Comparing Short-Term Outcomes of Three Problem Gambling Treatments: A Multi-Group Propensity Score Analysis

Adam David Soberay
University of Denver

Follow this and additional works at: https://digitalcommons.du.edu/etd
Part of the Counseling Commons, Counseling Psychology Commons, and the Experimental Analysis of Behavior Commons

Recommended Citation
https://digitalcommons.du.edu/etd/615

This Dissertation is brought to you for free and open access by the Graduate Studies at Digital Commons @ DU. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of Digital Commons @ DU. For more information, please contact jennifer.cox@du.edu,dig-commons@du.edu.
COMPARING SHORT-TERM OUTCOMES OF THREE PROBLEM GAMBLING TREATMENTS: A MULTI-GROUP PROPENSITY SCORE ANALYSIS

A Dissertation
Presented to
The Faculty of the Morgridge College of Education
University of Denver

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

by
Adam D. Soberay
March 2015
Advisor: Dr. Antonio Olmos
Abstract

This study applied a multi-group form of propensity score analysis to the study of outcomes related to problem gambling treatment. Across various treatment settings, it is often unfeasible or unethical to randomly assign participants to different treatment conditions, particularly when one of the conditions involves not receiving treatment. Additionally, evaluative practices often involve assessing outcomes from a primarily treatment focused setting, in which case clients are likely not randomly assigned to treatment. Consequently, where randomization does not exist, methods such as propensity score matching need to be implemented to separate what part of the observed outcomes is attributable to treatment and what part may be due to preexisting differences between the comparison groups. Traditional propensity score matching procedures involve matching and comparing across two groups, typically a treatment and a control group. This study applied newly developed methods for matching participants on propensity scores across three groups.

This study uses archival treatment data to compare three psychotherapeutic problem gambling treatments (cognitive-behavioral therapy, solution-focused brief therapy, and time-limited dynamic psychotherapy) where outcomes were likely influenced by self-selection of form of therapy. Specifically, this study looked at whether
participants improved their psychosocial functioning through five weeks of treatment, and, if so, are the three forms of treatment equally effective.

The results of this study support the utility of multi-group propensity score matching procedures. Covariate imbalance was improved through each of the four implemented matching procedures, though two of the matching procedures (caliper matching and 3:2:n matching) were more effective in reducing bias. The matching procedures also indicate that there may be a difference between treatment effects that was not observed through an unmatched analysis. The matching procedures consistently estimated the treatment effect for cognitive-behavioral therapy to be greater than that of the time-limited dynamic psychotherapy. This difference was found to be statistically significant on two of the four matching methods. Limitations of this study and recommendations for future research are also discussed.
Table of Contents

Chapter One: Introduction ...........................................................................................................1
  Background..............................................................................................................................1
  Statement of the Problem.........................................................................................................5
  Purpose of the Study ...............................................................................................................6
  Research Questions..............................................................................................................7
  Definitions.............................................................................................................................7
  Review of the Literature ........................................................................................................8
    Propensity Score Analysis ....................................................................................................8
    Problem Gambling .............................................................................................................18
  Summary...............................................................................................................................52

Chapter Two: Method..................................................................................................................54
  Participants..............................................................................................................................54
  Procedure...............................................................................................................................54
  Instruments.............................................................................................................................58
    URICA ................................................................................................................................58
    NODS.................................................................................................................................59
    OQ-45 .................................................................................................................................60
    WAI-S ................................................................................................................................61
    HANDS ...............................................................................................................................62
    MDQ ..................................................................................................................................62
    CDGAD ..............................................................................................................................63
    SPRINT-4 ...........................................................................................................................64
  Analytic Strategy...................................................................................................................64
    Missing Values Analysis ....................................................................................................64
    Research Question One ......................................................................................................65
    Research Question Two ......................................................................................................66
    Research Question Three .................................................................................................66

Chapter Three: Results..............................................................................................................70
  Missing Values Analysis.........................................................................................................70
  Characteristics of the Sample...............................................................................................72
  Assessing Treatment Selection ............................................................................................72
  Comparing Treatment Selection Groups ...............................................................................73
  Comparing Overall Treatment Groups ..............................................................................75
  Relationships Between Covariates and Outcomes .............................................................76
  Unmatched Analysis ............................................................................................................77
  Propensity Score Analysis ....................................................................................................78
    Specifying the Propensity Score Model ............................................................................78
    Covariate Balance Across Propensity Score Estimation Models ....................................79
    Maximum Treat Matching ...............................................................................................80
    Caliper Matching ................................................................................................................86
2:1:n Matching .......................................................................................................................... 91
3:2:n Matching .......................................................................................................................... 95
Comparing Matching Procedures ............................................................................................. 100

Chapter Four: Discussion .............................................................................................................. 102
  Overview .................................................................................................................................. 102
  Research Question One .......................................................................................................... 104
  Research Question Two .......................................................................................................... 107
  Research Question Three ....................................................................................................... 109
  Limitations ............................................................................................................................ 117
  Recommendations for Future Research ................................................................................. 119

References ....................................................................................................................................... 122

Appendix A .................................................................................................................................... 170
Appendix B .................................................................................................................................... 175
Appendix C .................................................................................................................................... 180
Appendix D .................................................................................................................................... 185
List of Tables

