers, & Brown, 2000; Kelly, Brown, Abrantes, Kahler, & Myers, 2008; Kennedy & Minami, 1993), adolescent participation is quite low due to varying hypothesized reasons (e.g., substance use severity or religious/spiritual affiliation: Kelly, Dow, Yeterian, & Kahler, 2010; Kelly, Myers, & Brown, 2002; Kelly, Pagano, Stout, & Johnson, 2011). William White argued that “those with the most enmeshed styles of involvement in a culture of addiction may require an equally enmeshed style of involvement in a culture of recovery to successfully avoid relapse and readdiction” (White, 2009, p. 150). Thus, a culture of recovery is fostered in the community through accessible recovery supports and RHSs appear to fit this need for certain populations of youth in recovery.

The logistic regression, classification tree, and random forest methods largely agreed in their identification of variables important in predicting RHS attendance, although variable importance did vary by method. For example, variables identified as significant in the logistic regression models were often identified as important in either the single classification tree or the random forest; however, neither interactions nor potential non-linear variable relationships were included in the logistic regression models. Thus, some of these finer variable relationship nuances were not identified and could have altered the overall model fit and structure. In addition, the variables that the SEARCH method identified, although it only identified one or two variables from each domain, were also consistent with these methods. This supports the finding that these particular variable interactions are critical for understanding RHS attendance in this sample. The difference in model complexity between SEARCH and the other methods is possibly due to the need to reduce the complexity of the dataset by making most continuous predictors categorical as well as the small sample size of this particular dataset.

Single classification trees had slightly better prediction rates than the random forests, but results from the single classification trees may not be generalizable to other samples. Single trees tend to overfit the data, thus building a less parsimonious model (Breiman et al., 1994; Breiman, 2001). Although pruning addresses this issue, if the initial roots of the tree are misspecified, the resulting model will suffer in accuracy. Indeed the exact results from each single tree were never completely matched in the corresponding forest. The single trees were, however, useful for getting a general sense of the structure of variable relationships in this particular sample, while the forests, due to the aggregation of results across many trees, instead provided variable importance measures that are useful in considering generalizability to other samples. Thus, factors identified as important in both methods should be considered for further attention in future research with other samples, while those identified in just the single trees should be cautiously interpreted and considered as resulting from within-sample variation. Given these nuances, those who employ exploratory methods, especially with smaller samples where splitting data into separate training and test datasets is not possible, may want to consider using the two exploratory methods in combination.

Limitations

The small sample size used in this study is a limitation since the data mining techniques employed here are often used with larger samples where the data are split into a training and test data set. This is not the first study to use these methods with such a small sample (e.g., see Scott, Jackson, & Bergeman, 2011), however, and thus may provide insight for researchers with similar size samples of real-world social science issues. The use of 10-fold cross validation and pruning in the classification tree approach was one way to address this limitation, which permitted the building of a more parsimonious model, taking relative error into account. The models were also better at predicting non-RHS attendance, indicating a great deal of heterogeneity among the RHS student population. Also, given that youth who did not attend an RHS attended many different types of schools, there may be some additional important differences among non-RHS students that could not be explored in these models. For example, the dichotomous outcome of two school groups might be more appropriately conceptualized as three or four, with youth attending traditional, alternative, private/charter specialty, or RHS schools. Thus, future research should explore these covariates with larger samples and account for a greater diversity of school type.

Implications

Although factors relevant to domains of recovery capital have been explored among adolescents, such as mental health comorbidities, spirituality, socioeconomic status, family functioning, and social pressure (e.g., Brown, Myers, Mott, & Vik, 1994; Chi, Kaskutas, Sterling, Campbell, & Weisner, 2009; Ramo & Brown, 2008; Ramo et al., 2012; Ritt-Olson, Milam, Unger, Trinidad, Teran, Dent, & Sussman, 2004; Rohde, Waldron, Turner, Brody, & Jorgensen, 2014; Winters, Stinchfield, Latimer, & Lee, 2007; Yu, Buka, Fitzmaurice, & McCormick, 2006), this study presents one of the first to explore predictors of one highly relevant community recovery capital support, recovery high schools. Using several non-parametric, exploratory methods enabled a more complex picture of recovery supports to emerge, such that interactions that may not have been hypothesized a priori could be identified.

This is one of the first empirical studies to address adolescent recovery within the broader framework of recovery capital, which focuses on multiple domains of resources. Separating analyses by recovery capital domains highlighted important variables and their interactions, which provided a more nuanced understanding of key domain-specific factors among this sample of recovering adolescents. Given that two of the problem solving factor scores, rational problem solving style and positive problem orientation were highly predictive of RHS attendance and that they are based on individual orientation and skills that would affect decisions in the social environment, their interaction with other predictors identified via exploratory methods are an interesting area to explore in future research. In addition, when viewing the individual models, no single recovery capital domain completely predicted RHS attendance, although the human recovery capital model came the closest. Thus, future research must consider the multiple interacting factors that affect adolescent recovery processes and the recovery capital model, an ecological framework that addresses multiple and multi-levels of domains, appears a useful tool for continued research on this issue.

References

American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental Disorders - Fourth Edition (DSM-IV)*. Washington D.C.: American Psychiatric Association.

Ary, D. V., Tildesley, E., Hops, H., & Andrews, J. (1993). The influence of parent, sibling, and peer modeling and attitudes on adolescent use of alcohol. *The International Journal of the Addictions*, *28*(9), 853-880.

Association of Recovery High Schools. (2016). *State of recovery high schools, 2016 biennial report*. Denton, TX: Association of Recovery High Schools.

Becker, S. J., Stein, G. L., Curry, J. F., & Hersh, J. (2012). Ethnic differences among substance-abusing adolescents in a treatment dissemination project. *Journal of Substance Abuse Treatment*, *42*(3), 328-36. doi:10.1016/j.jsat.2011.08.007

Best, D., Gow, J., Taylor, A., Knox, A., & White, W. (2011). Recovery from heroin or alcohol dependence: A qualitative account of the recovery experience in Glasgow. *Journal of Drug Issues, 41*(3), 359-378.

Best, D., Gow, J., Knox, T., Taylor, A., Groshkova, T., & White, W. (2012). Mapping the recovery stories of drinkers and drug users in glasgow: Quality of life and its associations with measures of recovery capital. *Drug and Alcohol Review*, *31*(3), 334-41. doi:10.1111/j.1465-3362.2011.00321.x

Blankers, M., Koeter, M. W., & Schippers, G. M. (2013). Baseline predictors of treatment outcome in internet-based alcohol interventions: A recursive partitioning analysis alongside a randomized trial. *BMC Public Health*, *13*, 455. doi:10.1186/1471-2458-13-455.

Breiman, L. (2001). Random forests. *Machine Learning, 45,* 5-32. doi:10.1023/A:1010933404324

Breiman, L. (1996). Bagging predictors. *Machine Learning*, *24*(2), 123-140. doi:10.1007/BF00058655

Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classification and regression trees*. Boca Raton, Florida: CRC press.

Brown, S. A., D'Amico, E. J., McCarthy, D. M., & Tapert, S. F. (2001). Four-year outcomes from adolescent alcohol and drug treatment. *Journal of Studies on Alcohol*, *62*(3), 381-388.

Brown, S. A., Myers, M. G., Mott, M. A., & Vik, P. W. (1994). Correlates of success following treatment for adolescent substance abuse. *Applied & Preventive Psychology, 3,* 61-73.

Brown, S. A., & Ramo, D. E. (2006). Clinical course of youth following treatment for alcohol and drug problems. In H. A. Liddle & C. L. Rowe (Eds.), *Adolescent substance abuse. Research and clinical advances* (pp. 79-103). New York: Cambridge University Press.

Burns, J., & Marks, D. (2013). Can recovery capital predict addiction problem severity? *Alcoholism Treatment Quarterly, 31*(3), 303-320.

Cattell, R. B. (1966). Psychological theory and scientific method. In R. B. Cattell (Ed.), *Handbook of multivariate experimental psychology* (pp. 1-18). Chicago: Rand McNally.

Cerdá, M., Wall, M., Keyes, K. M., Galea, S., & Hasin, D. (2012). Medical marijuana laws in 50 states: Investigating the relationship between state legalization of medical marijuana and marijuana use, abuse and dependence. *Drug and Alcohol Dependence*, *120*(1), 22-27. doi:10.1016/j.drugalcdep.2011.06.011

Centers for Disease Control and Prevention. [Youth Risk Behavior Surveillance- United States, 2011]. MMWR 2012; *61*(4): 24. Available from: http://www.cdc.gov/mmwr/pdf/ss/ss6104.pdf

Chi, F. W., Kaskutas, L. A., Sterling, S., Campbell, C. I., & Weisner, C. (2009). Twelve-Step affiliation and 3-year substance use outcomes among adolescents: Social support and religious service attendance as potential mediators. *Addiction*, *104*(6), 927-939. doi:doi:10.1111/j.1360-0443.2009.02524.x

DeCato, L. A., Donohue, B., Azrin, N. A., & Teichner, G. A. (2001). Satisfaction of conduct disordered and substance abusing youth with their parents. *Behavior Modification, 25*(1),44-52.

Dennis, M.L. (2010). Global Appraisal of Individual Needs (GAIN): A Standardized Biopsychosocial Assessment Tool. Bloomington, IL: Chestnut Health Systems.

Dierker, L. C., Avenevoli, S., Goldberg, A., & Glantz, M. (2004). Defining subgroups of adolescents at risk for experimental and regular smoking. *Prevention Science*, *5*(3), 169-183.

Doyle, S. R., & Donovan, D. M. (2014). Applying an ensemble classification tree approach to the prediction of completion of a 12-step facilitation intervention with stimulant abusers. *Psychology of Addictive Behaviors, 28*(4), 1127-43. doi:10.1037/a0037235

Duffy, P., & Baldwin, H. (2013). Recovery post treatment: Plans, barriers and motivators. *Substance Abuse Treatment, Prevention, & Policy*, *8*, 6. doi:10.1186/1747-597X-8-6

D'Zurilla, T. J., & Nezu, A. M. (1990). Development and preliminary evaluation of the Social Problem-Solving Inventory. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 2,* 156-163.

D'Zurilla T, Nezu A, & Maydeu-Olivares A. (2002) *Social Problem-Solving Inventory-Revised (SPSI-R): Manual*. North Tonawanda, NY: Multi-Health Systems.

Finch, A. J., & Frieden, G. (2014). The ecological and developmental role of recovery high schools. *Peabody Journal of Education*, *89*(2), 271-287. doi:10.1080/0161956X.2014.897106

Finch, A., Karakos, H., & Hennessy, E. A. (2016). Exploring the policy context around Recovery High Schools: A case study of schools in Minnesota and Massachusetts. In T. Reid (Ed.), *Substance Abuse: Influences, Treatment Options and Health Effects* (pp. 109-138). New York: Nova Science Publishers.

Frick, P. J., & Wootton, J. (1996). The assessment of parenting practices in families of elementary school-aged children. *Journal of Clinical Child Psychology, 25,* 317-327.

Gonzales, R., Anglin, M. D., Beattie, R., Ong, C. A., & Glik, D. C. (2012). Understanding recovery barriers: Youth perceptions about substance use relapse. *American Journal of Health Behavior*, *36*(5), 602-14. doi:10.5993/AJHB.36.5.3

Granfield, R., & Cloud, W. (1999). *Coming clean: Overcoming addiction without treatment.* New York: New York University Press.

Granfield, R., & Cloud, W. (2001). Social context and "natural recovery": The role of social capital in the resolution of drug-associated problems. *Substance Use & Misuse, 36*(11), 1543-1570.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). Random Forests (p. 587-604). In J. Friedman (Ed.). *The elements of statistical learning* (2nd). New York, New York: Springer.

Haug, M. R., & Sussman, M. B. (1971). The indiscriminate state of social class measurement. *Social Forces*, *49*(4), 549-563.

Hennessy, E. A., & Fisher, B. W. (2015). A meta-analysis exploring the relationship between 12-Step attendance and adolescent substance use relapse. *Journal of Groups in Addiction and Recovery, 10,* 79-96.

Johnston, L. D., O'Malley, P. M., Bachman, J. G., & Schulenberg, J. E. (2011). *Monitoring the Future national results on adolescent drug use: Overview of key findings, 2010*. Ann Arbor: Institute for Social Research, The University of Michigan.

Karakos, H. L. (2014). Positive peer support or negative peer influence? The role of peers among adolescents in recovery high schools. *Peabody Journal of Education*, *89*(2), 214-228. doi:10.1080/0161956X.2014.897094

Keane, M. (2011). *The role of education in developing recovery capital in recovery from substance addiction*. Dublin: Soilse Drug Rehabilitation Project. Archived by WebCite® at http://www.webcitation.org/6UZxzeZpo

Kelly, J. F., Brown, S. A., Abrantes, A., Kahler, C. W., & Myers, M. (2008). Social recovery model: An 8-year investigation of adolescent 12-step group involvement following inpatient treatment. *Alcoholism, Clinical and Experimental Research*, *32*(8), 1468-78. doi:10.1111/j.1530-0277.2008.00712.x

Kelly, J. F., Dow, S. J., Yeterian, J. D., & Kahler, C. W. (2010). Can 12-step group participation strengthen and extend the benefits of adolescent addiction treatment? A prospective analysis. *Drug and Alcohol Dependence*, *110*(1–2), 117 - 125. doi:10.1016/j.drugalcdep.2010.02.019

Kelly, J. F., Pagano, M. E., Stout, R. L., & Johnson, S. M. (2011). Influence of religiosity on 12-step participation and treatment response among substance-dependent adolescents. *Journal of Studies on Alcohol and Drugs*, *72*(6), 1000-1011.

Kelly, J. F., Myers, M. G., & Brown, S. A. (2000). A multivariate process model of adolescent 12-step attendance and substance use outcome following inpatient treatment. *Psychology of Addictive Behaviors*, *14*(4), 376-389.

Kelly, J. F., Myers, M. G., & Brown, S. A. (2002). Do adolescents affiliate with 12-step groups? A multivariate process model of effects. *Journal of Studies on Alcohol*, *63*(3), 293-304.

Kennedy, B. P., & Minami, M. (1993). The Beech Hill hospital/outward bound adolescent chemical dependency treatment program. *Journal of Substance Abuse Treatment, 10*, 395-406.

Kochanek, T. T. (2008). *Recovery high schools in Massachusetts: A promising, comprehensive model for adolescent substance abuse and dependence*. Retrieved from http://massrecoveryhs.org/documents/RecoveryHighSchooloverview.pdf

Lemon, S. C., Roy, J., Clark, M. A., Friedmann, P. D., & Rakowski, W. (2003). Classification and regression tree analysis in public health: Methodological review and comparison with logistic regression. *Annals of Behavioral Medicine*, *26*(3), 172-181.

Lewis, T. F., & Mobley, A. K. (2010). Substance abuse and dependency risk: The role of peer perceptions, marijuana involvement, and attitudes toward substance use among college students. *Journal of Drug Education*, *40*(3), 299-314.

Liaw, A., & Wiener, M. (2015). Breiman and Cutler's random forests for classification and regression [Version 4.6.-12] (Software). Available from https://cran.r-project.org/web/packages/randomForest/randomForest.pdf

Lim, T. – S., Loh, W. – Y., & Shih, Y. – S. (2000). A comparison of prediction accuracy, complexity, and training time of thirty-three old and new classification algorithms. *Machine Learning, 40*, 203-229.

Mason, M. J., Mennis, J., Linker, J., Bares, C., & Zaharakis, N. (2014). Peer attitudes effects on adolescent substance use: The moderating role of race and gender. *Prevention Science*, *15*(1), 56-64. doi:10.1007/s11121-012-0353-7

McArdle, J. J. (2014). Exploratory data mining using decision trees in the behavioral sciences. In J. J. McArdle & G. Ritschard (Eds.), *Quantitative Methodology Series: Contemporary issues in exploratory mining in the behavioral sciences* (pp. 1-47). New York, NY: Routledge.

Moberg, D. P., & Finch, A. J. (2007). Recovery high schools: A descriptive study of school programs and students. *Journal of Groups in Addiction & Recovery*, *2*, 128-161. doi:10.1080/15560350802081314

Morgan, J. N. (2005). History and potential of binary segmentation for exploratory data analysis. *Journal of Data Science*, *3*(2), 123-136.

Morgan, J. N., Solenberger, P. W., & Nagara, P. R. (2013). History and future of binary segmentation programs. Retrieved from http://www.isr.umich.edu/src/smp/search/search_paper.html

Morgan, J. N., & Sonquist, J. A. (1963). Problems in the analysis of survey data, and a proposal. *Journal of the American Statistical Association*, *58*(302), 415-434.

National Alcohol Beverage Control Association. (2014). *Grocery stores as alcohol outlets*. Alexandria, VA: NABCA Research. Retrieved from http://www.nabca.org/assets/Docs/Research/White%20Papers/Grocery%20Stores%20Paper%20Sept%202014.pdf

National Institute on Drug Abuse (NIDA). (2014, January). Principles of adolescent substance use disorder treatment: A research-based guide. Retrieved from http://www.drugabuse.gov/sites/default/files/podata_1_17_14.pdf

Neale, J., Nettleton, S., & Pickering, L. (2014). Gender sameness and difference in recovery from heroin dependence: A qualitative exploration. *International Journal of Drug Policy*, *25*(1), 3-12. doi:10.1016/j.drugpo.2013.08.002

Neale, J., & Stevenson, C. (2014). Social and recovery capital amongst homeless hostel residents who use drugs and alcohol. *International Journal of Drug Policy*, in press. http://dx.doi.org/10.1016/j.drugpo.2014.09012

Neighbors, C., Dillard, A. J., Lewis, M. A., Bergstrom, R. L., & Neil, T. A. (2006). Normative misperceptions and temporal precedence of perceived norms and drinking. *Journal of Studies on Alcohol*, *67*(2), 290-9.

Office of National Drug Control Policy (ONDCP). (2014). *National drug control strategy*. Washington, D.C.: Office of National Drug Control Policy. Retrieved from http://www.whitehouse.gov/sites/default/files/ondcp/policy-and-research/ndcs_2014.pdf

Oser, R., Karakos, H. L., & Hennessy, E. A. (2016). Disparities in youth access to substance abuse treatment and recovery services: How one recovery school initiative is helping students "Change Tracks". *Journal of Groups in Addiction and Recovery, 11*(4), 267-281. DOI:10.1080/1556035X.2016.1211056

Piper, D. L., Moberg, D. P., & King, M. J. (2000). The Healthy for Life project: Behavioral outcomes. *Journal of Primary Prevention 21(1)*, 47-73.

Ramo, D. E., & Brown, S. A. (2008). Classes of substance abuse relapse situations: A comparison of adolescents and adults. *Psychology of Addictive Behaviors*, *22*(3), 372-9. doi:10.1037/0893-164X.22.3.372

Ramo, D. E., Prince, M. A., Roesch, S. C., & Brown, S. A. (2012). Variation in substance use relapse episodes among adolescents: A longitudinal investigation. *Journal of Substance Abuse Treatment*, *43*(1), 44-52. doi:10.1016/j.jsat.2011.10.003

Reynolds, C. R., & Kamphaus, R. W. (1992). *Behavior Assessment System for Children: Manual*. Circle Pines, MN: American Guidance.