Table 1  Description of Therapy Choices .................................................................57
Table 2  Missing Values...........................................................................................71
Table 3  Observed Therapy Selections .....................................................................73
Table 4  Baseline Covariates Across Treatment Selection Groups .........................74
Table 5  Comparing Covariate Balance Across Treatment Groups .......................76
Table 6  Bivariate Correlations Between Baseline Characteristics and Outcome Measure ...........................................................................................................77
Table 7  Pairwise Differences of Covariates Prior to Matching ...............................79
Table 8  Unmatched Participants Within Maximum Treat Matching .......................81
Table 9  Covariate Values Following Maximum Treat Matching ..............................81
Table 10 Inferential Testing of Maximum Treat Post-Matching Covariate Balance .........................................................................................................................82
Table 11 Pairwise Differences of Covariate Balance Following Maximum Treat Matching ........................................................................................................83
Table 12 Post-Hoc Comparisons of Outcome Following Maximum Treat Matching .................................................................................................................84
Table 13 Sensitivity Analysis for CBT-TLDP Comparison Following Maximum Treat Matching ........................................................................................................86
Table 14 Unmatched Participants Within Caliper Matching .....................................86
Table 15 Covariate Values Following Caliper Matching ...........................................87
Table 16 Inferential Testing of Caliper Post-Matching Covariate Balance ................87
Table 17 Pairwise Differences of Covariate Balance Following Caliper Matching ......88
Table 18 Post-Hoc Pairwise Comparisons of Outcome Following Caliper Matching .................................................................................................................89
<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 19</td>
<td>Sensitivity Analysis for SFBT-TLDP Comparison Following Caliper Matching</td>
<td>91</td>
</tr>
<tr>
<td>Table 20</td>
<td>Sensitivity Analysis for CBT-TLDP Comparison Following Caliper Matching</td>
<td>91</td>
</tr>
<tr>
<td>Table 21</td>
<td>Unmatched Participants Within 2:1:n Matching</td>
<td>92</td>
</tr>
<tr>
<td>Table 22</td>
<td>Covariate Values Following 2:1:n Matching</td>
<td>92</td>
</tr>
<tr>
<td>Table 23</td>
<td>Inferential Testing of 2:1:n Post-Matching Covariate Balance</td>
<td>93</td>
</tr>
<tr>
<td>Table 24</td>
<td>Pairwise Differences of Covariate Balance Following 2:1:n Matching</td>
<td>93</td>
</tr>
<tr>
<td>Table 25</td>
<td>Unmatched Participants Within 3:2:n Matching</td>
<td>96</td>
</tr>
<tr>
<td>Table 26</td>
<td>Covariate Values Following 3:2:n Matching</td>
<td>97</td>
</tr>
<tr>
<td>Table 27</td>
<td>Inferential Testing of 3:2:n Post-Matching Covariate Balance</td>
<td>97</td>
</tr>
<tr>
<td>Table 28</td>
<td>Pairwise Differences of Covariate Balance Following 3:2:n Matching</td>
<td>98</td>
</tr>
<tr>
<td>Table 29</td>
<td>Comparing Matching Procedures</td>
<td>101</td>
</tr>
</tbody>
</table>
### List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1</td>
<td>Covariate Balance from the Three Propensity Scores Estimation Models</td>
<td>80</td>
</tr>
<tr>
<td>Figure 2</td>
<td>Boxplots of OQ-45 Change for Each Group After Maximum Treat Matching</td>
<td>84</td>
</tr>
<tr>
<td>Figure 3</td>
<td>Boxplots of Pairwise Differences of OQ-45 Change After Maximum Treat Matching</td>
<td>85</td>
</tr>
<tr>
<td>Figure 4</td>
<td>Boxplots of OQ-45 Change for Each Group After Caliper Matching</td>
<td>89</td>
</tr>
<tr>
<td>Figure 5</td>
<td>Boxplots of Pairwise Differences of OQ-45 Change After Caliper Matching</td>
<td>90</td>
</tr>
<tr>
<td>Figure 6</td>
<td>Boxplots of OQ-45 Change for Each Group After 2:1:n Matching</td>
<td>94</td>
</tr>
<tr>
<td>Figure 7</td>
<td>Boxplots of Pairwise Differences of OQ-45 Change After 2:1:n Matching</td>
<td>95</td>
</tr>
<tr>
<td>Figure 8</td>
<td>Boxplots of OQ-45 Change for Each Group After 3:2:n Matching</td>
<td>99</td>
</tr>
<tr>
<td>Figure 9</td>
<td>Boxplots of Pairwise Differences of OQ-45 Change After 3:2:n Matching</td>
<td>99</td>
</tr>
</tbody>
</table>
Chapter One: Introduction

Background

Quasi-experimental or observational methods are utilized frequently in social science research, as randomized experimental methods are often unfeasible or unethical. For instance, it may be unethical in certain circumstances to randomly assign participants to a no-treatment control group, particularly when denying treatment would put the individuals at risk. Consequently, such studies often involve self-selection of participants into treatment groups. This self-selection comes at the price of the balancing of confounding variables across groups that is associated with randomization (Heckman, Ichimura, Smith, & Todd, 1998). In addition, when self-selection occurs systematic differences may exist between the groups that may confound the relationship between treatment and outcome. Therefore, treatment assignment is likely to violate the ignorable treatment assignment assumption (ITAA), which states that assignment to any treatment condition is independent of outcome (Rosenbaum, 1984). Consequently, violation of the ITAA constitutes a selection bias that can pose a significant threat to the internal validity of the research findings. Moreover, this selection bias, if left unchecked, weakens the ability to draw conclusions on the causal counterfactual, which is the outcome of participants had they been in a different treatment condition.
Traditional approaches to controlling for these potentially confounding influences often involve multivariate analyses, which statistically control for the influence of covariates on the outcome. This approach will lead to unbiased estimates of the treatment effect so long as the ITAA is not violated and the model is appropriately specified (Rosenbaum, 1994). However, the ITAA is analogous to the ordinary least squares regression assumption of independence of the error term from all independent variables. When treatment is included in a multivariate analysis as a dichotomous independent variable, this assumption of independence is violated when ITAA does not hold, because it results in dependence between the error term and an independent variable. Consequently, estimates of treatment effect will be biased (Guo & Fraser, 2010).

Therefore, this traditional method of statistical control of covariates is often inappropriate for quasi-experimental and observational designs. Propensity score analysis was developed as a means to address such a problem (Rosenbaum & Rubin, 1983). It is a set of procedures that allow the researcher to account for the differences between groups that are typically associated with observational and quasi-experimental research.

Within the general public, and even within the treatment field, addiction is a concept that virtually everyone has an opinion on, yet where agreement is lacking regarding what it truly is. Accordingly, the varying conceptions of addiction reflect the underlying assumptions regarding the factors contributing to the origins, continuation, and elimination of the addictive behavior. There also does not seem to be consistent agreement on which behavioral disorders can and cannot constitute an addiction.
The National Association of Alcoholism and Drug Abuse Counselors defines addiction as a brain disease (National Association of Alcoholism and Drug Abuse Counselors, 2013). Such a conceptualization reflects a growing body of research that shows the role of, and effect on the brain associated with addiction.

The American Society of Addiction Medicine also focuses on the role that brain circuitry plays in addictive behaviors (American Society of Addiction Medicine, 2013). They indicate that addiction brain circuitry manifests itself in biological, psychological, and social dysfunction, which leads an individual to seek reward and/or relief through their addictive behavior. This definition suggests that addiction has physiological, cognitive, and emotional, as well as social components. Kranzler and Li (2008) indicates that due to the multiple components of addiction, it crosses disciplines, including biology, psychology, sociology, and pharmacology.

Various twelve-step organizations, such as Alcoholics Anonymous and Gamblers Anonymous, define addiction as an illness (Alcoholics Anonymous, 1972; Gamblers Anonymous, 2013). These organizations hold that the addictive behavior will continue to progressively worsen unless the behavior is completely abstained from. This view is in contrast to others that hold that the addictive behavior can become responsibly moderated.