Ritt-Olson, A., Milam, J., Unger, J. B., Trinidad, D., Teran, L., Dent, C. W., & Sussman, S. (2004). The protective influence of spirituality and "health-as-a-value" against monthly substance use among adolescents varying in risk. *The Journal of Adolescent Health*, *34*(3), 192-199. doi:10.1016/j.jadohealth.2003.07.009

Rohde, P., Waldron, H. B., Turner, C. W., Brody, J., & Jorgensen, J. (2014). Sequenced versus coordinated treatment for adolescents with comorbid depressive and substance use disorders. *Journal of Consulting and Clinical Psychology*, *82*(2), 342-8. doi:10.1037/a003580

Scott, S. B., Whitehead, B. R., Bergeman, C. S., & Pitzer, L. (2014). Understanding global perceptions of stress in adulthood through tree-based exploratory data mining. In J. J. McArdle & G. Ritschard (Eds.), *Quantitative Methodology Series: Contemporary issues in exploratory mining in the behavioral sciences* (pp. 371-404). New York, NY: Routledge.

Sheehan, D. V., Janavs, R., Baker, R., Harnett- Sheehan, K., Knapp, E., & Sheehan, M. (1999). *Mini International Neuropsychiatric Interview.* Tampa: University of South Florida Press.

Sobell, L.C., & Sobell, M.B., (1992). Timeline Followback: A technique for assessing self-reported alcohol consumption. In R.Z. Litten & J. Allen (Eds.), *Measuring alcohol consumption: Psychosocial and biological methods* (pp. 41-72). New Jersey: Humana Press.

Sonquist, J. A., Baker, E. L., & Morgan, J. N. (1974). *Searching for structure*. Michigan: The University of Michigan.

Sonquist, J. A., & Morgan, J. N. (1964). *The detection of interaction effects*. Michigan: Survey Research Center, University of Michigan.

Spear, S. F., Ciesla, J. R., & Skala, S. Y. (1999). Relapse patterns among adolescents treated for chemical dependency. *Substance Use & Misuse*, *34*(13), 1795-1815.

StataCorp (2015). *Stata Statistical Software: Release 14*. College Station, TX: StataCorp LP.

Sterling, S., Chi, F., Campbell, C., & Weisner, C. (2009). Three-year chemical dependency and mental health treatment outcomes among adolescents: The role of continuing care. *Alcoholism, Clinical and Experimental Research*, *33*(8), 1417-29. doi:10.1111/j.1530-0277.2009.00972.x

Stevens, S. J., Estrada, B., Murphy, B. S., McKnight, K. M., & Tims, F. (2004). Gender difference in substance use, mental health, and criminal justice involvement of adolescents at treatment entry and at three, six, 12 and 30 month follow-up. *Journal of Psychoactive Drugs*, *36*(1), 13-25.

Substance Abuse and Mental Health Services Administration. (2013). *Results from the 2012 National Survey on Drug Use and Health: Summary of National Findings*, NSDUH Series H-46, HHS Publication No. (SMA) 13-4795. Rockville, MD: Substance Abuse and Mental Health Services Administration.

Substance Abuse and Mental Health Services Administration (2014). *Results from the 2013 National Survey on Drug Use and Health: Summary of National Findings*, NSDUH Series H-48, HHS Publication No. (SMA) 14-4863. Rockville, MD: Substance Abuse and Mental Health Services Administration.

Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality (2016). *Treatment Episode Data Set (TEDS): 2004-2014. National Admissions to Substance Abuse Treatment Services.* BHSIS Series S-84, HHS Publication No. (SMA) 16-4986. Rockville, MD: Substance Abuse and Mental Health Services Administration.

Terrion, J. L. (2012). The experience of post-secondary education for students in recovery from addiction to drugs or alcohol: Relationships and recovery capital. *Journal of Social and Personal Relationships,* 1-21. doi:10.1177/0265407512448276

Terrion, J. L. (2014). Building recovery capital in recovering substance abusers through a service-learning partnership: A qualitative evaluation of a communication skills training program. *Journal of Service-Learning in Higher Education*, *3*, 47-63.

Therneau, T., Atkinson, B, & Ripley, B. (2015). Recursive partitioning and regression trees [Version 4.1.-10] (Software). Available from: https://cran.r-project.org/web/packages/rpart/

Therneau, T. M., Atkinson, E. J., & Mayo Foundation (2015). An introduction to recursive partitioning using the RPART routines. Available from https://cran.r-project.org/web/packages/rpart/vignettes/longintro.pdf

Torgo, L. F. R. A. (1999). Inductive learning of tree-based regression models.

Tukey, J. W. (1962). The future of data analysis. *Annals of Mathematical Statistics*, 33, 1-67.

Waldron, H. B., & Kaminer, Y. (2004). On the learning curve: The emerging evidence supporting cognitive--behavioral therapies for adolescent substance abuse. *Addiction*, *99*(s2), 93-105.

Wambeam, R. A., Canen, E. L., Linkenbach, J., & Otto, J. (2013). Youth misperceptions of peer substance use norms: A hidden risk factor in state and community prevention. *Prevention Science*, *15*(1), 75-84. doi:10.1007/s11121-013-0384-8

Wellman, R. J., Contreras, G. A., Dugas, E. N., O'Loughlin, E. K., & O'Loughlin, J. L. (2014). Determinants of sustained binge drinking in young adults. *Alcoholism, Clinical and Experimental Research*, *38*(5), 1409-15. doi:10.1111/acer.12365

White, W. L. (2009). The mobilization of community resources to support long-term addiction recovery. *Journal of Substance Abuse Treatment*, *36*(2), 146-58. doi:10.1016/j.jsat.2008.10.006

White, W., & Cloud, W. (2008). Recovery capital: A primer for addictions professionals. *Counselor*, *9*(5), 22-27.

White, W. L., & Finch, A. J. (2006). The recovery school movement: Its history and future. *Counselor*, *7*(2), 54-57. Retrieved from www.bhrm.org

White, A. M., Jordan, J. D., Schroeder, K. M., Acheson, S. K., Georgi, B. D., Sauls, G., . . . Swartzwelder, H. S. (2004). Predictors of relapse during treatment and treatment completion among marijuana-dependent adolescents in an intensive outpatient substance abuse program. *Substance Abuse*, *25*(1), 53-9. doi:10.1300/J465v25n01_08

Wills, T. A., Sandy, J. M., Yaeger, A. M., Cleary, S. D., & Shinar, O. (2001). Coping dimensions, life stress, and adolescent substance use: A latent growth analysis. *Journal of Abnormal Psychology*, *110*(2), 309.

Winters, K. & Henly, G. (1993). *Adolescent Diagnostic Interview (ADI) Manual*. Los Angeles, CA: Western Psychological Services.

Winters, K.C. & Henly, G.A. (1989). *Personal Experience Inventory (PEI) Test and Manual*. Los Angeles: Western Psychological Services.

Winters, K. C., Stinchfield, R., Latimer, W. W., & Lee, S. (2007). Long-term outcome of substance-dependent youth following 12-step treatment. *Journal of Substance Abuse Treatment*, *33*(1), 61-9. doi:10.1016/j.jsat.2006.12.003

Winters, K. C., Tanner-Smith, E. E., Bresani, E., & Meyers, K. (2014). Current advances in the treatment of adolescent drug use. *Adolescent Health, Medicine, and Therapeutics*, *5*, 199-210. doi:10.2147/AHMT.S48053

Yu, J. W., Buka, S. L., Fitzmaurice, G. M., & McCormick, M. C. (2006). Treatment outcomes for substance abuse among adolescents with learning disorders. *The Journal of Behavioral Health Services & Research*, *33*(3), 275-286. doi:10.1007/s11414-006-9023-5

Table 1. Recovery Capital Variables – Baseline Measurements by Group

	RHS M / n	RHS SD/(%)	nonR M / n	nonR SD/(%)	Total M / n	Total SD/(%)	n
Variables in each model (k = 11)							
Age	16.47	1.00	16.24	1.05	16.34	1.03	260
Age first treated for AOD	15.29	1.20	15.50	1.20	15.40	1.20	259
Race/ethnicity – White[a]	103	(86)	116	(83)	219	(84)	260
Race/ethnicity – Black[a]	9	(8)	13	(9)	22	(8)	260
Race/ethnicity – Hispanic[a]	14	(12)	13	(9)	27	(10)	260
Race/ethnicity – Asian[a]	4	(3)	7	(5)	11	(4)	260
Race/ethnicity - Native American[a]	4	(3)	7	(5)	11	(4)	260
Race/ethnicity – Other[a]	2	(2)	4	(3)	6	(2)	260
Sex (male = 1)	56	(47)	81	(58)	137	(53)	260
Days of alcohol use	18.84	25.36	15.69	22.62	17.15	23.93	260
Days of marijuana use	53.45	35.55	55.88	34.51	54.76	34.94	260
Days of other drug use	30.58	36.17	20.53	28.99	25.17	32.82	260
SUD diagnosis – alcohol	83	(69)	82	(59)	165	(63)	260
SUD diagnosis – other drugs	114	(95)	132	(94)	246	(95)	260
MH service receipt (# times)	6.12	20.32	3.71	3.84	4.84	14.18	255
AOD treatment (# times)	5.13	27.62	1.86	1.35	3.35	18.69	257
Human Capital (k = 19)							
SU expectancies – psychological benefits	21.84	4.28	21.97	3.35	21.91	3.80	260
SU expectancies – social benefits	22.24	4.42	21.47	3.82	21.83	4.12	259
Negative attitudes towards school	0.60	0.26	0.55	0.25	0.57	0.26	258
General (life) satisfaction	3.70	0.60	3.44	0.66	3.56	0.64	260
Physical health	3.62	1.63	3.21	1.77	3.40	1.71	260
School attendance, last 12 months (days)	162.22	38.91	147.34	51.47	154.20	46.61	260
Cumulative GPA (entire year)	2.51	0.88	2.43	0.88	2.47	0.88	41
Stress	3.48	1.76	3.01	1.77	3.22	1.78	260
School problems	12.72	3.17	13.06	3.07	12.91	3.12	255
Crime and violence screener, lifetime	2.47	1.51	2.60	1.48	2.54	1.49	260
Positive problem orientation (F1)	0.04	0.90	-0.03	0.87	0.00	0.88	244
Negative problem orientation (F2)	0.06	0.79	-0.05	0.88	0.00	0.84	244
Rational problem solving (F3)	-0.23	0.84	0.20	0.82	0.00	0.86	244
Impulsivity/carelessness style (F4)	0.10	0.80	-0.08	0.79	0.00	0.80	244
Avoidance style (F5)	-0.05	0.78	0.04	0.78	0.00	0.78	244
DSM-IV diagnosis: major depression	101	(84)	95	(68)	196	(75)	260
DSM-IV diagnosis: manic episode	16	(13)	28	(20)	44	(17)	260
DSM-IV diagnosis: PTSD	46	(38)	30	(21)	76	(29)	260
MH screens: eating disorder	36	(30)	41	(29)	77	(30)	260
Financial Capital (k = 4)							
Family income level	5.50	1.32	5.25	1.46	5.36	1.40	243
Parental social position score	33.37	12.93	36.32	15.10	35.00	14.22	245
Ever homeless	29	(25)	28	(20)	57	(22)	257
Any health insurance	105	(97)	124	(93)	229	(95)	241
Social capital (k = 8)							
General social support index	8.22	1.04	8.16	1.20	8.18	1.12	260
Social competence index	3.10	0.35	3.07	0.34	3.08	0.35	258
Neighborhood social connections index	2.94	0.73	2.97	0.78	2.95	0.75	257

Spiritual social support index	1.23	2.03	1.68	2.26	1.47	2.16	260
Youth-parent relationship	61.38	22.44	57.35	20.59	59.19	21.51	258
Substance approving peer attitudes	3.13	0.57	3.01	0.45	3.06	0.51	257
Immediate family AOD history	84	(70)	95	(68)	179	(69)	260
Sibling AOD history	34	(28)	37	(26)	71	(27)	260
Parent AOD history	77	(64)	87	(62)	164	(63)	260
Immediate family MH history	78	(65)	93	(66)	171	(66)	260
Sibling MH history	28	(23)	41	(29)	69	(27)	260
Parent MH history	74	(62)	86	(61)	160	(62)	260
Community Capital (k = 5)							
Perceived availability of drugs	4.28	0.62	4.19	0.66	4.23	0.64	260
Parent knowledge of RHS prior to TX	36	(33)	54	(42)	90	(38)	239
Youth knowledge of RHS prior to TX	35	(31)	56	(44)	91	(38)	240
AOD/MH counseling outside school	88	(74)	109	(78)	197	(76)	259
AA/NA/12-Step meeting attendance	3.69	1.56	3.10	1.66	3.37	1.64	259

Note. AOD = alcohol and other drug. MH = Mental Health. RHS = Recovery High School. Non-RHS = non Recovery High School. SU = Substance use. TX = treatment.

[a]Ethnic/racial demographic categories are presented for each collected category; however, only a binary variable indicating White/Non-White was used in the analysis.

Table 2. Summary of Logistic Regression, SEARCH, and Classification Tree Recovery Capital Results

	Financial			Human			Social			Community			Overall		
	LR	S	CT	LR	S	CT	LR	S	CT	LR	S	CT	LR	S	CT
Variables included in each unique domain model as control variables (k = 11)															
Age	+														
Age first treated for AOD															
Race/ethnicity – White[a]															
Sex (male = 1)							+			+					
Days of alcohol use															
Days of marijuana use															
Days of other drug use												+			
SUD diagnosis – alcohol															
SUD diagnosis – other drugs															
MH service receipt (# times)						+						+			
AOD treatment (# times)		+	+		+			+	+		+	+		+	
Financial Capital (k = 4)															
Family income level															
Parental social position score															
Ever homeless															
Any health insurance															
Human Capital (k = 19)															
SU expectancies – psychological benefits															
SU expectancies – social benefits						+									
Negative attitudes towards school				+									+		
General (life) satisfaction				+									+		
Physical health						+									
School attendance, last 12 months (days)				+		+									+
Cumulative GPA (entire year)															
Stress				+											
School problems															
Crime and violence screener, lifetime															
Positive problem orientation (F1)						+									+
Negative problem orientation (F2)						+									
Rational problem solving (F3)				+	+	+							+	+	+
Impulsivity/carelessness style (F4)				+											
Avoidance style (F5)				+											
DSM-IV diagnosis: major depression						+							+		
DSM-IV diagnosis: manic episode						+									+
DSM-IV diagnosis: PTSD						+									+
MH screens: eating disorder															
Social capital (k = 8)															
General social support index															

Social competence index															
Neighborhood social connections index									+						
Spiritual social support index															
Youth-parent relationship															
Substance approving peer attitudes									+						
Immediate family AOD history															
Immediate family MH history															
Community Capital (k = 5)															
Perceived availability of drugs												+			
Parent knowledge of RHS prior to TX															
Youth knowledge of RHS prior to TX										+		+	+		
AOD/MH counseling outside school												+			+
AA/NA/12-Step meeting attendance										+		+	+		+

Note. A "+" indicates that the variable was found to be significant in the model: non-significant variables that were included in the model are simply blank cells. Cells in grey indicate that the variable was not used in the model. LR = Logistic Regression model. S = SEARCH model. CT = Classification Tree Model.

Table 3. Results from Logistic Regressions of Recovery Capital

	Financial Recovery Capital (k = 15)			Human Recovery Capital (k = 30)			Social Recovery Capital (k = 19)			Community Recovery Capital (k = 16)			Overall Recovery Capital (k = 11)		
	(N = 230)			(N = 189)			(N = 247)			(N = 212)			(N = 226)		
	OR	95% CI		OR	95% CI		OR	95% CI		OR	95% CI		OR	95% CI	
Age	**1.48**	**[1.02**	**2.17]**	1.16	[0.70	1.94]	1.33	[0.93	1.89]	1.17	[0.80	1.72]	1.08	[0.79	1.49]
Age first treated for AOD	0.75	[0.54	1.03]	0.91	[0.59	1.42]	0.81	[0.59	1.10]	0.86	[0.62	1.20]			
Caucasian	0.96	[0.43	2.15]	1.29	[0.41	4.05]	1.21	[0.54	2.72]	0.68	[0.28	1.63]			
Sex (Male=1)	0.55	[0.30	1.00]	0.51	[0.20	1.28]	**0.51**	**[0.29**	**0.90]**	**0.43**	**[0.22**	**0.84]**	0.60	[0.31	1.16]
Alcohol use	1.00	[0.99	1.02]	1.00	[0.99	1.02]	1.00	[0.99	1.01]	1.00	[0.99	1.02]			
Marijuana use	1.00	[0.99	1.01]	1.00	[0.99	1.01]	1.00	[0.99	1.01]	1.00	[0.99	1.01]			
Other drug use	1.01	[1.00	1.01]	1.00	[0.99	1.01]	1.01	[1.00	1.02]	1.01	[1.00	1.02]			
Alcohol dependence/abuse	1.23	[0.62	2.42]	1.04	[0.42	2.57]	1.15	[0.59	2.25]	1.22	[0.58	2.59]			
Other drug dependence/abuse	0.99	[0.57	1.70]	0.83	[0.37	1.85]	0.94	[0.55	1.61]	0.93	[0.51	1.70]			
No. times received MH services	0.93	[0.85	1.01]	0.92	[0.81	1.04]	0.96	[0.89	1.04]	0.94	[0.85	1.03]			
No. times received AOD TX	1.12	[0.96	1.30]	1.26	[0.98	1.61]	1.11	[0.95	1.30]	1.11	[0.93	1.32]			
Parental income	1.07	[0.86	1.34]												
Parental social position score	0.99	[0.97	1.01]												
Ever homeless	0.96	[0.46	2.00]												
Any health insurance	2.62	[0.61	11.3]												
SU expectancies - psychological				0.98	[0.86	1.12]									
SU expectancies - social				1.06	[0.92	1.21]									
Negative attitudes towards school				**5.60**	**[1.01**	**30.90]**							**8.27**	**[2.24**	**30.53]**
General life satisfaction				**2.77**	**[1.32**	**5.83]**							**2.21**	**[1.28**	**3.81]**
Physical health				0.83	[0.64	1.09]									
Days of school attendance				**1.02**	**[1.00**	**1.03]**							1.01	[1.00	1.01]

Cumulative GPA	1.00	[0.61	1.64]									
Stress	**1.35**	**[1.03**	**1.78]**							1.18	[0.98	1.43]
Parent-reported school problems	1.06	[0.93	1.21]									
Crime and violence (lifetime)	1.07	[0.79	1.46]									
Positive problem orientation (F1)	1.12	[0.69	1.82]									
Negative problem orientation (F2)	1.11	[0.70	1.77]									
Rational problem solving (F3)	**0.53**	**[0.33**	**0.87]**							**0.65**	**[0.44**	**0.96]**
Impulsivity/carelessness style (F4)	**1.74**	**[1.02**	**2.98]**							1.16	[0.78	1.74]
Avoidance style (F5)	0.76	[0.45	1.27]									
Major depression	**3.19**	**[1.15**	**8.88]**							**2.50**	**[1.11**	**5.68]**
Manic episode	0.53	[0.19	1.54]									
PTSD	1.78	[0.70	4.53]									
Positive eating disorder screen	0.72	[0.28	1.86]									
Social support				1.12	[0.85	1.47]						
Social competence				1.16	[0.44	3.10]						
Neighborhood social connection				0.79	[0.54	1.16]						
Spiritual social support				0.93	[0.82	1.06]						
Youth-parent relationship				1.00	[0.99	1.02]						
Substance-approving peers				1.51	[0.86	2.67]						
Immediate family AOD history				0.99	[0.53	1.86]						
Immediate family MH history				0.96	[0.51	1.82]						
Perceived availability of drugs							0.93	[0.54	1.61]			
Parent knowledge of RHS							0.75	[0.39	1.45]			