Addiction is sometimes viewed as an umbrella term that covers disordered usage of drugs and alcohol as well as other behaviors such as gambling, shopping, and sexual activity. However, in other instances addiction can be more narrowly defined as only pertaining to drugs and alcohol. The American Psychological Association defines
addiction as a condition in which a person must take a substance to avoid psychological
and physical withdrawal (American Psychological Association, 2013). This
conceptualization also includes issues such as a physical dependence, as well as building
up a tolerance to the substance, which are concepts that are generally associated with
alcohol and substance abuse. However, issues of withdrawal have also been associated
with other behavioral addictions, such as pathological gambling (Rosenthal & Lesieur,

The American Psychiatric Association’s Diagnostic and Statistical Manual of
Mental Disorders (DSM) through its iterations has shown an evolution of the
conceptualization of addictive disorders. Prior to the latest edition, the DSM-5, the DSM
has emphasized the dependence aspects of addictive disorders; in fact, it used the term
dependence in lieu of the term addiction (Kranzler & Li, 2008). However, the fifth
edition of the DSM has seen significant changes to these diagnoses. There has been a
growing disenchantment with the choice of the term dependence in the DSM, as physical
dependence and its associated tolerance and withdrawal is a phenomenon often also
found with appropriate usage of medication (Courtwright, 2011; O’Brien, Volkow, & Li,
2006; O’Brien, 2011). Moreover, this conceptualization largely excluded non-substance
behavioral addictions, such as gambling. Among the major changes in the DSM-5 was
the introduction of a new category of behavioral addictions, to which gambling was the
first addition (American Psychiatric Association, 2013). This move is often viewed as the
first step in the inclusion of other behavioral addictions, such as internet addiction
(Holden, 2010).
When viewing addiction as a phenomenon that transcends alcohol and substance misuse researchers have identified elements that seem to be hallmarks of addiction. Shaffer (2011) identified three aspects of a behavior that go towards identifying it as an addiction. According to Shaffer, for an addiction to exist, the following three characteristics must be present: the addictive behavior is motivated by emotions, the behavior is continued despite negative consequences, and the person is experiencing some level of loss of control over their behavior.

**Statement of the Problem**

Until very recently, the implementation of propensity score matching has dealt with matching across two groups. This lent itself to studies that involved the comparison of a treatment and a control group. However, these methods were not fully conducive to the comparison of multiple treatment groups. Recent advances in statistics have provided approaches to generate propensity score matching procedures that can address comparisons of more than two groups (Bryer, 2013; Rassen et al., 2013). As these methods are relatively new, the current literature lacks examples of practical implementations of these methods.

Across the social sciences, research related to the treatment of problem gambling continues to lag far behind the evidence base surrounding other addictive disorders, particularly alcohol and substance abuse. Moreover, gambling addiction is a problem that can have significant adverse consequences, individually, interpersonally, as well as socially. Therefore, treatment must be available that utilizes evidence-based practices, so as to ensure that this problem is addressed in an efficacious manner.
Purpose of the Study

As an application of a multi-group propensity score analysis, this study seeks to compare the effectiveness of three common psychotherapeutic orientations, solution focused brief therapy (SFBT), time limited psychodynamic therapy (TLPD), and cognitive-behavioral therapy (CBT) for the treatment of gambling problems. CBT treatment methods have well-documented effects in the treatment of gambling problems. However, very little empirically based knowledge exists on the application of SFBT and TLPD in this area. Consequently, in addition to adding to the literature regarding a three-group propensity score analysis, this study also seeks to add to the literature by determining whether these relatively unstudied problem gambling treatments can be equally as effective as the well-studied cognitive-behavioral approach.

The first component of this study will explore whether various factors help to explain the particular therapeutic selections of the participants. The study will include the following variables in relation to treatment selection: age, gender, stages of change, co-occurring psychological disorders (depression, post-traumatic stress disorder, anxiety, and mood disorder), initial level of psychosocial functioning, and severity of gambling problems. Moreover, it will explore whether the problem gamblers in this study are disproportionately selecting any of the treatment options.

The second component will analyze whether each of the treatments is effective and the relative effectiveness of each treatment. Particularly, the study will examine short-term improvement in psychosocial functioning through the first five sessions of treatment.
Research Questions

The following three research questions will be addressed:

1) Is therapeutic selection explained by various demographic and psychological indicators? Are there certain subgroups of problem gamblers that may prefer one type of therapy relative to others?

2) Overall, do problem gambling therapy clients have particular therapeutic preferences? Are they selecting each of the three therapies in equal proportions?

3) Are these three psychological interventions for problem gambling effective in reducing overall psychosocial distress? Are the interventions equally effective?

Definitions

Propensity scores are the conditional probabilities of receiving a particular level of treatment given a set of pre-treatment covariates. Propensity score analysis refers to any of a set of statistical procedures involving the utilization of propensity scores in the calculation of a treatment effect.

Pathological gambling is a disorder from the Diagnostic and Statistical Manual IV-Text Revisions (DSM-IV TR) and is assigned according to a set of criterion therein (American Psychiatric Association, 2000). Gambling disorder is the diagnosis in the current, 5th edition of the DSM (American Psychiatric Association, 2014). Problem gambling is defined as gambling behavior that is having significant adverse consequences on the psychosocial functioning of the individual engaging in the behavior. Therefore, problem gambling is a term that encompasses both diagnosable as well as sub-clinical levels of disordered gambling behavior.
For the purposes of this paper, *addiction* will be viewed as a psychological disorder with multiple epidemiological influences. It will also be viewed as having psychological, interpersonal, physical, and social consequences. It is also viewed as a concept that applies to behaviors beyond substance abuse, which share common characteristics, as outlined by Shaffer (2011). Further, addiction will be held as a treatable disorder. Therefore, *gambling addiction* is used to describe the underlying psychological condition that leads to disordered gambling behavior.

*Psychotherapy* refers to any of various forms of treatment consisting primarily of talk-based interventions for psychological disorders, including problem gambling.

**Review of the Literature**

**Propensity Score Analysis**

Propensity score matching is a set of procedures designed to statistically counteract the selection bias associated with non-experimental research (Rosenbaum & Rubin, 1983). The procedure generally consists of two or three steps (Guo & Fraser, 2010). The first step in either case is the estimation of the propensity scores, which is typically achieved through logistic regression. The logistic regression produces a conditional probability that an individual will be in the treatment, as opposed to the comparison group. These probabilities become the propensity scores. Propensity scores reduce multidimensional covariates into a single value. Therefore, a treatment and a comparison participant with comparable propensity scores are considered to be balanced on the vector created by the combination of the covariates. It should also be noted that in addition to logistic regression, propensity scores can also be estimated through
discriminant function analysis or by OLS regression where the propensity score is the predicted value when a binary, dummy outcome representing treatment group is regressed upon the observed covariates (Holmes, 2014).