Youth knowledge of RHS										**0.37**	**[0.19**	**0.73]**	**0.40**	**[0.21**	**0.76]**
AOD/MH counseling outside school										0.73	[0.35	1.54]			
12-step meeting attendance frequency										**1.39**	**[1.11**	**1.74]**	**1.29**	**[1.06**	**1.58]**
Constant	0.06	0.00	9.44	**0.00**	**0.00**	**0.06**	0.02	0.00	5.81	0.89	0.00	242.03	**0.00**	**0.00**	**0.09**
R^2	0.09			0.28			0.09			0.13			0.20		
Chi^2	**28.63**			**71.99**			**30.44**			**37.78**			**63.19**		
p-value	0.02			0.00			0.05			0.00			0.00		

Table 4. Results from Random Forests: Variable Importance Rankings

Financial Recovery Capital		Human Recovery Capital		Social Recovery Capital		Community Recovery Capital		Overall Recovery Capital	
OOB error: 45%		OOB error: 36.92%		OOB error: 45.38%		OOB error: 42.69%		OOB error: 36.54%	
Parental social position	11.29	Rational problem solving (F3)	6.22	Substance approving peers	8.81	Other drug use (days)	10.38	Rational problem solving (F3)	6.97
Other drug use (days)	10.57	School attendance (days)	4.88	Youth-parent relationship	8.65	Alcohol use (days)	8.88	School attendance (days)	5.09
Marijuana use (days)	9.51	Impulsivity/ carelessness (F4)	4.65	Other drug use (days)	7.19	Marijuana use (days)	8.87	Impulsivity/ carelessness (F4)	4.95
Alcohol use (days)	9.00	Positive problem orientation (F1)	4.56	Neighborhood social connections	6.84	Perceived drug availability	8.32	Positive problem orientation (F1)	4.47
MH services (#)	7.76	Negative problem orientation (F2)	4.30	Alcohol use (days)	6.50	MH services (#)	7.42	Substance approving peers	4.42
Family income	6.89	Life satisfaction	4.19	Marijuana use (days)	6.29	12-Step attendance	6.98	Youth-parent relationship	4.37
Age first treated for AOD	5.71	Avoidance style (F5)	4.12	MH services (#)	5.62	Age first treated for AOD	5.53	Negative problem orientation (F2)	4.34
AOD treatment (#)	5.64	Other drug use (days)	3.89	Social competence	5.59	AOD treatment (#)	5.45	Avoidance style (F5)	4.32
Age	4.45	SU expectancies - social benefits	3.34	AOD treatment (#)	4.26	Age	4.44	Parental social position	4.08
Sex	2.32	School problems	3.23	Age first treated for AOD	3.86	Sex	2.58	Other drug use (days)	3.98
SUD - other drugs	1.87	SU expectancies - psychological benefits	3.19	Age	3.64	Parent knowledge of RHS before TX	2.35	Life satisfaction	3.95
SUD - alcohol	1.60	Alcohol use (days)	3.17	Spiritual social support	3.47	Youth knowledge of RHS before TX	2.10	Neighborhood social connections	3.62
Caucasian	1.49	Marijuana use (days)	3.06	General social support	3.04	SUD - other drugs	1.95	12-Step attendance	3.40
Ever homeless	1.33	Cumulative GPA	2.98	Sex	1.57	Caucasian	1.86	Alcohol use	3.40

								(days)	
Any health insurance	0.79	Negative attitudes towards school	2.98	SUD - other drugs	1.48	SUD - alcohol	1.65	Marijuana use (days)	3.17
		MH services (#)	2.67	SUD - alcohol	1.26	AOD/MH counseling outside school	1.50	Perceived drug availability	2.81
		Stress	2.62	Family MH history	1.12			MH services (#)	2.76
		AOD treatment (#)	2.57	Family AOD history	1.06			Family income	2.75
		Physical Health	2.49	Caucasian	1.00			Social competence	2.74
		Crime and violence	2.31					AOD treatment (#)	2.69
		Age first treated for AOD	2.14					Age first treated for AOD	1.94
		Age	1.86					Age	1.66
		Major depression	1.29						
		PTSD	1.19						
		Manic episode	0.77						
		SUD - other drugs	0.76						
		Sex	0.71						
		SUD - alcohol	0.58						
		Caucasian	0.50						
		Eating disorder	0.45						

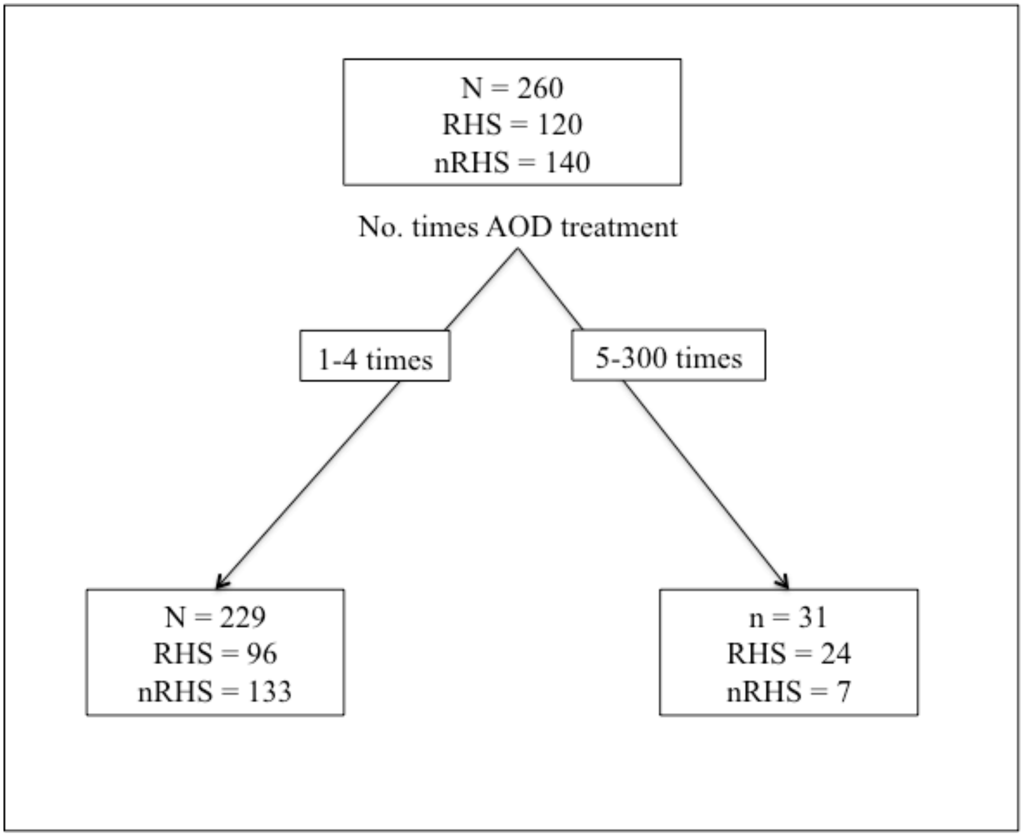

Figure 1. SEARCH Tree representing Financial, Social, and Community Recovery Capital results