A critical component of this step of the process is adequate specification of the logistic regression model. In other words, the selection of variables to include in this model plays a large role in determining how well the propensity scores will balance groups. Data simulations have shown that propensity score matching can exacerbate the bias produced by unmeasured covariates, so long as these unmeasured covariates vary independently of those being measured and included in the propensity score estimation model (Brooks & Ohsfeldt, 2013). There is some debate as to whether to include in the propensity score estimation model all measured covariates, only those related to treatment selection, only those related to outcomes, or only those related to assignment and outcome (Austin, 2011). However, it has been demonstrated that propensity scores can be estimated more effectively when including variables associated with the outcome variable, and not only those that are believed to simultaneously impact selection and outcome (Cuong, 2013). Further, it has also been argued that a characteristic that varies between treatment groups and has no influence over the outcome(s) will not bias the estimates of treatment effect (Tanner-Smith & Lipsey, 2014).

After estimating the propensity scores, determining if balance has been achieved is another important step in the process. There are a number of ways to check balance, including: significance tests, standardized differences, percent reduction, and graphical methods (Holmes, 2014). Significance tests, for continuous variables, include F and t
tests, or the Mann-Whitney test can be used if the distribution is non-normal. For categorical covariates, chi-square tests of independence can determine if significant differences exist. Standardized mean differences of each covariate between groups can reveal which covariates significantly differ between groups, as well as be used to determine whether or not post-matched groups are balanced on the observed covariates (Kuss, 2013). Additionally, an advantage of assessing standardized differences is the ability to determine relative imbalance among the covariates. Percent reduction deals with calculating the proportion of the differences in observed covariates between groups that is reduced through matching. Finally, graphical procedures, such as Q-Q plotting determine balance through comparing the distribution of the propensity scores across groups.

If through matching the systematic differences in the baseline covariates have been removed then the propensity score model is considered to be adequately specified (Austin, 2011). Sensitivity analysis can also be conducted to search for hidden bias in the propensity score estimation that may indicate the omission of a significant covariate (Rosenbaum, 2002). If satisfactory balance is not achieved, Holmes (2014) recommends adding relevant covariates or assessing and correcting for non-linear relationships, which may require data transformations.

For the three-step procedures, the second step is the selection and implementation of a matching procedure. Once propensity scores have been estimated, there are various categories of strategies for matching participants once the propensity scores have been
calculated, including: greedy matching, optimal matching, and fine balancing (Guo & Fraser, 2010).

Greedy matching procedures include: Mahalanobis metric, nearest neighbor, and caliper matching (Guo & Fraser, 2010). Mahalanobis metric matching procedures predate propensity score matching (Cochran & Rubin, 1973). These initial procedures involved matching participants based on Mahalanobis distances, or the multivariate distance of a participant from a common point. With the introduction of propensity scores, Mahalanobis metric matching was expanded to include propensity scores as an additional covariate. Nearest neighbor matching procedures involve matching participants across groups based on minimum absolute difference in propensity scores. These procedures can either involve a 1:1 match, or a 1:n match where a participant in the treated group can be matched to multiple controls. Caliper matching extends nearest neighbor procedures by placing a constraint on the maximum allowable distance between propensity scores. In traditional nearest neighbor approaches, a match will still be made even if there is a relatively large discrepancy between propensity scores so long as it is the closest possible match. This combination of nearest neighbor and caliper matching begins by randomly ordering treatment and comparison participants. Then, in order, treatment participants are matched with the closest control so long as the match is within the predetermined caliper. The caliper can vary, though 0.25 standard deviations of the propensity scores is a typical value, with tighter calipers resulting in reduced bias but increased unmatched participants (Lunt, 2014). Further, an analysis of propensity score matching among three groups compared calipers in increments of 0.1 and found that 0.2 was optimal in yielding
estimates of treatment effect (Wang et al., 2013). Related to this idea of increasing the precision of matches, it has been shown that the more precisely propensity score matching balances groups on the observed covariates, the greater the risk for loss of data, which can be both a statistical and financial issue (Golinelli, Ridgeway, Rhoades, Tucker, & Wenzel, 2012). This nearest neighbor within a caliper procedure has gained popularity as the matched samples can be analyzed using virtually any multivariate statistical procedure. In fact, a major advantage of these procedures is the ability to implement relatively simple post-matching analyses, such as a paired samples t-test, which further makes these procedures accessible to most clinical researchers (Cotton, Cuerden, & Cook, 2011). Finally, greedy matching procedures can also include a combination of the nearest neighbor within a caliper and Mahalanobis metric matching procedures.

One of the primary drawbacks to the greedy matching procedures is the requirement of significant overlap in propensity scores between the treated and comparison participants (Heckman, Ichimura, Smith, & Todd, 1996). Without this common support region of scores there can be a significant loss in data through unmatched participants.

As opposed to greedy matching procedures, which match each treatment participant sequentially with its closest match without regard for future matches, optimal matching procedures do not take this sequential approach, which can result in more precise overall matching (Rosenbaum, 1989). In other words, optimal matching procedures are focused upon creating matches that contribute to the minimum total distance in propensity scores, as opposed to minimum distances within individual
matches. There are three general strategies to optimal matching: pair matching, variable matching, and full matching. Similar to nearest neighbor matching, pair matching and variable matching create 1:1 and 1:n matches, respectively. With full matching, not only are treated participants matched to one or more control participants, but control participants are matched to one or more treated participants (Rosenbaum, 2002).

Fine balancing is a relatively new approach to matching (Rosenbaum, Ross, & Silber, 2007). Unlike greedy and optimal matching, which use a single propensity score to balance groups, fine balancing involves balancing groups based on a nominal variable. Fine balancing is a procedure that is used in conjunction with propensity score matching, and involves balancing a particular covariate across groups rather than between individual participants.

In addition to the aforementioned matching procedures, there are also nonparametric approaches to propensity score matching. A primary advantage of nonparametric statistical procedures is their ability to provide accurate statistical estimations when the assumptions regarding the normality of the underlying populations cannot be upheld (Hollander, Wolfe, & Chicken, 2013). Kernel-based matching is an approach to propensity score matching that utilizes nonparametric regression (Heckman, Ichimura, & Todd, 1998). Kernel-based matching produces one-to-many matches, which involves weighting matches based on closeness of propensity scores. Kernel-based matching has been shown to be equally effective as 1:1 matching when there are sufficient control participants but more effective when number of control participants are equal to or less than the number of treated participants (Berg, 2011). Semiparametric
propensity score estimation techniques have also demonstrated value when assumptions related to normality are violated (Lehrer & Kordas, 2013). Random forest classification is a nonparametric procedure designed to predict group membership based on a series of covariates, which has successfully been applied to propensity score estimation (Watkins et al., 2013).

The final of the three steps is the post-matching statistical analysis. As discussed earlier, a primary benefit to greedy matching is that multivariate analyses may be applied to determine whether treatment effects truly exist after balancing the groups on the observed covariates. Another analytical approach using propensity scores involves the stratification of participants. Through this approach, participants are divided into strata, often quintiles, based on propensity score. Mean differences are then calculated within each stratum, which are then averaged together to estimate the overall treatment effect. With optimal matching, unlike greedy matching, special considerations must be taken as to which post-matching analyses are appropriate (Rosenbaum, 1989). For full and variable matching procedures, the Hodges-Lehmann aligned rank test can be used to estimate the average treatment effect. Finally, optimal match pairs can be subject to a special type of regression adjustment whereby the difference on the outcome variable between a matched treatment and comparison participant is regressed on the difference between the observed covariate(s). There is also an analytic technique that combines the Hodges-Lehman procedure with the regression adjustment (Fraser & Guo, 2010).

However, not all propensity score techniques involve a matching procedure. For the two-step procedures, statistical analyses are conducted using propensity scores as
regression weights (Freedman & Berk 2008). However, if the causal model and/or the propensity score estimation model are not adequately specified then this procedure can be counterproductive through increasing random error.

Another facet of propensity score matching that has been explored is the usage of bootstrapping procedures. A bootstrap procedure, which uses the mean propensity score across bootstrap samples, has been shown to reduce the bias resulting from multiple propensity score matching procedures (Bai, 2013).

In terms of practical applicability, propensity score matching has demonstrated the ability to result in relatively unbiased estimates of treatment effect even for sample sizes as small as forty (Pirracchio, Resche-Rigon, & Chevret, 2012). Therefore, this procedure can be conducted in various settings where large amounts of data are not available.

If conducted properly, propensity score analyses can be very beneficial to quasi-experimental and observational research. It is a sophisticated yet elegant approach to counteracting the selection bias that exists in the absence of randomization. By removing this selection bias, the researcher is able to increase the internal validity of their findings through better isolation of the effect of the treatment on the outcome, which is a critical component of drawing conclusions regarding the causal effects of the treatment (Campbell, 1957).

One of the major practical drawbacks of traditional propensity score analysis procedures is the limitation of comparing only two treatment conditions, typically a treatment and a control group. However, in practical research settings, comparisons are
often drawn between multiple treatment groups. One approach to comparing multiple
groups using binary propensity score approaches is through the implementation of
pairwise comparisons of each treatment condition. However, simulation study has
demonstrated a significant bias associated with this procedure when significant treatment
effect heterogeneity exists (Rassen et al., 2013).

Imbens (2000) is generally credited with the initial work in extending propensity
score analysis to applications for multiple treatment groups. Imbens’ definition of the
generalised propensity score is the conditional probability of receiving a particular level
of the treatment given the observed covariates. This extension of the definition of
propensity scores allows for multi-group comparisons. Propensity scores in this context
are generally calculated through multinomial logistic regression, particularly when the
levels of treatment are qualitatively different (Baser, 2008). However, implementation of
multiple group propensity score analyses is relatively limited, particularly with regard to
practical applications.

As with binary propensity score analysis, multiple strategies can be used in the
implementation of the propensity score. One approach to utilizing multiple propensity
scores is to include them as independent variables within a multiple regression equation
along with dummy variables related to treatment received (Spreeuwenberg et al., 2010).
Additionally, both propensity score regression adjustment and propensity score weighting
procedures have been demonstrated to be effective in multiple group comparisons (Feng,
Recently, statistical software packages have been created to conduct propensity score matching across multiple groups (Bryer, 2013; McCaffrey et al., 2013; Rassen et al., 2013). In particular, the TriMatch package was created for the R statistical software specifically to compare three non-equivalent treatment groups (Bryer, 2013). TriMatch begins by using logistic regression to calculate propensity scores. Several matching strategies then exist based on minimizing the total standardized distance between propensity scores within a given matched triplet. Matching options include matching within a specified caliper and with specified ratios (e.g. 3:2:n or 2:1:n). The TriMatch package also provides multiple means for checking the covariate balance achieved through the matching process, including multiple graphical outputs. Finally, this package performs analysis on the matched groups, using repeated measures ANOVA and a Friedman Rank Sum Test. If either of these tests is found to be significant, a Wilcoxon Signed Rank Test and three dependent sample t-tests are performed to assess differences between individual groups.

Utilizing one of these multi-group matching procedures may result in more precise estimates of treatment effect. Simulation has shown that such matching procedures that utilize an algorithm to create matches across all comparison groups have been shown to be preferable to pairwise approaches of comparing multiple groups in terms of reducing bias and increasing covariate balance, particularly when treatment effect heterogeneity exists (Rassen et al., 2013). Therefore, the TriMatch procedures have been chosen for this application.
Problem Gambling

Addiction, in a broad sense, affects a relatively large proportion of the general population. Recent estimates indicate that over 10% of the adult American population will develop a drug use disorder at some point during their lifetime (Compton, Thomas, Stinson, & Grant, 2007). However, when expanding the definition of addiction to include behavioral addictions, such as gambling, sex, exercise, and shopping, it has been estimated that approximately 47% of the U.S. adult population may exhibit addiction-like behavior (Sussman, Lisha, & Griffiths, 2011). This underscores the need for adequate identification and treatment of addictive disorders, particularly given the profoundly negative consequences these behaviors can have on the individuals experiencing them, those around them, as well as society as a whole.

Gambling addiction is also a relatively common phenomenon. In fact, given the often hidden nature of the disorder, rates of gambling addiction are likely much higher than the general public perceives. Current estimates suggest that from 2% to 6% of the population will develop significant gambling related problems during their lifetime; moreover, an estimated additional 2% will develop diagnosable levels of gambling problems during their lifetime (Kessler et al., 2008; Park et al., 2010; Penfold et al., 2006a, b; Shaffer & Hall, 2001). Therefore, as many as approximately one in twelve people will experience significant problems as a result of gambling at some point during their lifetime. Moreover, these estimates are likely to increase with the introduction of the changes to the gambling diagnosis in the DSM-V, most notably the decreasing of the threshold for diagnosis (Weinstock et al., 2013).
Gambling has become ubiquitous in the United States over the past several decades. Casino gambling, in particular, has seen a dramatic increase. Annual revenue from legal casinos in the United States has increased substantially over the past several decades as more and more states have legalized casino gambling (Eadington, 1999). In fact, between 1991 and 2006, casino revenue in the United States increased from 9 billion dollars annually to 32 billion (Ozurumba, 2009). Although the total gross gambling revenues in the United States decreased approximately 8% between 2007 and 2009, coinciding with an overall downturn in the American economy, the revenue was still greater in 2009 than in 2004 indicating a continuing upward trend in gambling (Eadington, 2011). This dramatic increase in gambling revenues is being observed outside of the United States as well, particularly in Europe, Asia, South Africa, and Australia (Wynne & Shaffer, 2003).