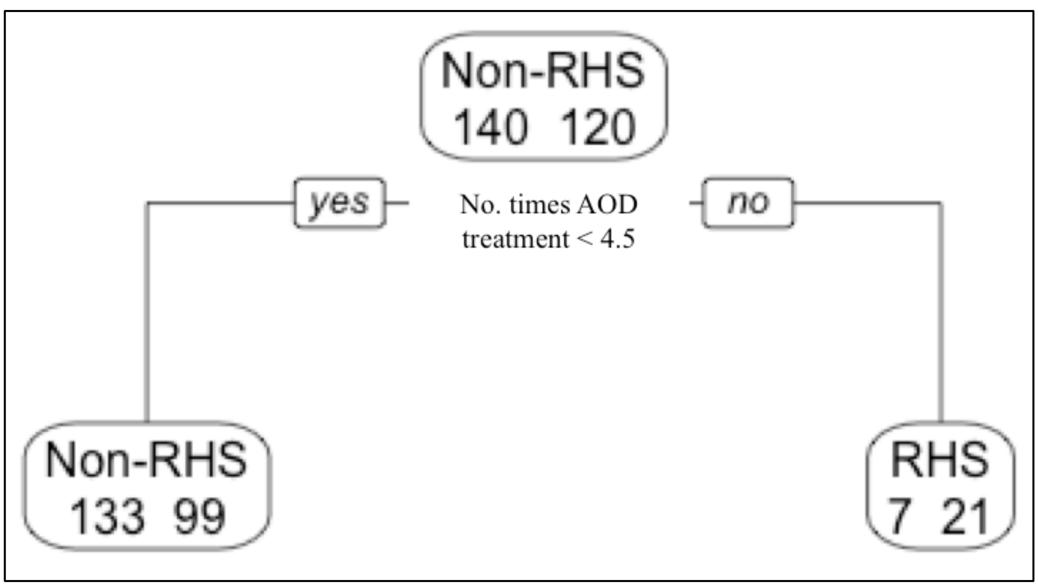

Figure 2. Financial Recovery Capital Classification Tree

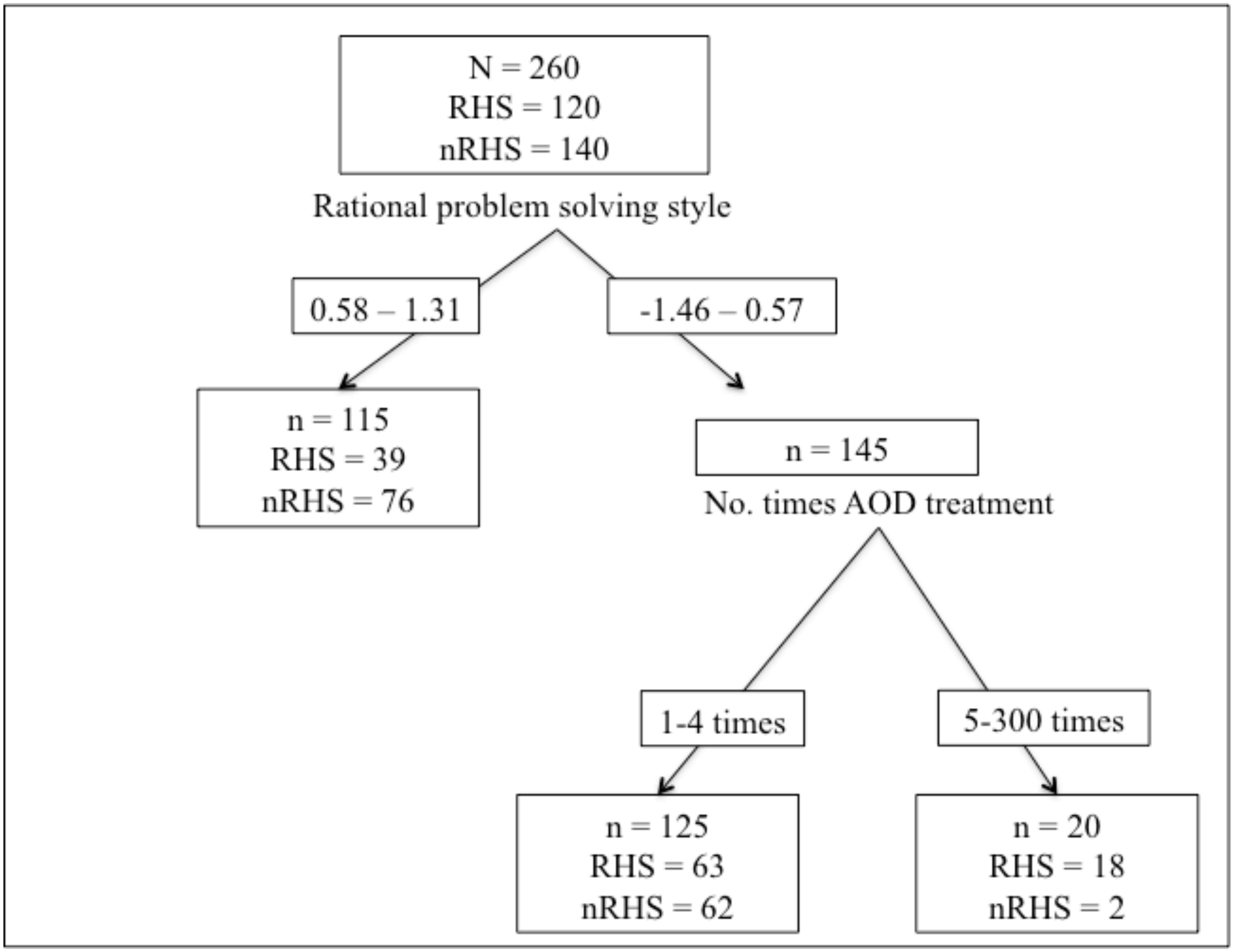

Figure 3. SEARCH Tree representing Human and Overall Recovery Capital results

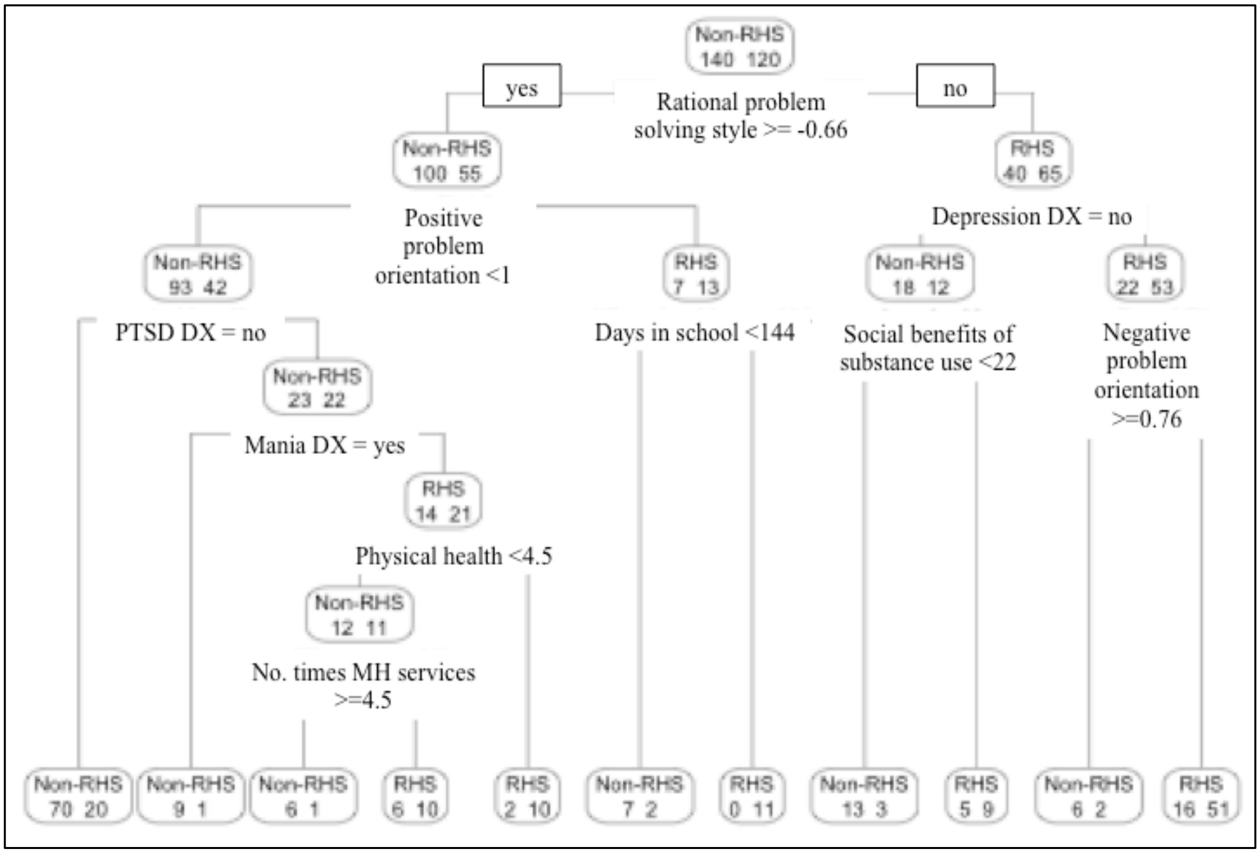

Figure 4. Human Recovery Capital Classification Tree

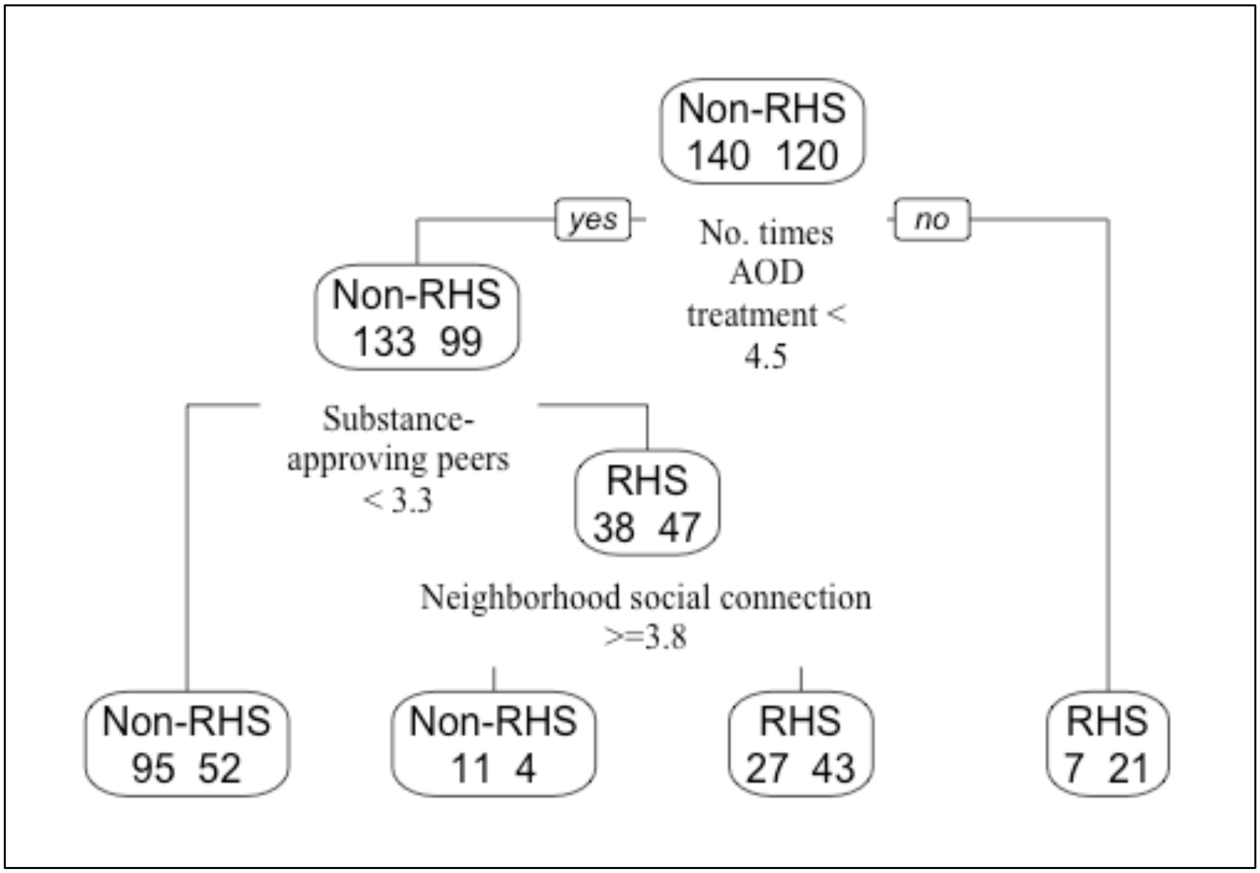

Figure 5. Social Recovery Capital Classification Tree

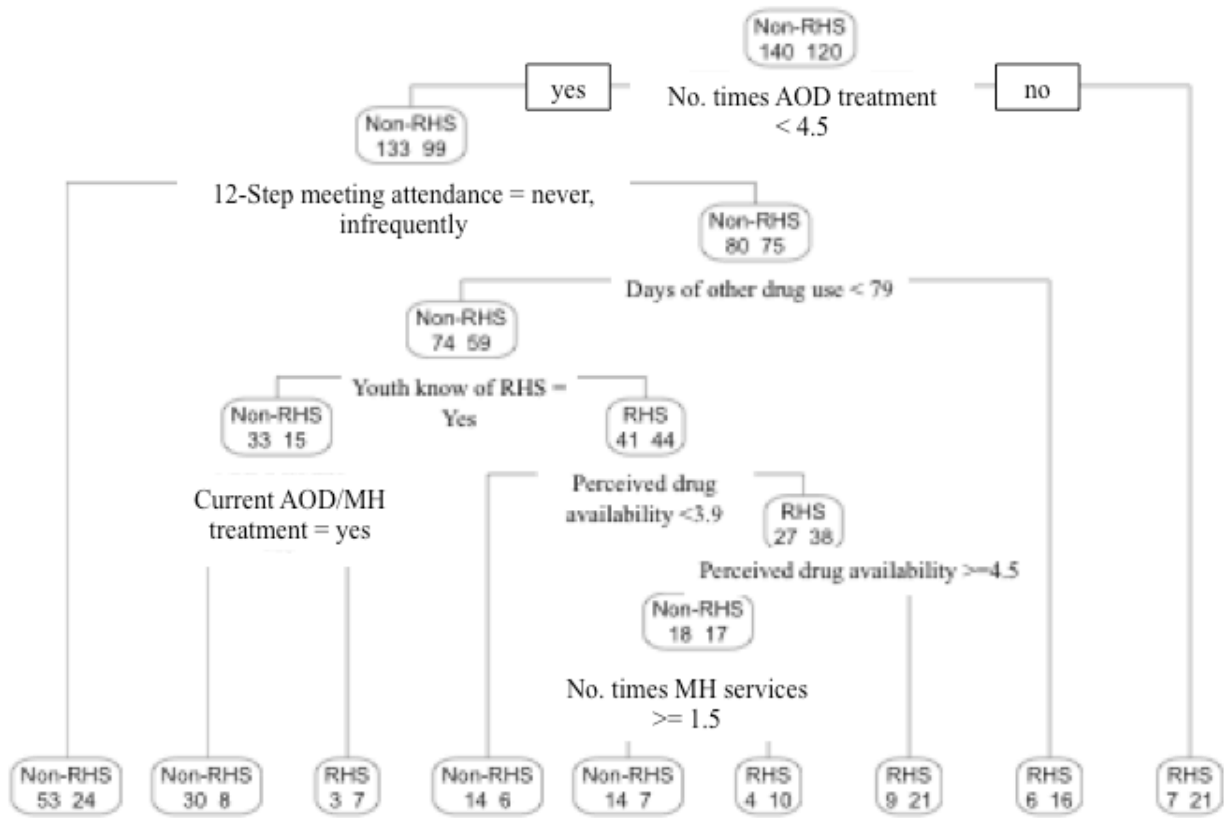

Figure 6. Community Recovery Capital Classification Tree

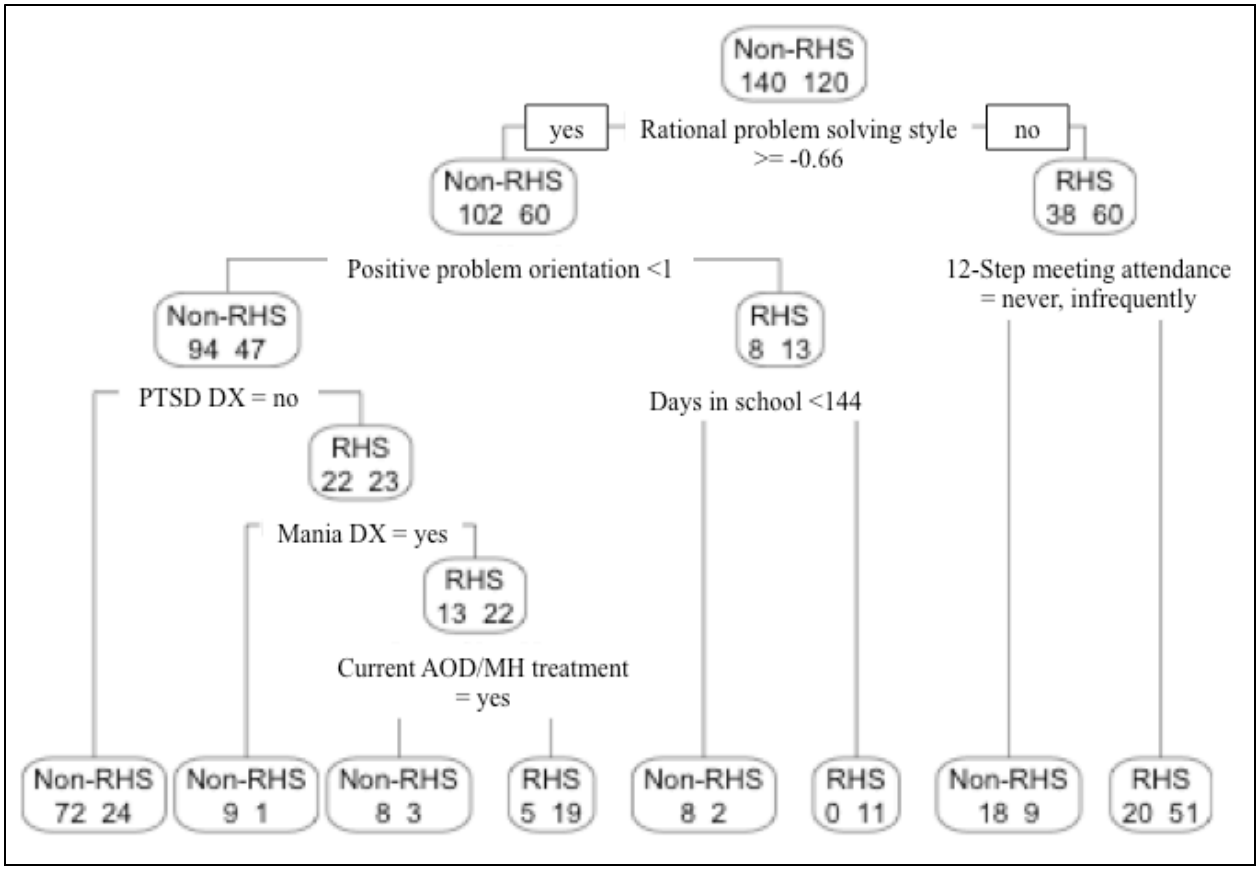

Figure 7. Overall Recovery Capital Classification Tree

Appendix A

Table 1. Variables and measurement tools organized by Recovery Capital domain

Variable Name	Scale	Range	Scale Properties
Variables added to each model			
Age	ADI (Winters & Henly, 1993)		
Race/ethnicity	ADI (Winters & Henly, 1993)	0 - 1	Non-exclusive categories
Sex	ADI (Winters & Henly, 1993)	0 - 1	
Age first treated for AOD	ADI (Winters & Henly, 1993)		
AOD treatment (# times)	ADI (Winters & Henly, 1993)		
MH service receipt (# times)	ADI (Winters & Henly, 1993)		
Days of alcohol use	TLFB (Sobell & Sobell, 1992)		
Days of marijuana use	TLFB (Sobell & Sobell, 1992)		
Days of other drug use	TLFB (Sobell & Sobell, 1992)		
SUD abuse or dependence diagnosis - alcohol	MINI-SCID (Sheehan et al., 1999)	0 - 1	
SUD abuse or dependence diagnosis - other drugs	MINI-SCID (Sheehan et al., 1999)	0 - 1	
Human Capital			
Substance use expectancies - psychological benefits	PEI (Winters & Henly, 1989), modified	0 - 28	7 item scale (1 = "strongly disagree", 4 = "strongly agree")
Substance use expectancies - social benefits	PEI (Winters & Henly, 1989)	0 - 28	7 item scale (1 = "strongly disagree", 4 = "strongly agree")
Negative attitudes towards school	BASC (Reynolds & Kamphaus, 1992), modified	0 - 1	10 item scale, average of all items
General (life) satisfaction	GAIN-Q3, LSI (Dennis, 2010), modified	1 - 5	6-item scale (1 = "very dissatisfied", 5 = "very satisfied")
Physical health	GAIN-Q3 (Dennis, 2010)	0 - 6	6 item, number of problems (count)
School attendance, last 12 months (days)	Services timeline followback, created from TLFB		
Cumulative GPA (entire year)	HSQ (Moberg & Finch, 2008)	0 - 4	
Sources of stress screener, lifetime	GAIN-Q3: SSScrL (Dennis, 2010)	0 - 8	8 item scale (rescaled so that 0 = "never", 4 = past month")
School problems	GAIN-Q3: School Problem Screener Lifetime (SPScrL; Dennis, 2010), modified	5 - 20	4 item summary (1 = "never", 5 = past month")
Crime and violence screener, lifetime	GAIN-Q3 (Dennis, 2010)	0 - 5	5 items, (rescaled to that 0 = "never", 4 = past month")
Social problem solving inventory	SPSI-R SF (D'Zurilla & Nezu, 1990)		5-factor, 25 item scale (D'Zurilla et al., 2002): each factor included separately in analysis
DSM-IV MH diagnoses	MINI-SCID (Sheehan et al., 1999)	0 - 1	Each diagnosis included separately in analysis
MH screen – eating disorder	MINI-SCID (Sheehan et al., 1999)	0 - 1	
Financial Capital			
Family income level	Single item	1 - 7	1 = <$5,000; 2 = >$5,000 - $10,000; 3 = >$10,000 - $25,000; 4 = >$25,000 - $40,000; 5 = >$40,000 - $75,000; 6 = >$75,000 - $100,000; 7 =

			>$100,000
Parental social position score	Social position index (Haug & Sussman, 1970)		Combination of parent participant and spouse education and occupation
Ever homeless	GAIN-Q3, single item (Dennis, 2010)	1 - 5	1 = "never", 5 = "past month"
Any health insurance	Single item	0 - 1	Measured at 3 months, assumed no changes since baseline
		Social Capital	
General social support index	GAIN - GSS (Dennis, 2010)	0 - 9	9 item scale (yes/no)
Social competence index	HSQ (Moberg & Finch, 2008)		8 item scale (higher indicates higher social competence)
Neighborhood social connections index	PSL-Adapted (Leffert et al., 1998)		6 item scale (higher indicates higher ratings of neighborhood social connections)
Spiritual social support index	GAIN - SSSI (Dennis, 2010)	0 - 7	7 item scale (yes/no)
Youth-parent relationship	YHPS (De Cato et al., 2001)	0 - 100	11 item index (0 = "completely unhappy", 100 = "completely happy")
Substance approving peer attitudes	PEI (Winters & Henly, 1989)		13 item scale (1 = "strongly disagree", 4 = "strongly agree")
Immediate family AOD history	ADI (Winters & Henly, 1993)	0 - 1	Parent and sibling AOD history combined for a family AOD history
Immediate family MH history	ADI (Winters & Henly, 1993)	0 - 1	Parent and sibling MH history combined for a family AOD history
		Community Capital	
Perceived Availability of Drugs	MTF (Johnson et al., 2011), modified	1 - 5	1 = "impossible", 5 = "fairly easy"
Parent knowledge of RHS prior to treatment	Single item	0 - 1	
Youth knowledge of RHS prior to treatment	HSQ (Moberg & Finch, 2008)	0 - 1	
Receipt of AOD or mental health counseling outside school	HSQ (Moberg & Finch, 2008)	0 - 1	
Attendance at AA/NA/other 12-Step meetings (frequency)	HSQ (Moberg & Finch, 2008)	1 - 6	1 = "never", 6 = "daily"

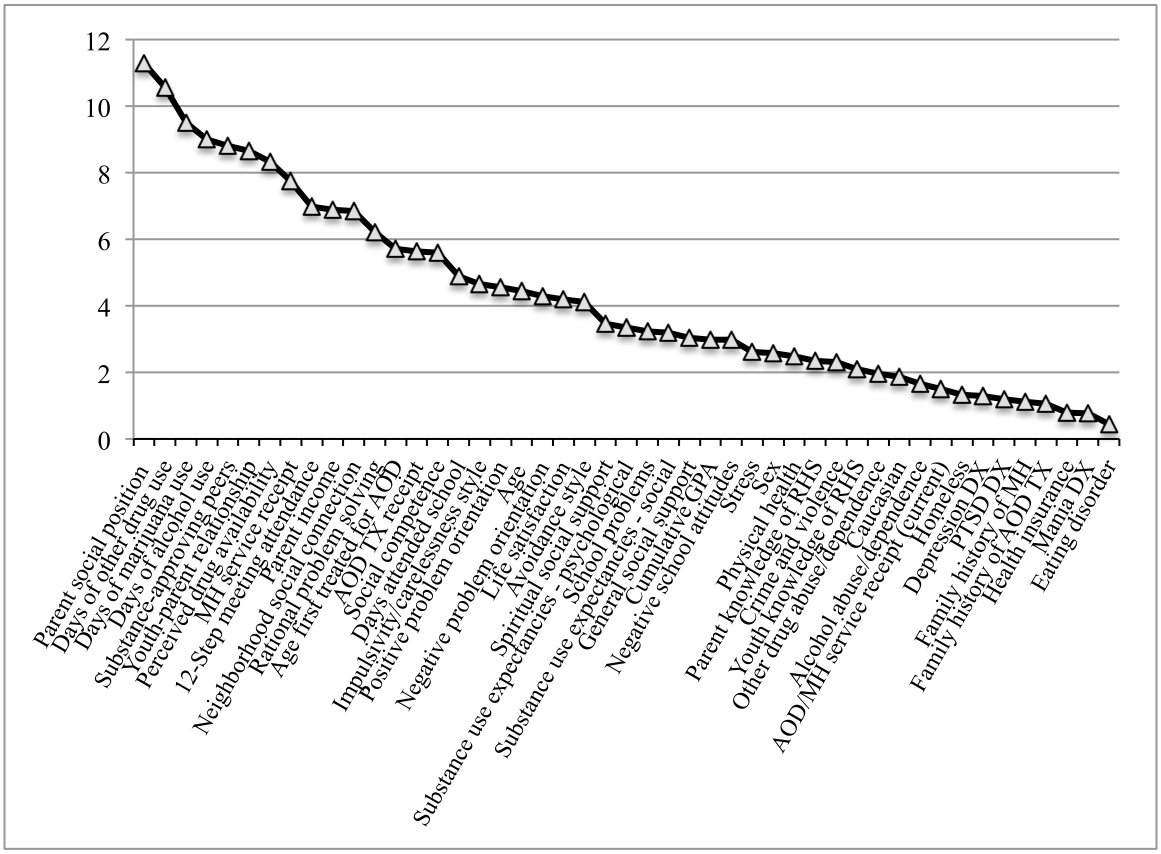

Figure 1. Variable Importance from Individual Random Forest Recovery Capital Domains

Note. Variables have been combined and reordered so that if a variable was deemed important in multiple individual recovery capital models, the highest variable importance was included and the other values were dropped.

CHAPTER IV

COVARIATE SELECTION FOR PROPENSITY SCORE ESTIMATION: A COMPARISON OF EXPLORATORY APPROACHES AND APPLICATION TO ADOLESCENT RECOVERY

Introduction

Adolescent recovery from a substance use disorder (SUD), like other social and public health issues, is a complex experience to understand because from an ecological perspective there are many levels of interacting factors that affect the onset of an SUD and the recovery process (Bronfenbrenner, 1977; 1994; Brown, Myers, Mott, & Vik, 1994; Chi, Kaskutas, Sterling, Campbell, & Weisner, 2009; Ramo & Brown, 2008; Rohde, Waldron, Turner, Brody, & Jorgensen, 2014). Studies of adolescent treatment and continuing care supports have often used randomized controlled trials (RCTs) to enable causal inferences about treatment effects on recovery outcomes. However, some aspects of the recovery experience cannot be studied using an RCT design for ethical or practical reasons, and thus many researchers have turned to nonrandomized quasi-experimental designs (QEDs) and their accompanying analytic methods. For example, a researcher interested in the effect of family support on recovery outcomes could not ethically conduct a study where adolescents are randomly assigned to a recovery-supportive family versus a non-supportive family. Similarly, for researchers interested in the effect of educational supports after treatment, it would be difficult to randomly assign adolescents and then mandate attendance to specific types of schools with varying degrees of recovery supports.

Thus, for social science questions where RCTs are not possible, QEDs may be implemented and during the analysis phase researchers attempt to reduce potential selection biases introduced due to the lack of randomization. Although there are a variety of ways to conduct analyses in QEDs to address the limitations of non-random allocation of participants, propensity scores are one of the most frequently used methods to address selection bias in QEDs, and thus these methods are the focus of this paper. There are, however, multiple steps in developing and using propensity scores (Austin, 2011b; Guo & Fraser, 2015) and there is ongoing work in the field of propensity score estimation and analysis to perfect each of these methodological decisions. There is not, however, one best method for choosing variables to include in the propensity score estimation model, which is one of the first and perhaps most important steps in a propensity score analysis. Thus, the focus of this paper is on the utility of exploratory methods, guided by a theoretical framework, for choosing covariates to be included in a propensity score estimation model.

This paper will first give a brief overview of the propensity score and key issues of covariate selection and estimation. Then, through the use of an empirical example utilizing data from an observational study, this paper will demonstrate and compare how traditional and data mining techniques choose covariates for the estimation of propensity scores, and how successful each method is at achieving balance among potentially confounding covariates. Using this empirical example, different sets of estimated propensity scores will next be used to predict outcomes. Finally, implications of using each type of propensity score estimation model will be discussed, with the primary focus on differences between the methods in covariate selection and balance for those interested in using these methods in the future for propensity score estimation.

Propensity Scores

A propensity score is a composite score derived from covariates and is used to balance different groups on pretreatment differences so that casual inferences about the effect(s) of the treatment of interest can be made (Luellen, Shadish, & Clark, 2005; Steiner & Cook, 2013). The propensity score is the probability that an individual will be in one condition (treatment) versus another (control) given a set of relevant, measured variables. Formally, the propensity score is defined as $e(\boldsymbol{x}_i) = pr(W_i = 1|\boldsymbol{X}_i = \boldsymbol{x}_i)$. That is, for participant i $(i = 1, \dots N)$ the propensity score is the conditional probability of treatment $(W_i = 1)$ versus nontreatment $(W_i = 1)$ given a vector $(\boldsymbol{x}_i)$ of observed covariates (Guo & Fraser, 2015; Luellen et al., 2005; Rosenbaum & Rubin, 1983). A key assumption for making causal inferences is that treatment assignment and the covariates measured at baseline are conditionally independent given the propensity score. For treatment assignment to be independent from the covariates measured at baseline, no baseline factor should affect whether or not a participant is in a particular group: if any variables that alter treatment assignment are not included in the propensity score estimation, they will introduce bias to the estimation of the effect of treatment.

Covariate Selection and Estimation of Propensity Scores

Because a propensity score is a composite score created from researcher-selected covariates, one must carefully measure and select appropriate variables that will be used in its estimation (Steiner & Cook, 2013). Any variable that could predict treatment assignment or differences on the outcome variable must be considered for inclusion in the propensity score model. Covariate selection for the propensity score estimation model is a burgeoning area of research and authors have reported a variety of ways to select covariates (Thoemmes & Kim, 2011), including by assessing bivariate relationships between proposed covariates and treatment and choosing covariates that have significant relationships with the treatment and the outcome of interest, using all measured variables and then testing for balance among covariates, using a stepwise regression procedure, and choosing key variables from correlation matrices of prior relevant meta-analyses (Tanner-Smith & Lipsey, 2014). There are not strict guidelines on how to select variables for inclusion or on the appropriate number of variables to include in the propensity score estimation model: a review of social science literature found that anywhere from 3 to 238 covariates have been used in propensity score estimation models (Thoemmes & Kim, 2011).

Additional considerations during variable selection for propensity score estimation is whether included variables have a non-linear relationship with the outcome or whether they interact with each other. Articles on propensity score analysis tend to rely on broad recommendations, such as ensuring that all potential confounders with the dependent variable are included in the analysis, without exploring estimation approaches that might maximize knowledge of the existing data. For example, a recent practical guide on propensity scores discussed various options in propensity score analysis, but did not explore alternative approaches for variable selection, for example, in cases of interactions between variables, which is problematic given the likely interactions in data addressing social science questions (Lanza, Moore, & Butera, 2013).

Logistic regression is frequently used to estimate propensity scores, where the key observed covariates are included as predictors and the treatment assignment is the binary dependent variable (Luellen et al., 2005). Specific interactions and quadratic or cubic predictors can be included in the logistic regression model; however, these terms must be selected carefully with the usual attention to building a parsimonious model. Thus, it is likely that important interactions or non-linear relationships could be ignored when logistic regressions are used to create propensity scores. Non-linearity and interactions between variables are not issues that have been adequately addressed in the social sciences for those interested in using propensity scores to achieve balance and those that wish to address this issue in their data need to undergo a series of checking each potential interaction in the model, which may be cumbersome for studies with a large number of potential covariates.

Studies in other fields, however, have explored the use of data mining techniques in the estimation of propensity scores. There are a variety of data mining methods available, but in general methods such as classification and regression trees (CART) allow pre-specified algorithms to determine the best covariates for propensity score estimation within a particular dataset (Luellen et al., 2005; Stone, Obrosky, Singer, Kapoor, & Fine, 1995; Thoemmes & Kim, 2011). These data mining techniques are also particularly useful for addressing interactions and potential non-linear relationships; covariate selection with these methods involves choosing covariates that minimize the error variance while simultaneously identifying important interactions between chosen variables that predict group membership (Lee, Lessler, & Stuart, 2009; McArdle, 2014; Morgan, 2005; Setoguchi, Schneeweiss, Brookhart, Glynn, & Cook, 2008). Simulation studies have demonstrated the utility of data mining techniques compared to more traditional models of propensity score estimation (Lee et al., 2009; Setoguchi et al., 2008). For example, in one simulation study, data mining approaches had consistently better reduction in bias for propensity score weighting than standard logistic regression models, especially in data that included covariate interactions and non-linearity (Lee et al., 2009). In another simulation study, logistic regression performed well, but introduced more bias than other approaches; however, no estimation method singly outperformed the other methods (Setoguchi et al., 2008).

This paper will focus on three specific data mining techniques, comparing results from these three methods to the traditional logistic regression approach: SEARCH (Morgan, 2005; Morgan, Solenberger, & Nagara, 2013; Morgan & Sonquist, 1963), classification trees (McArdle, 2014), and the ensemble method of random forests (Breiman, 1996; 2001). Using the available data, each of these exploratory methods sequentially tests available predictors and attempts to categorize participants into subgroups by covariate(s) that best predict the outcome of interest (Morgan & Sonquist, 1963; Sonquist et al., 1974). The resulting subgroups can then be used to create predicted group membership and propensity scores based on the odds of group membership. Both SEARCH and classification tree methods utilize covariates to create a single tree from the data while the random forest approach involves bootstrap aggregations, creating many (often 500-1000) trees and averaging results across trees. Although classification trees and forests have been used in propensity score creation (e.g., see Cham, 2013; McCaffrey, Ridgeway, & Morral, 2004; Stone et al., 1995), the SEARCH application has not been used and compared to these methods. Previous studies comparing different propensity score estimation techniques were simulation studies (Lee et al., 2009; Setoguchi et al., 2008), so it is unclear how these methods would perform with a smaller, non-simulated dataset. As each dataset is unique and may produce different results, a comparison of performance between these methods for a dataset collected from participants in an observational study will highlight analysis issues for social scientists interested in using propensity scores to achieve balance.

Motivating Example: Adolescents in Recovery

The motivating empirical example used in this paper is from a QED study of adolescents with SUDs followed longitudinally for 12 months after SUD treatment receipt. The aim of the study was to understand whether certain kinds of educational supports, Recovery High Schools

(RHSs; Moberg & Finch, 2007), were more effective in maintaining abstinence and sustaining school attendance than other educational supports, such as traditional or alternative schools not focused on recovery. Because the choice of where a student attends school after receiving SUD treatment may depend on factors that are also related to the outcomes of interest (relapse to substance use), bias will be introduced into the analysis if we do not control for these potential confounding variables. For example, adolescents who attend RHSs may have more supportive friends or family, may exhibit less (or greater) substance use severity, or may be more connected to the continuing care community. Alternatively, adolescents who attend non-RHSs could have fewer recovery needs or already be connected to a supportive peer community.

Thus, to make causal inferences about the effectiveness of RHSs compared to non-RHSs in this observational sample, it is necessary to identify variables that differentially predict RHS attendance and recovery outcomes and use these covariates in the propensity score estimation model. Without such an adjustment a comparison between the two groups on recovery outcomes would likely be biased given any baseline differences on these variables. The theoretical framework for the initial selection of covariates in this analysis is the ecological model of adolescent recovery capital; that is, the total amount of resources an individual has to recover from an SUD (Granfield & Cloud, 1999; Hennessy, 2017; White & Cloud, 2008). The adolescent recovery capital model is comprised of four primary domains including financial (e.g., income, health insurance), human (e.g., self-efficacy to recover, life skills), social (e.g., sober and supportive friends and family, connections to others), and community (e.g., access to community recovery supports such as 12-Step programming) recovery capital. These are all resources that may affect both the decision to engage in a recovery supportive environment such as the treatment of interest in this study, an RHS, as well as the ability to maintain abstinence from substance use. As such, they were considered primary covariates of interest, and an additional 11 control covariates, all variables that could potentially interact with recovery capital resources, such as age, sex, and treatment history, were included.

Methods

The aim of this study was to explore and compare how different methods of covariate selection in the creation of propensity scores changes the resulting estimated propensity score and covariate balance within the same dataset, with the emphasis on the potential utility of exploratory data mining approaches to reduce bias in the estimation of treatment effects. Logistic regression, SEARCH, classification trees, and random forests, were used to create and compare unique sets of estimated propensity scores to balance non-RHS and RHS groups. The resulting propensity scores from each method were then used to compare alcohol consumption and marijuana use outcomes by RHS condition.

Data

This analysis used data collected through the longitudinal, observational study titled, “Effectiveness of Recovery High Schools as Continuing Care” which received Institutional Review Board approval from the University of Minnesota. In this study, adolescents and their parent(s) in Minnesota, Wisconsin, and Texas were enrolled at the end of adolescents’ SUD treatment and data involving an extensive standardized interview protocol was collected at study enrollment and 3-, 6-, and 12-months later. Adolescents chose whether to enroll in school and if so, the type of school in which to enroll. Although there were a variety of schools that adolescents could choose to attend, the purpose of this study was to assess the effectiveness of RHSs compared to other types of schools. Thus, this analysis compared two non-randomly allocated groups: a treatment group (RHS attendance) and a comparison group (other school attendance, labeled the non-RHS group). The sample consisted of 260 adolescent participants[7], of whom 120 attended an RHS at 3-months post-baseline interview and 140 attended another type of school (non-RHS).

Analysis

In a previous paper using this sample, the relationship between individual recovery capital domains and predictors of school attendance were explored among 47 total variables. Each of the analyses discussed here first involved running a model that predicted RHS attendance by a specific recovery capital domain (e.g., financial, human, social, or community). Given the size of the sample (N = 260) and the number of available covariates, for each single domain model, those predictors that were deemed important (e.g., p value < .05 for logistic regression analysis or covariates that appeared in final tree for the classification tree analysis) were used to create a model predicting RHS attendance from multiple variables representing a measure of overall recovery capital. Thus, recovery capital and additional control variables that were found to be important in predicting RHS attendance in that paper were included in this analysis as potential covariates for propensity score estimation. As these variables were considered measures of recovery capital, they were also potentially important in predicting recovery-specific behavior outcomes, the outcomes examined in this paper. Of the potential 47 variables, 12 variables were not important in any of the analyses and were removed; the remaining 35 variables varied in their importance depending on the covariate selection method. This process thereby led to the estimation of one unique set of propensity scores from each separate propensity score estimation method, which were each used to create strata for covariate

[7] Although the sample from the observational study was slightly larger at 3-month follow-up (n = 322), youth were removed from the sample if it was unclear at the 3-month follow-up whether or not the youth was in school, and where they attended.

balance. If any method produced strata with either no treatment or no control participants, these strata were removed and participants in that strata were dropped from remaining analyses.

Logistic regression. Using the primary predictors of RHS attendance based on the recovery capital framework, propensity scores generated from logistic regressions were estimated in Stata SE (version 14.2; StataCorp., 2015). Given the number of covariates included, and the aim of this paper to demonstrate the potential utility and ease of employing exploratory versus traditional methods, only main effects were included in these models. Results from this analysis generated a single propensity score for each participant and these scores were used to create five strata of equal sample size. Five is the recommended number of strata to use in propensity score estimation as in most cases this number of strata should remove 90% of the bias (Cochran, 1968). Although studies have tested up to 10 strata (e.g., see Austin, 2011a), given the small sample in this study and the potential for creating sparse cells when more strata are used, this analysis used five strata.

Classification tree. Using the primary predictors of RHS attendance based on the recovery capital framework, one classification tree was created in the RStudio statistical environment (0.99.486) with the Recursive PARTitioning (`rpart`) package (Therneau, Atkinson, & Ripley, 2015). Given the small sample size, 10-fold cross validation was used in place of splitting data into a training and test data set (Hastie, Tibshairani, & Friedman, 2008). No stopping rules were used, but the complexity parameter, a measure of which splitting a variable node improves the relative error (Therneau et al., 2015) was used to prune each tree after growth to ensure that the model was not overfit (e.g., see Breiman et al., 1994; Torgo, 1999). Results from the pruned tree included the creation of propensity score strata, where each participant was assigned to a single stratum with the same propensity score. It is important to note that this method automatically produces the number of strata based on how many splits are chosen to create the tree, so the researcher has less choice in determining number of strata unless strata are combined. Strata were not combined in this analysis because it may have created imbalanced within-strata results across covariates.

Random forests. The method of random forests, which utilizes bootstrap aggregations (Breiman, 1996; 2001; Scott et al., 2014) in the R package, `randomForest`, was used to predict RHS attendance based on overall recovery capital variables (Liaw & Wiener, 2015). The forest method is similar to the sequential variable selection used in the classification trees but uses boosted aggregations of the sample to generate X number of trees while randomly choosing a set of variables (k) to attempt at each split: based on recommendations, 1000 trees were generated and k was set to three (e.g., see Hastie et al., 2009; Liaw & Wiener, 2002; Strobl et al., 2008). Results from this analysis generated a single propensity score for each participant and these scores were used to create five strata of equal sample size.

SEARCH. Analyses to predict RHS attendance from the recovery capital framework were initially conducted using the SEARCH program in Stata SE (14.2; StataCorp, 2015); however, these models resulted in too few covariates for propensity score estimation to achieve adequate balance on important variables as it only resulted in the emergence of two important covariates to predict RHS attendance: thus, the SEARCH method was not used to create propensity score strata for this sample.

Assessing Balance

Demonstrating that the propensity scores have produced balanced groups is an important part of assessing the performance of the estimated propensity score (Shadish & Steiner, 2010). To demonstrate balance, both groups must have similar distributions on the covariates measured at baseline and on the estimated propensity score. Balance can be assessed in several ways; however, one recommended method is to compare the two groups on effect size metrics including the standardized mean difference (Cohen's *d*) for continuous outcomes and the odds ratio for binary outcomes (Austin 2011a; Lanza et al., 2013; Shadish & Steiner, 2010). Effect sizes for each continuous and binary covariate were thus estimated to assess baseline differences between the RHS and non-RHS groups. Odds ratios for binary covariates were first estimated and then the Cox transformation outlined by Sánchez-Meca and colleagues was used to transform log odds ratios into Cohen's *d* so that all balance measures could be assessed on the same metric (Sánchez-Meca, Marín-Martínez, & Chacón-Moscoso, 2003). For each of the propensity score estimation models, effect sizes were calculated between groups in each stratum and then an overall average difference across the strata was produced to assess balance. Guidelines vary in terms of the recommended cutoff values used for assessing (im)balance: some scholars have recommend that a standardized mean difference greater than 0.10 between the two groups indicates that balance was not achieved (Steiner & Cook, 2013); however, this guideline may be too strict and others have recommended using 0.25 (Harder, Stuart, & Anthony, 2010). Thus, covariate balance will be discussed using both levels.

Additionally, the region of common support was assessed to determine whether the propensity scores demonstrated adequate overlap among the RHS and non-RHS groups and whether trimming participants outside of the common support region was necessary. Standardized mean difference effect sizes of the propensity scores within strata were calculated to statistically assess this in the propensity score estimates generated from the logistic regression and random forest models. Cell size and balance within each strata was used to assess the region of common support in the propensity score estimates from the single classification tree as each strata had a single propensity score.

Estimated Treatment Effects

For each covariate selection method, the Stata program, `mixed`, was used to estimate the (1) mean days of alcohol use and (2) mean days of marijuana use at the 6-month follow-up for participants in each group in each propensity score strata. Given that these data are from students in different schools, multilevel models with random intercepts for schools were used to account for clustering. Stratum-specific differences were then calculated and the overall estimated treatment effect was the mean of the stratum-specific differences in days of use for each substance use outcome.

Results

The baseline characteristics of those who attended RHSs versus those who did not are described in Table 1. Of the 35 covariates measured at baseline and used in this analysis due to their importance in predicting RHS attendance, 28 had standardized mean differences that exceeded 0.10 (six exceeded 0.25). These differences indicate systematic baseline differences between students who attended different school types at 3-months. Thus, the focus of the presentation of results will center on how well the different methods were able to balance these baseline covariates by their respective strata prior to propensity score estimation.

Covariate Selection using Logistic Regression

There were 11 variables that were included in the original logistic regression model predicting RHS attendance from overall recovery capital. However, propensity scores should take into account predictors that may impact the treatment outcome (i.e., in this case substance use) in addition to choice of treatment and the logistic regression method allows for including additional variables in propensity score estimation. Thus, the following baseline variables were added to the propensity score estimation model given their significant correlation ($p < .05$) with alcohol and/or marijuana use outcomes at six months: number of days of alcohol use, number of days of marijuana use, diagnosis of an alcohol use disorder, diagnosis of mania, positive screening for an eating disorder, crime and violence episodes, and youth-parent relationships. A total of 18 variables were included in this propensity score estimation model.

Due to missing values on some covariates, 226 of 260 participants were assigned a propensity score and the remaining 34 were dropped from further analysis. The quintiles of the estimated propensity score from the logistic regression were 0.008, 0.198, 0.378, 0.550, 0.740, respectively and each quintile was comprised of 45 or 46 participants (see Table 2). The proportion of participants within each stratum who attended RHSs ranged from a low of 8.7% in the strata with the lowest propensity score to a high of 86.7% in the stratum with the highest propensity score. Examining the propensity scores across strata and by treatment group, the lowest and highest strata exhibited less overlap between the two groups than the other strata, although none of the within-strata effect sizes were statistically significantly different. However, given the great degree of participants outside the region of common support in these strata, participants with propensity score values lower than 0.1 were removed from the analysis (quintile 1: non-RHS = 18, RHS = 1)[8]. This resulted in 207 participants to estimate the propensity scores and assess balance.

Assessing balance. Of the 35 covariates to assess for balance, 22 achieved balance with $d < 0.10$ and 12 with d values ranging from 0.10-0.25. Regarding only the 18 variables included in

[8] Originally, participants outside the range of 0.9 were also removed; however, this removed 8 RHS participants and the resulting within strata effect size between RHS and non-RHS groups was greater than before the removal of the participants and also statistically significant (SMD = -0.97, 95% CI [-1.87, -0.07]). Thus, no trimming for the highest propensity score strata was conducted and these 8 RHS participants were retained for further analysis.

the actual propensity score estimation, 17 achieved balance of $d < 0.10$, while one covariate (diagnosis of major depression) had a reduced effect size from $d = 0.24$ to $d = 0.14$. Thus, overall, the logistic regressions performed well in reducing imbalance across most of the covariates, including those that were not included in the propensity score estimation model, but were still potentially important confounders to address (See Table 3).

Covariate Selection using Classification Tree

There were 18 variables that were included in the original classification tree predicting RHS attendance from overall recovery capital. This original tree produced 11 strata using 11 of the 18 variables; however, one strata had only RHS students and another included only non-RHS students and the remaining strata were sparse, so the tree was pruned (see Table 4). In the results from the pruned tree, eight strata were produced using only seven of the included variables: one of these strata had 11 RHS students and no non-RHS students so this strata was dropped and these participants were removed from further analyses using the classification tree strata. The values of the estimated propensity score from this final model were 0.10, 0.20, 0.25, 0.27, 0.33, 0.718, 0.79, and 1.00, respectively. The proportion of participants within each stratum who attended RHSs ranged from a low of 10% in the stratum with the lowest propensity score to a high of 71.8% in stratum six. Conversely, the proportion of participants within each stratum who attended a non-RHS ranged from a low of 20.8% in stratum seven to a high of 75% in stratum three. Thus, there was somewhat unbalanced overlap in the propensity score for RHS and non-RHS participants across the strata.