Casino gambling, of course, is not the only form of gambling that is generating a tremendous amount of revenue in the United States. State sponsored lotteries continue to be the largest source of gambling revenue. According to the North American Association of State and Provincial Lotteries (2014), in 2011 total lottery revenues in the United States exceeded 63 billion dollars. Further, charitable gambling activities, those conducted by non-profit organizations, generated approximately 1.3 billion dollars in gambling revenues in 1997; however, it has been shown that as the availability of for-profit gambling opportunities increases, the ability of non-profit organizations to generate revenue through charitable gambling decreases (Dolan & Landers, 2006).
Another major element related to the increase in overall gambling is the proliferation of online gambling opportunities. Online gambling has increased dramatically over the past few years, and internet gamblers may be particularly susceptible to developing gambling problems (Wood & Williams, 2007). Furthermore, internet gambling may be particularly attractive to younger people, which may suggest that it will continue to grow as a preferred means of gambling (Wood & Williams, 2011).

Regardless of the particular form it takes, gambling in the United States is an activity that people are spending a tremendous amount of money on each year. Further, gambling behavior and the monies associated with it continue to increase. Therefore, gambling is an activity that must be taken seriously, particularly as it relates to the potential for a continued increase in associated problems.

Until very recently, pathological gambling was listed in the Diagnostic and Statistical Manual of Mental Disorders within the category of impulse-control disorders, along with other behavioral disorders such as kleptomania, pyromania, and intermittent explosive disorder (American Psychiatric Association, 2000). However, clinical and research experience have shown that disordered gambling behavior shares many common characteristics with drug and alcohol misuse, including psychological and treatment components (Ashley & Boehlke, 2012; El-Guebaly et al., 2012; Potenza, 2006). Further, similar neurological pathways and vulnerabilities have been implicated for both substance abuse and gambling disorders (de Ruiter, Oosterlaan, Veltman, van den Brink, & Goudriaan, 2012; Frascella, Potenza, Brown, & Childress, 2010; Koehler et al., 2013). The alignment of gambling addiction with alcohol and substance use disorders will
hopefully have the impact of further legitimizing gambling disorder as a valid condition that will receive the recognition and resources it deserves.

Proposed changes have been enacted whereby the disorder was renamed from “pathological gambling” to “disordered gambling.” It has also become re-categorized with substance-related disorders (Holden, 2010; Petry, 2010). Additionally, the criterion related to the committing of illegal acts has been eliminated from the list of diagnostic criteria, and the threshold for diagnosis has been reduced from five of ten to four of nine criteria (Mitzner et al., 2011). Other proposed changes, however, were not accepted into the revisions of the DSM-V. For instance, Cunningham-Williams et al. (2009) advocated for the expansion of criteria related to the withdrawal-like symptoms associated with gambling.

Initial investigation of the impact of the removal of the illegal acts criterion and reduction of the threshold for diagnosis indicates an improvement in diagnostic classification without a reduction in the psychometric properties of existing measures, such as the National Opinion Research Center DSM-IV Screen for Gambling Problems (NODS; Petry et al., 2012). An initial investigation using gambling helpline callers estimates that the changes will result in an increase in individuals meeting the diagnostic criteria of as much as 11% depending on the setting (Petry et al., 2012; Weinstock et al., 2013).

The legalization and growing availability of gambling is often viewed as having various and far reaching social consequences (Barmaki, 2010). The most obvious consequence is that as gambling opportunities increase so too will gambling behaviors.
The opening of casinos has been associated with an increased in gambling frequency, amounts of money lost gambling, as well as problem gambling behavior in the surrounding areas (Jacques, Ladouceur, & Ferland, 2000; Room, Turner, & Ialomiteanu, 1999). The fallout from the increased availability of gambling may occur relatively quickly, and it will likely continue to increase over time. It has been estimated that a majority of treatment seeking problem gamblers had developed significant gambling-related problems within 2 years of beginning gambling (Lahti, Halme, Pankakoski, Sinclair, & Alho, 2012). Further, the prevalence of pathological gamblers has been found to continue to increase in areas where legal gambling has been available for extended periods of time. Rates of pathological gambling may be as much as three times higher in areas where legal gambling has been available for more than 20 years compared to areas where legal gambling has been available for less than 10 years (Volberg, 1994).

This then raises the issue of whether states or countries that allow legalized gambling have a social obligation to address issues of gambling addiction. Unfortunately, governmental support for research and treatment are far from commensurate with the growth in legal gambling (Pavalko, 2004). At the mental health agency level, there may be a general lack of preparation by the agency for the anticipated increase in gambling addiction that accompanies gambling expansion (Engel, Rosen, Weaver, & Soska, 2010).

Economic factors, such as high unemployment rates, are commonly cited as reasons for the legalization of gambling (Argusa, Lema, Asage, Maples & George, 2010; Richard, 2010). It is assumed that legalized gambling will generate much needed revenues that can be allocated towards social betterment. However, the expansion of
legalized gambling has been linked to negative financial impacts, such as a decreased ability for families to financially save, particularly for low-income, low-education, and urban families (MacDonald, McMullan, & Perrier, 2004). Further, legal casino gambling has been found to be associated with increased local bankruptcy rates (Barron, Staten, & Wilshusen, 2008; Nichols, Stitt, & Giacopassi, 2000). It has been estimated that the availability of casino gambling increases bankruptcy rates by between 2% and 9% (Boardman & Perry, 2007; Daraban & Thies, 2011). Research has also demonstrated a link between states that have larger proportions of residents who visit out-of-state casinos and increased bankruptcy rates (Garrett & Nichols, 2008). Gambling, in particular state lotteries have been associated with greater social economic inequity (Freund & Morris, 2006). However, in certain jurisdictions, the introduction of legal casino gambling has been associated with positive social and economic changes, including: increased funding for public education, increased availability of health care services, and a decreased rate of households relying on government financial assistance (Long, Johnson, & Oakley, 2011). Therefore, though gambling may in fact generate revenues that are put towards good causes, it is often the case that these revenues come at the cost of the overall financial well-being of the constituent populations.