Assessing balance. Only seven variables were used in the creation of the final pruned tree, indicating that only these variables were predictive of RHS attendance and would be balanced (see Table 5). For these seven variables, the lowest strata, which had only one RHS student and 10 non-RHS students did not demonstrate good balance. Although three variables, mania diagnosis, PTSD diagnosis, and AOD/MH counseling outside school, all had the same responses for both RHS and non-RHS attendees, the remaining comparisons between the two groups demonstrated imbalance. Thus, given the size of this lowest strata and these covariate results, strata 1 was removed from treatment effects analyses, along with the eighth strata, which was removed for being outside the region of common support as it only contained RHS attendees.

Regarding the mean effect from strata 2-7, 10 of the effect sizes resulted in an effect size of $d < 0.10$, while 13 had d values ranging from 0.10-0.25, demonstrating reductions in 23 of 35 effect sizes comparing the RHS and non-RHS students on baseline covariates. Worth noting, four of the binary variables (i.e., diagnosis of major depression, manic episode, PTSD, and youth knowledge of RHS prior to treatment) had some strata where the outcome did not vary between the two groups. However, of the 18 covariates included in the propensity score estimation model, only five had a reduction to $d < 0.10$ and only five were reduced to d between 0.10 to 0.25: of those eight covariates that were utilized in estimating propensity scores from the pruned tree, only six covariates were reduced to $d < 0.25$. Additionally, some covariates demonstrated greater imbalance across the strata than was evidenced at baseline (e.g., number of times of previous AOD treatment, physical health, neighborhood social connections).

Covariate Selection using Random Forest

There were 22 variables that were included in the random forest analysis that predicted RHS attendance from overall recovery capital. Additionally, because this method estimates multiple potential trees, a measure of variable importance in predicting the outcome is given for all included variables. The quintiles of the estimated propensity score from the random forest model were 0.096, 0.346, 0.422, 0.493, 0.581, respectively and each quintile was comprised of 52 participants (see Table 6). The proportion of participants within each stratum who attended RHSs ranged from a low of 34.6% in both of the strata with the lowest propensity scores to a high of 65.4% in the stratum with the highest propensity score. The range of the propensity score for each group within each stratum was fairly similar and the mean for each group within each stratum only had slight differences: effect sizes for propensity score differences by group within each stratum were nonsignificant. Thus, there was reasonable overlap in the propensity score between treated and untreated participants and no participants were dropped from the analysis.[9]

Assessing balance. Regarding the mean effect from strata 1-5, 11 of the effect sizes resulted in an effect size of $d < 0.10$, while 15 were d ranging from 0.10-0.25, demonstrating reductions in 26 of 35 effect sizes comparing the RHS and non-RHS students on baseline covariates. Of the 35 potential covariates, 22 were included in the propensity score estimation model using the random forest method (see Table 7). For these 22 variables, all of the new effect sizes generated from the average of the five effect sizes across the random forest strata indicated a reduction in the effect size; however, nine of these new effect sizes resulted in $d < 0.10$ and ten resulted in d ranging from 0.10-0.20, indicating three covariates with remaining imbalance between the groups. When viewing the result of the effect sizes from the strata generated using random forests against the variable importance results (see Table 8), the three most important variables, rational problem solving style, days of school attendance, and impulsivity/carelessness problem solving style, demonstrated a reduction in the effect size; however, for each covariate d

[9] Austin (2011) does not recommend removing participants from strata generated in this way if it appears that there is reasonable overlap because this will change the population of the analyzed treatment group. Additionally, following Crump and colleagues' (2009) rule of thumb and trimming scores outside the range [0.1, 0.9] would have resulted in only trimming one participant from the sample.

> 0.18, indicating potential imbalance in these three major variables, especially for rational problem solving.

Estimation of Treatment Effects

Across all strata, the analysis from each propensity score estimation model demonstrated a consistent and strong, but non-significant reduction in number of days of use of marijuana at 6-months if participants attended an RHS versus a non-RHS after treatment; however the estimate of the effect varied by method with the logistic regression showing the smallest reduction in days of use and the random forests demonstrating a greater reduction in days of use (see Table 9).

The results for days of use of alcohol were less consistent across the methods, however, with the estimate using strata generated from the logistic regression model demonstrating a small, but insignificant increase in days of alcohol use at 6-months if participants attended an RHS versus a non-RHS after treatment. Both the models using strata generated from the single classification tree and the random forest demonstrated a small but insignificant decrease in days of alcohol use at 6-months if participants attended an RHS versus a non-RHS after treatment. Thus, RHS and non-RHS students did not demonstrate any significant differences on substance use outcomes at 6 months post baseline study entry.

Discussion

This paper used three different covariate selection models for choosing covariates to estimate propensity scores, using demonstration data from a quasi-experimental study designed to assess the effectiveness of RHSs for youth in recovery from SUDs. The recovery capital theoretical framework was used to identify relevant covariates, from among a large set of potential covariates, for inclusion in the propensity score estimation models. Each covariate selection method (logistic regression, classification tree, and random forest) identified a unique set of baseline covariates, although there was some overlap in identified covariates across the methods. For example, all methods included number of days attended school in the past 12 months, rational problem solving scores, and frequency of 12-Step meeting attendance. As would be expected, however, all methods produced different propensity score estimates, identified slightly different sets of participants to be included in unique stratum, achieved slightly different balance on baseline covariates, and estimated different treatment effects. In addition, although the initial analysis plan included a comparison of the SEARCH method, another recursive partitioning technique, this method did not identify enough covariates to reasonably address balance in a propensity score estimation model. This may be due to the small sample size of the empirical example, however, and future research with larger samples should compare this method alongside the methods utilized here.

Of primary importance in this analysis is how well the different methods achieved balance on baseline covariates to reduce bias in the estimation of treatment effects. It appears that the best balance on all included covariates was achieved using the logistic regression covariate selection model. Relative to the logistic regression results, where balance of $d < 0.25$ was achieved on almost all potentially important covariates (97%), including those not incorporated in the estimation of propensity scores, both the classification tree and random forest model were not as effective in achieving balance (66% and 74%, respectively, achieved adequate balance). Regarding the six covariates the logistic regressions and classification trees shared in their propensity score estimation models, the logistic regressions performed much better on balancing two covariates, but otherwise the methods were very similar in producing balance for the shared covariates. Regarding the 10 covariates the logistic regressions and random forests shared, both days of alcohol use and perceived availability of drugs were slightly better balanced in the random forests than the logistic regression; however, life satisfaction, rational problem solving style, 12-Step meeting attendance, and age were better balanced in the logistic regression than the random forest. The remaining shared four covariates had similar balance between the two methods. These details are important to note as anyone using propensity scores in their estimation of treatment effects must ensure adequate balance across potential confounding covariates prior to moving forward with any sort of treatment group comparisons.

However, the logistic regression model had the fewest cases for the final treatment effects analysis, only 68% of the available sample. Quite a bit of the sample was lost to missing values in some of the covariates and trimming the lowest strata resulted in one stratum with almost half the participants than the other four strata. This loss leads to an inefficient estimate of the treatment effect with decreased power. Alternatively, the random forest method retained 100% of the sample for the propensity score estimation and 85% of the sample in the treatment effects analysis. Indeed, the strata demonstrated equal sample sizes with fairly evenly distributed propensity scores between participants within a single stratum. If trimming participants had occurred based on previously established rules of thumb, only one participant would have been removed from the sample. Thus, for studies with smaller samples and some missing data, the random forest method of propensity score estimation may enable keeping more of the original sample and increased efficiency in estimating the treatment effect. However, given the differences in the size of the estimated treatment effect in this sample (although not in the statistical significance of the results) between covariate selection methods, it seems that both logistic regressions and random forests should be compared, especially on balance measures in small samples to ensure adequate balance prior to estimation of the treatment effect. For example, in this study, the balance achieved for covariates included in the random forests was slightly worse than balance from the logistic regression models, with three covariates demonstrating $d > 0.25$. The random forest method looks for and incorporates important interactions between covariates, which the logistic regression model cannot do, unless pre-specified interactions were included. So, although some balance appears to be worse in the random forest propensity score estimation model, this method may better account for key interactions between variables that were not included in the logistic regressions and may also allow for a more efficient estimate of treatment effects. To adjust for these differences in balance if they occur when using random forests as compared to logistic regression, researchers could then include any imbalanced covariates in the estimated treatment effect as control variables. Alternatively, it may be useful for researchers to first use the random forest method to identify important covariate interactions for inclusion in the logistic regression framework. This approach may reduce some of the burden for researchers in testing interactions when using logistic regression models to estimate propensity scores.

The single classification tree appeared to perform the worst out of all the models. The original tree had too many strata with empty cells and even after pruning the tree there was still one stratum with empty cells and another with poor balance that had to be trimmed from further analysis. Additionally, the estimated propensity scores were fairly crude in that each strata was assigned a single propensity score value, where both the logistic regression and the random forests had ranges of propensity score values within strata so that an assessment of ranges of scores between treatment and control participants could be calculated. Balance was achieved on some covariates but the final classification tree only included seven covariates, so other potentially important baseline differences that would affect treatment outcomes were not included and subsequently not adjusted. For example, compared to the logistic regression model, the classification tree did not adjust for age, negative attitudes towards school, life satisfaction, stress, and impulsivity/carelessness style of problem solving, all variables demonstrating $d >$ 0.20 between the two groups at baseline. In comparison to another non-parametric approach used, the random forests achieved better balance on more covariates. For example, the random forests included adjustments for days of alcohol use at baseline so that the difference between the two groups averaged across strata was $d = 0.12$, while the single tree did not adequately address baseline imbalance on this covariate (although there was still some reduction in the effect size between groups: $d = 0.19$). This is important to note because days of alcohol consumption at baseline was significantly and positively correlated with consumption at 6 months (alcohol use = 0.28, $p < .05$) and thus needed to be accounted for prior to estimating the treatment effect.

Finally, it is worth noting that despite accounting for baseline differences between the RHS and non-RHS groups via the use of propensity scores, there were no significant differences in substance use at 6 months across any of the three methods. Although treatment effects were not the explicit focus of this paper, and were used primarily for demonstration purposes, this bears some attention. It is possible given the small sample size that the study was underpowered to detect a significant treatment effect. Indeed, the original study design utilized a priori power analysis, which suggested a sample of 914 participants after 3-month follow-up and pruning based on propensity score analysis to detect an effect of 0.20; however, the study suffered from recruitment issues and enrolled a much smaller number of adolescent-caregiver dyads than originally projected. Thus, low recruitment may have resulted in the inability to detect an effect between the two groups.

In comparison to the treatment effect results of the primary study's outcomes analysis (Finch, Tanner-Smith, Hennessy, & Moberg, 2017), the methods used in this paper produced somewhat similar results in that differences in mean days of alcohol use were nonsignificant between groups although mean days of marijuana use between the two groups was significantly different in the primary outcomes paper. However, that paper had a slightly different sample size due to the timeframe used, defined RHS attendance using a dosage threshold of 20 or more days of school enrollment, and utilized propensity scores as covariate controls, a different estimation method than the method of stratification used in this paper.

Limitations

The small sample size from this QED is a potential limitation of this study, especially considering that data mining approaches recommend using a larger sample for more accurate estimates. Thus, with a larger sample, the propensity score estimates and resulting balance may have differed. Additionally, with a larger sample, it is likely that the SEARCH method would have identified more covariates and thus been included in the comparison of covariate balance. A similar assessment that compares these methods with a larger sample size should be undertaken in the future.

Given the focus on this study of exploring the potential utility of data mining methods, the logistic regression methods were fairly simplistic and interaction terms were not tested in propensity score estimation as might be done in an analysis that uses only this method. As one of the aims was to demonstrate how data mining techniques could identify and include interactions in the estimation of propensity scores, this limit is within the scope of the paper; thus this paper serves to demonstrate a less cumbersome, but methodologically appropriate way, for researchers estimating propensity scores to address these issues. Finally, there are many post-propensity score estimation options to use the propensity scores to assess balance and estimate treatment effects, of which stratification is just one (Guo & Fraser, 2015; Thoemmes & Kim, 2011). The use of stratification was chosen primarily because the predicted group membership from the classification tree results were only available as stratified results and thus to make this method comparable to the other methods used; however, it would be worthwhile for future research to compare the propensity scores generated from the logistic regressions and the random forests as weights, or another viable technique, to see if balance was improved with another method of incorporating the propensity score.

Implications

Given the many complex research questions of interest to social scientists and the increased recognition that the research "gold standard" of conducing RCTs is not always the most ethical or feasible method of research, QEDs are recognized as viable, and oftentimes more appropriate, research designs for understanding issues outside the laboratory setting. As a result, analytic methods to address imbalance on baseline variables that are confounded with treatment selection are increasingly being developed and used; these methods can also be incorporated to address issues of selection bias when randomization fails (e.g., see Williamson, Forbes, & White, 2014). Propensity scores are therefore one way for researchers to address selection bias that occurs in these natural settings, or to address selection bias in controlled trials, with all measured potential confounders.

However, there is a fair degree of subjectivity in deciding which covariates to include in these models, and thus, the selection of all important covariates remains an area for increased attention in evaluation research. Additionally, given the diversity of populations, treatments, and outcomes studied, each research question has the potential for very different sets of confounders and a variable degree of researcher subjectivity. Initially, using a theoretical framework to measure and select potentially confounding variables is extremely important. The theory-driven approach in this study via use of the recovery capital framework, is one way to reduce the subjectivity in choosing covariates to include from a set of many; however, it still enabled the consideration of a great number of potential covariates for the covariate selection models. Thus, in addition to the theoretical framework, the analytic method chosen becomes highly important in selecting covariates to reduce selection bias.

The results of this analysis demonstrated some potential benefits and challenges of using different data mining techniques to choose covariates for the estimation of propensity scores. Based on this empirical example, the logistic regression approach to choosing covariates for propensity score estimation may best balance the included covariates with small sample sized studies. However, with larger samples, the random forest approach deserves attention because it can account for many interactions that may not be tested in the logistic regression framework, and unlike the single classification tree, it averages across many trees to find the best fitting model and resulting propensity score. The method also produced propensity scores that were divided into equal strata, achieving the largest region of common support, leading to an estimation of the treatment effect with the largest proportion of the primary sample. It is worth noting here that the random forest method used in this analysis is just one among many data mining techniques that utilizes non-parametric techniques: other potentially viable data mining methods that have been used in the estimation of propensity scores include neural networks, bagged CART, and generalized boosted models (Lee et al., 2009; McCaffrey et al., 2004; Setoguchi et al., 2008) and are worth further exploration with similar sized samples.

To conclude, ensuring that any potential confounding variables leading to selection bias are mitigated prior to undergoing treatment effects analysis is necessary for evaluation researchers to properly understand whether a treatment is in fact effective for the population of interest. The methods proposed here aim to assist those interested in using propensity scores to address this issue and highlight some new approaches to choose the correct covariates to ensure balance that will be less burdensome for samples with a large number of potential covariates.

References

Austin, P. C. (2011a). A tutorial and case study in propensity score analysis: An application to estimating the effect of in-hospital smoking cessation counseling on mortality. *Multivariate Behavioral Research, 46*, 119-151. http://dx.doi.org/10.1080/00273171.2011.540480

Austin, P. C. (2011b). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research, 46*, 399-424. doi:10.1080/00273171.2011.568786

Breiman, L. (1996). Bagging predictors. *Machine Learning*, *24*(2), 123-140. doi:10.1007/BF00058655

Breiman, L. (2001). Random forests. *Machine Learning, 45,* 5-32. doi:10.1023/A:1010933404324

Bronfenbrenner, U. (1977). Towards an experimental ecology of human development. *American Psychologist,* 518-531.

Bronfenbrenner, U. (1994). Ecological models of human development. In *International Encyclopedia of Education, Vol. 3, 2nd Ed.* Oxford: Elsevier. Reprinted in: M. Gauvain & M. Cole (Eds.), *Readings on the Development of Children, 2nd Ed.* (1993, pp. 37-43). NY: Freeman.

Brown, S. A., Myers, M. G., Mott, M. A., & Vik, P. W. (1994). Correlates of success following treatment for adolescent substance abuse. *Applied & Preventive Psychology, 3,* 61-73.

Cham, H. N. (2013). *Propensity score estimation with random forests.* Dissertation. (UMI 3567836).

Chi, F. W., Kaskutas, L. A., Sterling, S., Campbell, C. I., & Weisner, C. (2009). Twelve-step affiliation and 3-year substance use outcomes among adolescents: Social support and religious service attendance as potential mediators. *Addiction*, *104*(6), 927-939. doi:doi:10.1111/j.1360-0443.2009.02524.x

Cochran, W. G. (1968). The effectiveness of adjustment by subclassification in removing bias in observational studies. *Biometrics*, 295-313.

Crump, R. K., Hotz, V. J., Imbens, G. W., & Mitnik, O. A. (2009). Dealing with limited overlap in estimation of average treatment effects. *Biometrika, 96,* 187-199.

Finch, A. J., Tanner-Smith, E. E., Hennessy, E. A., & Moberg, D. P. (2017). Recovery High Schools: Effect of schools supporting recovery from substance use disorders. *American Journal of Drug and Alcohol Abuse*, under review.

Granfield, R., & Cloud, W. (1999). *Coming clean: Overcoming addiction without treatment.* New York: New York University Press.

Harder, V. S., Stuart, E. A., & Anthony, J. C. (2010). Propensity score techniques and the assessment of measured covariate balance to test causal associations in psychological research. *Psychological Methods, 15*(3), 234.

Guo. S, & Fraser, M. W. (2015). *Propensity score analysis*. *Statistical methods and applications* (Second Edition). Thousand, Oaks, California: Sage.

Lanza, S. T., Moore, J. E., & Butera, N. M. (2013). Drawing causal inferences using propensity scores: a practical guide for community psychologists. *American Journal of Community Psychology*, *52*(3-4), 380-392.

Lee, B. K., Lessler, J., & Stuart, E. A. (2009). Improving propensity score weighting using machine learning. *Statistics in Medicine, 29*, 337-346.

Liaw, A., & Wiener, M. (2015). Breiman and Cutler's random forests for classification and regression [Version 4.6.-12] (Software). Available from https://cran.r-project.org/web/packages/randomForest/randomForest.pdf

Luellen, J. K., Shadish, W. R., & Clark, M. H. (2005). Propensity scores: An introduction and experimental test. *Evaluation Review*, *29*(6), 530-58. doi:10.1177/0193841X05275596

McArdle, J. J. (2014). Exploratory data mining using decision trees in the behavioral sciences. In J. J. McArdle & G. Ritschard (Eds.), *Quantitative methodology series: Contemporary issues in exploratory mining in the behavioral sciences* (pp. 1-47). New York, NY: Routledge.

McCaffrey, D. F., Ridgeway, G., & Morral, A. R. (2004). Propensity score estimation with boosted regression for evaluating causal effects in observational studies. *Psychological Methods*, *9*(4), 403-25. doi:10.1037/1082-989X.9.4.403

Moberg, D. P., & Finch, A. J. (2007). Recovery high schools: A descriptive study of school programs and students. *Journal of Groups in Addiction & Recovery*, *2*, 128-161. doi:10.1080/15560350802081314

Morgan, J. N. (2005). History and potential of binary segmentation for exploratory data analysis. *Journal of Data Science*, *3*(2), 123-136.

Morgan, J. N., Solenberger, P. W., & Nagara, P. R. (2013). History and future of binary segmentation programs. Survey Research Center, Institute for Social Research, University of Michigan. Retrieved from http://www.isr.umich.edu/src/smp/search/search_paper.html

Morgan, J. N., & Sonquist, J. A. (1963). Problems in the analysis of survey data, and a proposal. *Journal of the American Statistical Association*, *58*(302), 415-434.

Ramo, D. E., & Brown, S. A. (2008). Classes of substance abuse relapse situations: A comparison of adolescents and adults. *Psychology of Addictive Behaviors*, *22*(3), 372-9. doi:10.1037/0893-164X.22.3.372

Rohde, P., Waldron, H. B., Turner, C. W., Brody, J., & Jorgensen, J. (2014). Sequenced versus coordinated treatment for adolescents with comorbid depressive and substance use disorders. *Journal of Consulting and Clinical Psychology*, *82*(2), 342-8. doi:10.1037/a003580

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70*, 41-55.

Sánchez-Meca, J., Marín-Martínez, F., & Chacón-Moscoso, S. (2003). Effect-size indices for dichotomized outcomes in meta-analysis. *Psychological methods*, *8*(4), 448.

Setoguchi, S., Schneeweiss, S., Brookhart, M. A., Glynn, R. J., & Cook, E. F. (2008). Evaluating uses of data mining techniques in propensity score estimation: A simulation study. *Pharmacoepidemiology and Drug Safety*, *17*(6), 546-55. doi:10.1002/pds.1555

Shadish, W. R., & Steiner, P. M. (2010). A primer on propensity score analysis. *Newborn & Infant Nursing Reviews, 10*, 19-26.

StataCorp (2015). *Stata Statistical Software: Release 14*. College Station, TX: StataCorp LP.

Steiner, P. M., & Cook, D. (2013). Matching and Propensity Scores. In *The Oxford Handbook of Quantitative Methods (Vol 1)*, T. D. Little (Ed.). New York: Oxford University Press.

Stone, R. A., Obrosky, D. S., Singer, D. E., Kapoor, W. N., & Fine, M. J. (1995). Propensity score adjustment for pretreatment differences between hospitalized and ambulatory patients with community-acquired pneumonia. Pneumonia patient outcomes research team (PORT) investigators. *Medical Care*, *33*(4 Suppl), AS56-66.

Tanner-Smith, E. E., & Lipsey, M. W. (2014). Identifying baseline covariates for use in propensity scores: A novel approach illustrated for a nonrandomized study of recovery high schools. *Peabody Journal of Education*, *89*(2), 183-196. doi:10.1080/0161956X.2014.895647

Therneau, T., Atkinson, B, & Ripley, B. (2015). Recursive partitioning and regression trees [Version 4.1.-10] (Software). Available from: https://cran.r-project.org/web/packages/rpart/

Thoemmes, F. J., & Kim, E. S. (2011). A systematic review of propensity score methods in the social sciences. *Multivariate Behavioral Research*, *46*(1), 90-118. doi:10.1080/00273171.2011.540475

White, W., & Cloud, W. (2008). Recovery capital: A primer for addictions professionals. *Counselor*, *9*(5), 22-27.

Williamson, E. J., Forbes, A., & White, I. R. (2014). Variance reduction in randomised trials by inverse probability weighting using the propensity score. *Statistics in medicine*, *33*(5), 721-737.

Table 1. Baseline Characteristics by Treatment Group

	RHS M/n	RHS SD/%	nonR M/n	nonR SD/%	*d*	n
Control Variables (k = 9)						
Age	16.47	1.00	16.24	1.05	0.22	260
Age first treated for AOD	15.29	1.20	15.50	1.20	-0.18	259
Alcohol use disorder diagnosis	83	(69%)	82	(59%)	0.12	260
AOD treatment (# times)	5.13	27.62	1.86	1.35	0.17	257
Days of alcohol use	18.84	25.36	15.69	22.62	0.13	260
Days of marijuana use	53.45	35.55	55.88	34.51	-0.07	260
Days of other drug use	30.58	36.17	20.53	28.99	0.31	260
MH service receipt (# times)	6.12	20.32	3.71	3.84	0.17	255
Sex (male = 1)	56	(47%)	81	(58%)	-0.12	260
Human Capital (k = 16)						
Crime and violence screener, lifetime	2.47	1.51	2.60	1.48	-0.09	260
DSM-IV diagnosis: major depression	101	(84%)	95	(68%)	0.24	260
DSM-IV diagnosis: manic episode	16	(13%)	28	(20%)	-0.13	260
DSM-IV diagnosis: PTSD	46	(38%)	30	(21%)	0.22	260
Life satisfaction	3.70	0.60	3.44	0.66	0.40	260
MH screens: eating disorder	36	(30%)	41	(29%)	0.01	260
Negative attitudes towards school	0.60	0.26	0.55	0.25	0.21	258
Physical health	3.62	1.63	3.21	1.77	0.24	260
Problem solving - positive orientation	0.04	0.90	-0.03	0.87	0.08	244
Problem solving - negative orientation	0.06	0.79	-0.05	0.88	0.13	244
Problem solving - rational style	-0.23	0.84	0.20	0.82	-0.52	244
Problem solving - impulsivity/carelessness style	0.10	0.80	-0.08	0.79	0.23	244
Problem solving - avoidance style	-0.05	0.78	0.04	0.78	-0.12	244
School attendance, last 12 months (days)	162.22	38.91	147.34	51.47	0.32	260
Stress	3.48	1.76	3.01	1.77	0.26	260
Substance use expectancies - social benefits	22.24	4.42	21.47	3.82	0.19	259
Financial Capital (k = 2)						
Family income level	5.50	1.32	5.25	1.46	0.18	243
Parental social position score (average)	33.37	12.93	36.32	15.10	-0.21	245
Social Capital (k = 4)						
Neighborhood social connections index	2.94	0.73	2.97	0.78	-0.03	257
Social competence index	3.10	0.35	3.07	0.34	0.07	258
Substance approving peer attitudes	3.13	0.57	3.01	0.45	0.24	257
Youth-parent relationship	61.38	22.44	57.35	20.59	0.19	258
Community Capital (k = 4)						
AA/NA/12 step meeting attendance	3.69	1.56	3.10	1.66	0.36	259
AOD/MH counseling outside school	88	(73%)	109	(78%)	-0.06	259
Perceived availability of drugs	4.28	0.62	4.19	0.66	0.15	260
Youth knowledge of RHS prior to TX	35	(29%)	56	(40%)	-0.15	240

Note. *d* = Cohen's *d* standardized mean difference effect size

Table 2. Stratification Results from Logistic Regression Covariate Selection Model

	Logistic Regressions (pre-trimmed)								Logistic Regressions (trimmed)							
Strata	PS	RHS PS M	RHS PS range	nRHS PS M	nRHS PS range	SMD	RHS n	nRHS n	PS	RHS PS M	RHS PS range	nRHS PS M	nRHS PS range	SMD	RHS n	nRHS n
1	0.008	0.14	0.092-0.179	0.11	0.008-0.194	0.48	4	42	0.106	0.16	0.137-0.179	0.16	0.106-0.194	-0.07	3	24
2	0.198	0.29	0.198-0.369	0.28	0.199-0.368	0.26	15	30	0.198	0.29	0.198-0.369	0.28	0.199-0.368	0.26	15	30
3	0.378	0.45	0.378-0.534	0.46	0.378-0.548	-0.20	20	25	0.378	0.45	0.378-0.534	0.46	0.378-0.548	-0.20	20	25
4	0.550	0.66	0.550-0.738	0.63	0.558-0.735	0.55	28	17	0.550	0.66	0.550-0.738	0.63	0.558-0.735	0.55	28	17
5	0.740	0.84	0.740-0.960	0.86	0.826-0.897	-0.33	39	6	0.740	0.84	0.740-0.960	0.86	0.826-0.897	-0.33	39	6

Table 3. Balance Results from Logistic Regression Covariate Selection Model

	Pre-strata balance *d*	S1 n=27	S2 n=45	S3 n=45 *d*	S4 n=45	S5 n=45	S1-5 *d*
Control Variables (k = 9)							
Age	**0.22**	**-0.31**	**-0.18**	**0.50**	**-0.24**	**-0.25**	**-0.08**
Age first treated for AOD	-0.18	-0.57	-0.25	0.17	-0.26	-0.23	-0.20
Alcohol use disorder diagnosis	**0.12**	**0.00**	**-0.08**	**0.27**	**0.14**	**-0.15**	**0.04**
AOD treatment (# times)	0.17	0.28	-0.22	0.51	0.40	0.24	0.24
Days of alcohol use	**0.13**	**-0.10**	**0.11**	**0.10**	**-0.42**	**0.18**	**-0.02**
Days of marijuana use	**-0.07**	**-0.32**	**0.26**	**0.00**	**-0.29**	**0.12**	**-0.02**
Days of other drug use	0.31	-0.57	0.63	-0.14	0.46	-0.06	0.12
MH service receipt (# times)	0.17	-0.02	-0.16	-0.03	0.06	0.11	-0.01
Sex (male = 1)	**-0.12**	**0.00**	**0.08**	**0.05**	**-0.63**	**0.26**	**-0.05**
Human Capital (k = 16)							
Crime and violence screener, lifetime	**-0.09**	**-0.18**	**-0.19**	**0.06**	**0.12**	**-0.04**	**-0.03**
DSM-IV diagnosis: major depression	**0.24**	**0.00**	**0.59**	**-0.34**	**-0.40**	**0.79**	**0.14**
DSM-IV diagnosis: manic episode	**-0.13**	**0.12**	**0.00**	**0.06**	**0.40**	**-0.53**	**0.00**
DSM-IV diagnosis: PTSD	0.22	0.25	0.36	0.12	0.67	0.82	0.46
Life satisfaction	**0.40**	**0.92**	**0.14**	**-0.23**	**0.06**	**-0.22**	**0.06**
MH screens: eating disorder	**0.01**	**0.00**	**-0.23**	**0.72**	**0.17**	**-0.82**	**-0.03**
Negative attitudes towards school	**0.21**	**-0.85**	**-0.07**	**0.22**	**0.13**	**0.28**	**0.01**
Physical health	0.24	-0.10	0.24	0.27	0.17	-0.13	0.11
Problem solving - positive orientation	0.08	0.29	0.30	-0.33	-0.03	-0.20	-0.02
Problem solving - negative orientation	0.13	-0.18	0.29	0.02	-0.19	-0.08	-0.02
Problem solving - rational style	**-0.52**	**0.00**	**-0.07**	**0.33**	**-0.26**	**-0.25**	**-0.05**
Problem solving - impulsivity/carelessness style	**0.23**	**-0.84**	**-0.07**	**0.35**	**0.10**	**-0.11**	**-0.05**
Problem solving - avoidance style	-0.12	0.09	-0.15	0.45	0.11	-0.95	-0.11
School attendance, last 12 months (days)	**0.32**	**0.55**	**0.03**	**-0.17**	**0.03**	**-0.17**	**0.01**
Stress	**0.26**	**0.65**	**0.25**	**-0.27**	**0.23**	**-0.84**	**-0.05**
Substance use expectancies - social benefits	0.19	0.31	0.00	0.42	-0.12	0.42	0.20
Financial Capital (k = 2)							
Family income level	0.18	0.27	-0.99	-0.02	0.46	0.94	0.12
Parental social position score (average)	-0.21	0.56	0.46	-0.18	-0.50	-0.67	-0.12
Social Capital (k = 4)							
Neighborhood social connections	-0.03	0.42	-0.35	0.10	-0.12	-0.64	-0.16
Social competence	0.07	0.26	-0.42	-0.28	0.14	-0.03	-0.09
Substance approving peer attitudes	0.24	0.09	-0.16	-0.07	0.19	0.35	0.08
Youth-parent relationship	**0.19**	**-0.21**	**0.02**	**-0.14**	**0.25**	**0.28**	**0.06**
Community Capital (k = 4)							
AA/NA/12 step meeting attendance	**0.36**	**-0.02**	**-0.29**	**0.33**	**0.19**	**-0.53**	**-0.07**
AOD/MH counseling outside school	-0.06	-0.56	0.12	-0.19	-0.48	-0.15	-0.22
Perceived availability of drugs	0.15	0.04	-0.01	0.28	0.28	-0.07	0.11
Youth knowledge of RHS prior to TX	**-0.15**	**0.00**	**0.34**	**-0.59**	**0.45**	**-0.31**	**-0.02**

Note. Bolded values indicate variable was included in the model. *d* = Cohen's *d* standardized mean difference effect size.

Table 4. Stratification Results from Classification Tree Covariate Selection Model

	Unpruned			Pruned		
Strata	PS	RHS n	nRHS n	PS	RHS n	nRHS n
1	0.000	0	23	0.100	1	9
2	0.100	1	9	0.200	2	8
3	0.143	7	42	0.250	24	72
4	0.200	2	8	0.273	3	8
5	0.273	3	8	0.333	9	18
6	0.333	12	24	0.718	51	20
7	0.600	9	6	0.792	19	5
8	0.774	48	14	1.000	11	0
9	0.792	19	5			
10	0.889	8	1			
11	1.000	11	0			

Table 5. Balance Results from Classification Tree Covariate Selection Model

	Pre-strata balance	S1 n=10			S2 n=10	S3 n=96	S4 n=11	S5 n=27	S6 n=71	S7 n=24	
		RHS n=1	nRHS M	nRHS SD							S2-7
	d				*d*						*d*
Control Variables (k = 9)											
Age	0.22	17.00	16.44	0.88	1.12	-0.12	0.94	0.00	0.24	1.01	0.21
Age first treated for AOD	-0.18	15.00	15.89	0.93	0.37	-0.25	-0.49	-0.23	0.20	0.56	-0.02
Alcohol use disorder diagnosis	0.12	1.00	0.78	0.44	0.00	0.10	0.00	-0.42	0.08	1.05	0.12
AOD treatment (# times)	**0.17**	**2.00**	**1.89**	**1.17**	**0.38**	**0.32**	**1.12**	**-0.11**	**0.19**	**0.52**	**0.29**
Days of alcohol use	0.13	41.00	38.78	37.95	0.70	0.26	-0.49	0.30	0.01	0.58	0.21
Days of marijuana use	-0.07	41.00	42.89	41.40	0.25	0.26	-0.77	0.16	-0.29	0.27	0.04
Days of other drug use	**0.31**	**20.00**	**36.33**	**36.55**	**0.41**	**0.21**	**-0.17**	**0.24**	**0.20**	**0.19**	**0.20**
MH service receipt (# times)	**0.17**	**14.00**	**5.50**	**6.48**	**-0.34**	**-0.23**	**-0.09**	**-0.27**	**0.20**	**-0.04**	**-0.09**
Sex (male = 1)	-0.12	0.00	0.33	0.50	0.31	0.00	-0.11	-1.00	-0.54	-0.87	-0.35
Human Capital (k = 16)											
Crime and violence screener, lifetime	-0.09	0.00	3.00	1.41	-0.07	0.40	-0.92	-0.62	-0.04	0.40	0.07
DSM-IV diagnosis: major depression	**0.24**	**1.00**	**1.00**	**0.00**	**0.00**	**0.43**	**0.00**	**0.62**	**0.90**	**0.46**	**0.55**
DSM-IV diagnosis: manic episode	***-0.13***	***1.00***	***1.00***	***0.00***	***0.00***	***-0.42***	***0.00***	***0.00***	***0.03***	***0.00***	***-0.16***
DSM-IV diagnosis: PTSD	***0.22***	***1.00***	***1.00***	***0.00***	***0.31***	***0.00***	***0.00***	***-0.28***	***0.96***	***0.00***	***0.27***
Life satisfaction	0.40	3.17	3.57	0.65	0.67	0.40	-0.05	0.31	0.17	0.38	0.31
MH screens: eating disorder	0.01	1.00	0.56	0.53	0.31	-0.19	0.25	-0.28	-0.16	0.18	-0.11
Negative attitudes towards school	0.21	0.60	0.52	0.35	0	0.32	0.05	-0.08	0.33	-0.09	0.21
Physical health	**0.24**	**5.00**	**4.67**	**1.32**	**0.75**	**0.27**	**-0.28**	**-0.08**	**0.46**	**0.91**	**0.35**
Problem solving - positive orientation	***0.08***	***-0.76***	***0.13***	***0.79***	***0.58***	***0.29***	***-1.34***	***0.10***	***0.11***	***-0.52***	***0.07***
Problem solving - negative orientation	**0.13**	**0.31**	**0.05**	**1.33**	**0.03**	**0.11**	**-0.16**	**0.49**	**-0.41**	**0.26**	**0.00**
Problem solving - rational style	***-0.52***	***0.81***	***0.84***	***0.27***	***0.32***	***0.28***	***-1.71***	***0.39***	***-0.48***	***-0.01***	***-0.05***
Problem solving - impulsivity/carelessness style	0.23	-0.01	0.31	0.69	1.40	0.52	-0.39	0.26	0.08	-0.29	0.28
Problem solving - avoidance style	-0.12	-0.65	0.63	0.36	-0.42	0.00	-0.57	0.28	0.01	0.08	0.00
School attendance, last 12 months (days)	***0.32***	***219.00***	***168.78***	***47.52***	***0.62***	***0.14***	***1.27***	***0.09***	***0.26***	***-0.60***	***0.17***
Stress	0.26	3.00	4.33	1.32	0.06	0.05	0.89	0.81	0.14	0.63	0.26
Substance use expectancies - social benefits	**0.19**	**19.00**	**21.22**	**3.83**	**0.37**	**0.30**	**-0.59**	**-0.30**	**0.33**	**-0.06**	**0.17**

Financial Capital (k = 2)											
Family income level	0.18	4.00	5.88	0.83	0.63	-0.13	**	-0.27	0.42	0.50	0.12
Parental social position score (average)	-0.21	37.00	37.81	12.32	-0.47	0.16	***	-0.12	-0.47	-0.20	-0.14
Social Capital (k = 4)											
Neighborhood social connections	**-0.03**	**2.67**	**2.78**	**0.73**	**0.45**	**-0.15**	**-0.52**	**-0.08**	**-0.40**	**-0.51**	**-0.25**
Social competence	0.07	2.88	3.08	0.32	0.80	-0.04	-0.02	-0.66	-0.22	-0.34	-0.16
Substance approving peer attitudes	**0.24**	**3.08**	**3.13**	**0.45**	**0.42**	**0.55**	**1.24**	**-0.25**	**0.14**	**0.35**	**0.34**
Youth-parent relationship	0.19	64.55	57.78	22.22	1.22	-0.03	0.25	-0.59	0.41	-0.34	0.07
Community Capital (k = 4)											
AA/NA/12 step meeting attendance	***0.36***	***5.00***	***3.44***	***1.74***	***1.52***	***0.01***	***1.18***	***-0.52***	***-0.17***	***-0.29***	***-0.02***
AOD/MH counseling outside school	***-0.06***	***1.00***	***1.00***	***0.00***	***0.00***	***-0.56***	***0.00***	***-0.09***	***-0.80***	***0.00***	***-0.47***
Perceived availability of drugs	**0.15**	**4.60**	**4.44**	**0.61**	**0.00**	**0.41**	**-0.59**	**-0.74**	**0.14**	**0.51**	**0.15**
Youth knowledge of RHS prior to TX	**-0.15**	**0.00**	**0.67**	**0.50**	**0.61**	**-0.32**	**0.59**	**0.10**	**-0.65**	**-1.17**	**-0.38**