Research has also focused on how legalized gambling is perceived by local residents. However, these studies on the perceptions of those living near casinos have yielded mixed results. Despite the empirically demonstrated negative consequences of available legalized gambling, it has been shown that people living near casinos generally view them positively (Chhabra, 2007). However, other studies have demonstrated that
following the introduction of a casino the perceptions of residents became less favorable compared to prior to the casino’s introduction, as well as less favorable as the availability of gambling remains in the community (Jacques, Ladouceur, & Ferland, 2000).

Individual perceptions of whether legalized gambling has had a positive or negative impact seem to be linked to the individual’s perception of whether crime has increased and whether the community has economically benefitted as a result of the introduction of legalized gambling (Hsu, 2000). To this issue of increased crime, research has also demonstrated that the opening of legal casinos may be associated with an increase in the number of individuals with criminal histories coming to the area (Piscitelli & Albanese, 2000). Similarly, quality of life for those residents of area with casino gambling was found to be related to perceived social impacts of the casinos. The less educated and those living in urban areas reported perceiving more negative social impacts of the casinos (Roehl, 1999). Therefore, the individual perception of the legalized gambling seems to be related to the individual experience, which may relate to issues of social equity.

As the extreme proliferation of gambling is a relatively recent phenomenon, the full extent of the impact may not yet be fully understood. However, it seems relatively clear that it leads to increased rates of gambling problems and the associated social and economic difficulties. It also appears that these negative impacts of gambling are serving to increase existing social inequities.
For the individual, problem gambling is thought of as progressing through a series of phases: winning, losing, and desperation (Lesieur & Custer, 1984). In the winning phase the individual is beginning gambling behavior and often experiences varying degrees of winning. The losing phase, as the name suggests, is characterized by the loss of money, often lost in an attempt to recoup previous losses. In the desperation phase the gambling has caused significant life problems, and the gambler is often driven to extreme measures, such as illegal acts. Recently, these phases have been expanded to include a fourth phase, hopelessness (Denure, Ford, Hillyard, Moore, & Scherer, 2006). At this point, the problem gambler has become so overwhelmed by the consequences of their gambling behavior that suicide often becomes a likely outcome.

Gambling addiction can have a variety of profoundly negative consequences. Most obviously, the individual exhibiting the gambling behavior can struggle in a multitude of ways, including financially and psychologically. Additionally, interpersonal functioning is often significantly impacted. Also, individuals with whom the gambler has a personal relationship are in many ways affected by this behavior.

Interpersonal difficulties are often associated with gambling addiction, as both a precipitant and a consequent of the behavior (Callan et al., 2011; Downs & Wollrych, 2010; Reid et al., 2011). These interpersonal difficulties may relate to a number of factors. Gambling has been referred to as the hidden addiction, as the lack of recognizable physical signs as is seen with alcohol or drug use can facilitate concealment of the behavior (Ladouceur, 2004). Furthermore, internet gambling, with its ease of access and relative anonymity, may increase the extent to which gambling behavior can
be concealed from significant others (Valentine & Hughes, 2012). Therefore, the secrecy associated with gambling, both in terms of the behavior itself as well as the financial difficulties the behavior has created, can be a contributing factor to interpersonal difficulties. In fact, lying and concealing gambling behavior is a major factor associated with diagnosing gambling disorder, as it is one of the criteria used in its diagnosis (American Psychiatric Association, 2013). It is estimated that approximately 78% of pathological gamblers lie to those around them regarding the extent of their gambling (Toce-Gerstein et al., 2003). Pathological gambling has also been associated with increased guilt related to the consequent deteriorating quality of interpersonal relationships (Locke, Shillkret, Everett, & Petry, 2013). These interpersonal difficulties may further contribute to the significant psychological difficulties faced by the pathological gambler, which will be discussed later.

Another major issue related to gambling addiction is the large sums of financial debt that can be amassed (e.g., Ciarrochi, 2002; Downs & Woolrych, 2010; Yip et al., 2007). One study of pathological gamblers receiving inpatient treatment found that, excluding home mortgages, 50% were more than $25,000 in debt, 10% were between $50,000 and $100,000 in debt, and 18% were more than $100,000 in debt; furthermore, pathological gamblers are nearly five times more likely to have declared bankruptcy than individuals with no gambling problems (Ciarrocchi, 2002). In another study of pathological gamblers within treatment, Teo and colleagues (2007) found a mean reported debt of $102,735.45, with the largest debt being $1.5 million. Problem gamblers often reported borrowing money from friends and family, thereby receiving what is
termed a “bailout” (Tang et al., 2007). In addition to borrowing money directly from friends and family, gamblers often take out loans from financial institutions, often in response to the debt that they have incurred from their gambling (Brown, Dickerson, McHardy, & Taylor, 2012; Tang et al., 2007).

Debt among pathological gamblers has been linked to significant mental health issues. Individuals in debt have been found to exhibit more psychological disorders if they are experiencing gambling problems than if they are non-gamblers (Meltzer, Bebbington, Brugha, Farrell, & Jenkins, 2013). Further, these financial losses resulting from gambling have been linked to a strong sense of shame that is often experienced by pathological gamblers (Yi & Kanetkar, 2011). Research has also found a significant link between gamblers’ financial problems and suicidality, which is another major issue with gambling addiction that will be discussed later (Meltzer et al., 2011; Wong et al., 2010a, b). In fact, from posthumous studies of gamblers who have died by suicide, it is estimated that between 47.2% and 100% had amassed gambling related debt at the time of their death (Wong et al., 2010a, b). Further, among callers to a problem gambling hotline, those that reported a risk of suicide (25.6%) were more likely to acknowledge financial, as well as family, legal, mental, and substance abuse problems (Ledgerwood et al., 2005). Therefore, financial issues can play a major role in the development of, and fallout from gambling addiction.

Pathological gambling has also been associated with increased rates of criminal behavior, most commonly crimes designed to acquire revenue (Meyer & Stadler, 1999). Problem and pathological gamblers often resort to stealing in order to continue their
gambling behavior, and often this stealing is from sources familiar to the individual, such as friends, family, and the workplace (Cheah et al., 2008; Meyer & Stadler, 1999). In a study of pathological gamblers, Meyer and Stadler (1999) found the most common forms of illegal behavior admitted to were: fraud (37.7%), theft from the workplace (23.3%), embezzlement (21.7%), and theft from family (21%). Another study estimates that between 20.7% and 23.5% of pathological gamblers admit to a history of illegal acts (Lee et al., 2011). Estimates of histories of illegal acts among treatment-seeking gamblers have been found to be lower (11.6%), which may indicate that those with criminal histories may be less likely to seek treatment (Martin, MacDonald, & Ishiguro, 2013). The prevalence and nature of this criminal behavior seems to speak to the desperation of the situation that gamblers find themselves in, particularly as it relates to financial difficulty. However, it has been found that of the diagnostic criteria for pathological gambling in the DSM-IV-TR, the one related to the committing of illegal acts was the least commonly endorsed (Molde et al., 2010). Further, as previously discussed, the current set of criteria for gambling disorder in the DSM-V no longer contains a criterion related to illegal acts (American Psychiatric Association, 2013).