Note. *d* = Cohen's *d* standardized mean difference effect size. Strata 8 was removed because of the lack of non-RHS participants in this strata. Bolded values indicate variables were used in model to create classification tree: bolded and italicized values indicate variables were used in final classification tree model.
**Indicates only one value for RHS student: 3; non-RHS mean = 5.13 (SD = 1.64).
***Indicates only one value for RHS student: 51.5; non-RHS mean = 39.56 (SD = 21.97)

Table 6. Stratification Results from Random Forest Covariate Selection Model

Strata	PS	RHS PS M	RHS PS range	nRHS PS M	nRHS PS range	SMD	RHS n	nRHS n
1	0.096	0.28	0.096-0.345	0.29	0.18-0.343	-0.18	18	34
2	0.346	0.39	0.346-0.422	0.39	0.346-0.422	0.04	18	34
3	0.422	0.45	0.422-0.489	0.46	0.426-0.491	-0.20	20	32
4	0.493	0.53	0.494-0.580	0.53	0.493-0.559	0.15	30	22
5	0.581	0.68	0.581-0.851	0.67	0.590-0.806	0.16	34	18

Table 7. Balance Results from Random Forest Covariate Selection Model

	Pre-strata balance *d*	S1	S2	S3 *d*	S4	S5	S1-5 *d*
Control Variables (k = 9)							
Age	**0.22**	**0.01**	**-0.07**	**-0.03**	**0.63**	**0.23**	**0.15**
Age first treated for AOD	**-0.18**	**-0.07**	**-0.34**	**-0.37**	**0.25**	**-0.07**	**-0.12**
Alcohol use disorder diagnosis	0.12	0.22	0.44	-0.15	0.25	0.20	0.19
AOD treatment (# times)	**0.17**	**0.00**	**0.67**	**-0.40**	**0.27**	**0.25**	**0.16**
Days of alcohol use	**0.13**	**0.64**	**0.25**	**0.25**	**0.00**	**-0.55**	**0.12**
Days of marijuana use	**-0.07**	**0.05**	**-0.07**	**-0.10**	**0.08**	**0.09**	**0.01**
Days of other drug use	**0.31**	**0.18**	**0.18**	**0.54**	**-0.19**	**0.23**	**0.19**
MH service receipt (# times)	**0.17**	**-0.33**	**-0.12**	**-0.06**	**0.23**	**0.28**	**0.00**
Sex (male = 1)	-0.12	-0.28	-0.29	-0.43	-0.05	-0.27	-0.27
Human Capital (k = 16)							
Crime and violence screener, lifetime	-0.09	0.38	-0.30	-0.19	-0.26	0.17	-0.04
DSM-IV diagnosis: major depression	0.24	-0.40	0.97	0.90	0.80	0.25	0.50
DSM-IV diagnosis: manic episode	-0.13	-0.19	-0.78	-1.01	-0.39	-0.14	-0.50
DSM-IV diagnosis: PTSD	0.22	0.66	0.40	0.77	0.58	-0.02	0.48
Life satisfaction	**0.40**	**0.42**	**0.10**	**0.17**	**0.09**	**0.63**	**0.28**
MH screens: eating disorder	0.01	-0.05	0.20	0.15	-0.39	0.13	0.01
Negative attitudes towards school	0.21	0.31	-0.13	0.48	0.04	0.29	0.20
Physical health	0.24	0.23	0.12	0.16	0.14	0.36	0.20
Problem solving - positive orientation	**0.08**	**0.17**	**0.01**	**-0.10**	**0.02**	**-0.03**	**0.01**
Problem solving - negative orientation	**0.13**	**-0.41**	**0.32**	**-0.20**	**0.52**	**0.17**	**0.08**
Problem solving - rational style	**-0.52**	**-0.55**	**-0.22**	**-0.49**	**-0.30**	**0.03**	**-0.31**
Problem solving - impulsivity/carelessness style	**0.23**	**0.11**	**0.12**	**0.61**	**-0.07**	**0.14**	**0.18**
Problem solving - avoidance style	**-0.12**	**0.17**	**0.17**	**-0.44**	**-0.20**	**0.02**	**-0.06**
School attendance, last 12 months (days)	**0.32**	**0.25**	**0.43**	**-0.03**	**0.27**	**0.21**	**0.23**
Stress	0.26	0.07	0.61	-0.20	0.75	0.04	0.26
Substance use expectancies - social benefits	0.19	0.60	0.00	0.49	-0.10	-0.11	0.18
Financial Capital (k = 2)							
Family income level	**0.18**	**-0.23**	**0.06**	**0.23**	**-0.15**	**0.78**	**0.14**
Parental social position score (average)	**-0.21**	**-0.04**	**-0.03**	**-0.41**	**0.00**	**-0.56**	**-0.21**
Social Capital (k = 4)							
Neighborhood social connections	**-0.03**	**-0.11**	**-0.02**	**-0.23**	**0.19**	**-0.22**	**-0.08**
Social competence	**0.07**	**-0.21**	**-0.23**	**0.18**	**0.10**	**-0.14**	**-0.06**
Substance approving peer attitudes	**0.24**	**-0.04**	**0.51**	**-0.11**	**0.32**	**-0.02**	**0.13**
Youth-parent relationship	**0.19**	**-0.31**	**-0.09**	**-0.11**	**0.45**	**0.20**	**0.03**
Community Capital (k = 4)							
AA/NA/12 step meeting attendance	**0.36**	**-0.15**	**0.19**	**-0.04**	**0.71**	**0.69**	**0.28**
AOD/MH counseling outside school	-0.06	-0.17	-0.14	-0.22	0.29	-0.57	-0.16
Perceived availability of drugs	**0.15**	**0.28**	**0.21**	**-0.03**	**-0.10**	**-0.05**	**0.06**
Youth knowledge of RHS prior to TX	-0.15	-0.69	-0.15	-0.34	-0.76	-0.23	-0.43

Note. *d* = Cohen's *d* standardized mean difference effect size. Bolded values indicate covariate was used in the model.

Table 8. Random Forest Variable Importance and Balance Characteristics from Covariate Selection Model

	Variable Importance	Pre-strata *d*	Random Forest *d*
Rational problem solving (F3)	6.97	-0.52	**-0.31**
School attendance (days)	5.09	0.32	**0.23**
Impulsivity/carelessness (F4)	4.95	0.23	**0.18**
Positive problem orientation (F1)	4.47	0.08	**0.01**
Substance approving peers	4.42	0.24	**0.13**
Youth-parent relationship	4.37	0.19	**0.03**
Negative problem orientation (F2)	4.34	0.13	**0.08**
Avoidance style (F5)	4.32	-0.12	**-0.06**
Parental social position	4.08	-0.21	-0.21
Other drug use (days)	3.98	0.31	**0.19**
Life satisfaction	3.95	0.40	**0.28**
Neighborhood social connections	3.62	-0.03	-0.08
12-Step attendance	3.40	0.36	**0.28**
Alcohol use (days)	3.40	0.13	**0.12**
Marijuana use (days)	3.17	-0.07	**0.01**
Perceived drug availability	2.81	0.15	**0.06**
MH services (#)	2.76	0.17	**0.00**
Family income	2.75	0.18	**0.14**
Social competence	2.74	0.07	**-0.06**
AOD treatment (#)	2.69	0.17	**0.16**
Age first treated for AOD	1.94	-0.18	**-0.12**
Age	1.66	0.22	**0.15**

Note. *d* = Cohen's *d* standardized mean difference effect size. Bolded values indicate the reduction favors the adjusted *d*.

Table 9. Results from Treatment Effects Analysis after Stratification on Each Covariate Selection Model

													Summary			
Logistic Regression (n = 176)																
	Strata 1 (n = 25)		Strata 2 (n = 38)		Strata 3 (n = 39)		Strata 4 (n = 36)		Strata 5 (n = 38)						95% CI	
	b	SE	b	SE	b	SE	b	SE	b	SE			b	SE	LL	UL
Alcohol use (days)	-7.05	9.85	-0.65	2.05	-0.39	1.19	6.14	5.35	1.16	2.22			0.28	3.68	-6.93	7.49
MJ use (days)	-22.20	20.30	-11.05	9.53	-10.18	7.32	-3.62	13.33	5.53	9.82			-7.34	11.41	-29.71	15.02
Classification Tree (n = 203)																
	Strata 2 (n = 7)*		Strata 3 (n = 84)		Strata 4 (n = 9)		Strata 5 (n = 23)		Strata 6 (n = 59)		Strata 7 (n = 21)				95% CI	
	b	SE	b	SE	b	SE	b	SE	b	SE	b	SE	b	SE	LL	UL
Alcohol use (days)	-0.50	1.96	-4.97	3.64	-3.14	3.49	-1.73	2.08	-0.83	1.24	5.31	9.10	-2.10	3.26	-8.50	4.30
MJ use (days)	13.07	13.07	-19.13	10.29	-3.92	10.59	-32.26	14.36	1.07	2.79	0.72	15.21	-10.91	9.19	-28.92	7.11
Random Forest (n = 221)																
	Strata 1 (n = 44)		Strata 2 (n = 48)		Strata 3 (n = 42)		Strata 4 (n = 44)		Strata 5 (n = 43)						95% CI	
	b	SE	b	SE	b	SE	b	SE	b	SE			b	SE	LL	UL
Alcohol use (days)	2.58	2.88	-2.76	1.73	-0.66	1.57	-13.71	8.30	-0.16	1.69			-2.97	3.23	-9.30	3.36
MJ use (days)	-0.24	8.03	-29.15	9.01	-6.36	6.29	-35.99	14.07	-8.20	9.48			-16.35	9.40	-34.77	2.07

*Indicates model could not calculate numerical derivatives due to discontinuous region with missing value.

CHAPTER V

SUMMARY AND SYNTHESIS

This dissertation aimed to explore adolescent recovery processes using the adolescent recovery capital framework, an ecological model that addresses multiple factors at the individual, interindividual, and community levels that affect the adolescent recovery process. The purpose of doing so was twofold. First, despite the growing body of research on adolescent treatment and recovery, there are still many gaps in understanding how different factors at different ecological levels interact to affect the recovery experience, and ultimately recovery outcomes, throughout the cyclical process of recovery from a substance use disorder. Second, despite the large research base on recovery capital for adults, the recovery capital model had not been adapted or explored for adolescents (Hennessy, 2017), a population with different recovery patterns and processes than older individuals. This issue is not unique to recovery capital, however, and represents a historical trend in studying substance use disorders: often evidence-based adult models of treatment and recovery supports have been simply passed down to adolescents before assessing whether these programs are similarly effective for a new population. Thus, this dissertation is the first to explore an adapted model of recovery capital and its potential application to adolescents using different exploratory methods, while also highlighting the use of these methods as potential resources for those interested in applying them to similar types of complex social science research questions.

A Latent Class Exploration of Adolescent Recovery Capital

The first empirical chapter of this dissertation demonstrated that by using the recovery capital framework among a national sample of adolescents in need of substance use treatment, different and complex patterns of recovery capital emerged. That is, there was not simply a three-group model of low, medium, and high recovery capital, but instead qualitatively distinct subgroups of youth with different recovery capital characteristics. These results highlighted the diverse nature of adolescent substance use disorders as well as some demographic characteristics that were predictive of class membership, which may provide useful information for practitioners working with youth as well as researchers studying the treatment and recovery process.

The results also demonstrated that the different recovery capital domains and the variables that represented them were an appropriate initial step in beginning to use the recovery capital framework to study adolescent recovery processes. The latent class analysis was useful in identifying these qualitatively distinct groups; however, future research with more robust variables and with longitudinal data and analysis methods (e.g., latent transition analysis) might generate different or more nuanced results, as this study was mostly limited to binary, cross-sectional indicators of capital. In addition, some of the recovery capital factors included in this study, such as religious beliefs, were fairly crude indicators of more complex constructs, thus warranting further attention in future research with adolescent recovery populations.

Adolescent Recovery Capital and Recovery High School Attendance: An Exploratory Data Mining Approach

The second empirical chapter of the dissertation further explored the recovery capital framework among a smaller sample of youth who had already had at least one substance use treatment episode. This paper used data from an observational study and incorporated a number of additional measures of recovery capital domains enabling a more robust assessment of recovery capital resources than were available in the first empirical paper. The aim of this paper was to better understand what variables were predictive of attendance at a recovery high school, and if the recovery capital framework was useful in identifying the key predictors of recovery school attendance, based on two primary assumptions. The first assumption was that as a school specifically designed for youth in recovery, a recovery high school is a developmentally-appropriate, community recovery capital resource. The second assumption was based on arguments stated by the authors of the original recovery capital framework for adults: that individuals with some capital are more likely to build more capital and that building capital for recovery involves a dynamic process over time (Cloud & Granfield, 2008; Granfield & Cloud, 1999; White & Cloud, 2008). These arguments would suggest that those with many sources of recovery capital at one point in time, such as individual motivation to remain in recovery, access to treatment, a family supportive of recovery, and a sober peer network might be more likely to end up in a recovery high school because of the potential access through their family, peers, or treatment center and personal motivation to be part of a sober community.

The results from the different methods, however, demonstrated that recovery high schools appear to meet the needs of a diverse group of youth who may have a mix of recovery capital and not necessarily the highest levels of recovery capital resources. Indeed, youth who attended non-recovery high schools may have some additional human capital resources that enabled them to return to a non-recovery focused environment; however, given that the non-recovery high school group was actually comprised of students who attended traditional, alternative, charter or some other form of schooling, some of these students may simply have found other types of schools to address their recovery needs that were more convenient or more enticing than recovery high schools. Thus additional research with more clearly distinct non-recovery high school categories might highlight more nuanced understandings of school choices posttreatment among youth in recovery.

The use of the exploratory data mining techniques in this paper, i.e., SEARCH, classification trees, and random forests, also demonstrated the ability to capture interactions in recovery capital predictors of recovery high school attendance that were missed in the logistic regression analysis. Although interactions between variables could have been tested in the regression analysis, there were such a large number of potential interactions and little empirical reason to include one interaction over another in the traditional analysis, that the exploratory methods combined with the recovery capital model generated new areas for future research in adolescent recovery to explore. For example, one area future research could address is why the different problem solving orientations and styles interacted with mental health, 12-Step group attendance, and history of school attendance to predict attendance at a recovery high school.

Covariate Selection for Propensity Score Estimation: A Comparison of Exploratory Approaches and Application to Adolescent Recovery

The third empirical dissertation chapter used the techniques and sample from paper two to demonstrate the utility of classification tree and random forest approaches in choosing covariates for the estimation of propensity scores for use in treatment effects analysis. Similar to paper two, these data mining methods were also compared to traditional logistic regression approaches. The results demonstrated that with a smaller sample size, the logistic regression may be the safest approach to use as it balanced 97% of the included covariates. However, the results also demonstrated that one limitation to the logistic regression approach is that missing data can severely affect the resulting sample size after pruning out participants from the region of common support. This issue was remedied in the random forest method, as this approach had good balance on most covariates and also maintained most of the existing sample. In addition, because it is a recursive partitioning approach, the random forest method accounts for potential variable interaction, resulting in reduced burden on the researcher to test all potential interactions for the estimation of propensity scores.

Implications and Future Directions

The findings from these three empirical papers suggest several important directions for future research. First, overall the adolescent recovery capital framework was useful in considering the many factors, especially at the individual, microsystem, and mesosystem levels that affect the recovery process. Individual level factors from human and community recovery capital, especially, were important in predicting attendance at a recovery high school and both human and social recovery capital factors such as religious beliefs and close friend and parent support were key distinguishers between adolescents in need of substance use treatment.

However, there remain some gaps with the recovery capital model, in particular attention to broader ecological domains such as understanding the actual availability of tangible resources from particular microsystems beyond from the individual perspective, how microsystems (e.g., peer, family, and treatment) interact to support recovery, or the larger context of the macrosystem. Unfortunately, as an initial exploratory study of the applicability of recovery capital for adolescents, the adolescent recovery capital model was not adapted to the degree that would allow for these relationships to be directly tested and the data used here would not have addressed expansions across ecological levels in this way. Indeed, studies using the original adult-focused recovery capital models have not attended to these issues: for example, in their exploration of the recovery capital model for adults, Zschau and colleagues (Zschau, Collins, Lee, & Hatch, 2016) critiqued original models as blurring the distinction between individual-level and network-based components of recovery. The authors then measured network-based recovery capital by, for example, measuring perceived financial capital resources available through participant's networks such as assessing whether participants felt their friends were willing to lend money or a car if needed. There is also one large multi-site study beginning in Europe dedicated to further understanding recovery capital which includes a focus on how policy and practice change has influenced addiction stigma, access to services, and reintegration (D. Best, personal communication, February 26, 2017). Both of these examples highlight gaps in the current research and lead to a natural extension of the original recovery capital model to understanding nuances and broader ecological domains, which could be measured and tested in the future, especially for adolescent recovery populations.

The studies in this dissertation also illustrate additional areas for research including addressing adolescent recovery capital and changes over time to quantitatively assess the applicability of growth capital, that is, how the recovery process affects the generation (or reduction) of recovery capital resources over time. Overall, having a better understanding of how recovery capital affects initial and later recovery attempts and identifying some of the most important recovery capital indicators of sustained change would be useful for both researchers and practitioners who study and work with youth in recovery. Future research in this area could also focus on whether the generation of recovery capital is different for youth with problems with multiple substances or across diagnoses of disorders with specific substances (e.g., heroin versus marijuana use disorders) as these nuances were not explored in this analysis. Additionally, as has been done with some adult samples, qualitative research could further explore different aspects of recovery capital domains and their interaction. This might be especially relevant to gain a better understanding of the key indicators of cultural recovery capital, one element of the community recovery capital domain with the smallest research base, especially for adolescents.

There are also areas worth further exploration for researchers interested in utilizing the exploratory methods utilized here. For example, for more exploratory research questions that involve identifying potential relationships between complex constructs, data mining techniques can provide additional nuances for future research to explore and test. In addition, the use of non-parametric techniques for covariate selection in propensity score estimation is still underutilized, especially in the social sciences. Additional research is needed for investigating the best technique(s) among the many available options, especially to identify methods more suitable for smaller sized samples and for diverse outcomes.

To conclude, adolescent substance use disorders are a major public health problem, which affect healthy development and quality of life among youth and cost families as well as the health services system thousands of dollars every year in the United States. Given the cyclical nature of adolescent addiction and recovery, gaining a better understanding of the recovery process within context, and figuring out ways to more successfully intervene and reduce relapse rates should be a high priority among researchers and practitioners. Thus, the adapted adolescent recovery capital model is a useful preliminary framework that draws attention to individual, interindividual, and community factors to more comprehensively study the process of recovery and is one important step to moving the addiction treatment and recovery community forward in better understanding this complex health issue.

References

Cloud, W., & Granfield, R. (2008). Conceptualizing recovery capital: Expansion of a theoretical construct. *Substance Use & Misuse*, *43*(12-13), 1971-86. doi:10.1080/10826080802289762

Granfield, R., & Cloud, W. (1999). *Coming clean: Overcoming addiction without treatment.* New York: New York University Press.

Hennessy, E. A. (2017). Recovery Capital: A systematic review of the literature. *Addiction Research and Theory*, in press.

White, W., & Cloud, W. (2008). Recovery capital: A primer for addictions professionals. Counselor, 9(5), 22-27.

Zschau, T., Collins, C., Lee, H., & Hatch, D. L. (2016). The hidden challenge: Limited recovery capital of drug court participants' support networks. *Journal of Applied Social Science*, *10*(1), 22-43.