Those engaging in problematic levels of gambling behavior are often divided into two sub-groups based upon how they orient themselves towards the behavior. For many, gambling is a coping strategy whereby the individual engages in the behavior to avoid dealing with a negative emotional state (Wood & Griffiths, 2007). Such individuals, who are often labeled as “escape” gamblers, generally engage in the behavior as a way to numb themselves against some sort of underlying emotional dysregulation. These
individuals generally tend to favor games of chance, such as slot machines. Conversely, those labeled as “action” gamblers crave the high that they can receive from gambling, and they generally tend to favor games of perceived skill, such as poker. This distinction is often related to the gender of the gambler, with women and the elderly being more commonly associated with the escape-type gambling and men with the action-type (Li, 2007; Odlaug, Marsh, Kim, & Grant, 2011; Potenza et al., 2001). However, this distinction, also referred to as strategic and non-strategic gambling, is not always entirely clear-cut, as one study found that over 40% of pathological gamblers regularly engage in both types of gambling (Odlaug et al., 2011). However, identifying and understanding this distinction may have implications for the proper treatment of gambling addiction (Tang et al., 2007).

Certain demographic groups have been associated with an increased risk for developing gambling-related problems. Also, various demographic factors have been associated with differential gambling presentations. Moreover, factors have also been identified that can help to understand the manifestations and course of gambling addiction, as well as having implications for treatment.

Gender is a factor that has received much attention for its role in gambling. Typical gender differences are found in patterns of gambling activities, types of gambling related problems, and treatment seeking (e.g., Nelson et al., 2006; Nower & Blaszczynski, 2006; Tang et al., 2007). Compared to male gamblers, female gamblers begin gambling later in life, progress more quickly from first gambling experience to problematic levels of gambling, have less gambling-related debt, and tend to prefer
gambling on slot machines, as opposed to sports betting or card games (Crisp et al., 2004; Nower & Blaszczynski, 2006; Tang, Wu, & Tang, 2007). Female gamblers have also been found to bet in smaller denominations compared to men, and female pathological gamblers have been found to earn significantly less income than their male counterparts (Nower & Blaszczynski, 2006).

Gender also has implications for gambling treatment. Women have been found to have begun gambling later in life, first attempted to quit gambling later in life, and have gambled for a shorter duration by the time they enter treatment (Ladd & Petry, 2002). Women have also been associated with greater gambling and financial problems at the time of accessing treatment, as well as a greater readiness to change (Ledgerwood, Wiedeman, Moore, & Arfken, 2012). Women pathological gamblers have been found to be more likely to have accessed mental health treatment, and have been found to access treatment later in life (Potenza et al., 2001; Tang et al., 2007). This may indicate a need to more actively seek male pathological gamblers who are in need of treatment.

Women have also been found to present with differing gambling symptoms, particularly men were more likely to meet the DSM criteria related to preoccupation, illegal acts (DSM-IV-TR), and lost/hurt relationships/jobs, whereas women were more likely to meet the criterion related to escaping personal problems (Crisp et al., 2000). However, another study of adolescents found differences in symptom patterns in a community sample but failed to find any differences in symptom endorsement among male and female pathological gamblers in treatment (Faregh & Derevensky, 2011). In either case, clinicians may need to be aware that men and women may present for
treatment with different symptom patterns. Other clinical issues may also vary between men and women. Women have been found to be more likely to indicate a desire to prevent suicide as their reason for self-excluding from casinos, which is a process by which an individual can bar themselves from entering a casino (Nower & Blaszczynski, 2006). Also, female pathological gamblers have been found to attempt suicide more frequently than males; however, pathological gamblers who die by suicide have been found to have an increased likelihood of being male (Martins et al., 2004; Potenza et al., 2001; Wong et al., 2010a; Wong et al., 2010b). Further, women have been found to be more likely to report resolution of their problems through treatment than men (Crisp et al., 2000).

Ethnicity may also play a role in gambling addiction. A study of callers to a problem gambling helpline found that Caucasian callers were more likely than African-American callers to have previously sought mental health treatment; however, the study found no significant difference in proportions reporting financial problems related to gambling (Barry et al., 2008). One study looking at minority acculturation to the dominant culture indicates that successfully assimilating into the dominant culture can be a protective factor against problem gambling (Oei & Raylu, 2009). Also, ethnic minorities are greatly underrepresented in gambling treatment (Volberg, 1994).

Age also seem to play a significant role in gambling behaviors. Age of gambling behavior onset has been found to have an impact on gambling manifestation. Individuals who began gambling in pre/early adolescence reported greater severity of psychiatric, family/social, substance abuse problems, and suicidal ideation but no difference in
suicide attempts (Burge et al., 2006). Individuals who experience late onset of pathological gambling, after age 55, have been found to be less likely to have declared bankruptcy, less likely to have a parent with a gambling problem, and more likely to exhibit an anxiety disorder (Grant, Kim, Odlaug, Buchanan, & Potenza, 2009). Older onset of gambling problems has also been associated with more severe psychopathology, while younger onset has been associated with more severe gambling problems (Jimenez-Murcia et al., 2010).

Current age is also a factor in gambling manifestation. Older adult pathological gamblers have been found to have accumulated significantly higher debt (Kennedy et al., 2005). However, recreational gambling among older adults has been found to have no association with overall health and well-being (Desai, Maciejewski, Dausey, Caldarone, & Potenza, 2004). Older adults have also been found to more often cite suicide prevention as a reason for self-exclusion from casinos (Nower and Blaszczynski, 2008). Rates of gambling among adolescents are estimated to be far greater than in the general population (Gupta & Derevensky, 1998; LaBrie & Shaffer, 2007). Further, other researchers have found that younger individuals who gamble may actually be at a significantly higher risk for gambling-related problems, including suicidal ideation and attempts, and low rates of seeking help for their gambling problems (Afifi et al., 2007; Dowling, Clark, Memery, & Corney, 2005; Froberg et al., 2012; Martins et al., 2004; Nower et al., 2004). One particular age-related subset of individuals, student athletes, have been found to be associated with increased likelihood to engage in gambling behaviors and have gambling-related debt (Bovard, 2008; Stuhldreher et al., 2007). The
interplay between age and gambling is, therefore, obviously a complex one whose relationship may not necessarily be linear given the evidence that both the young and the elderly may be at increased risk.

Research has also implicated exposure to gambling as a risk factor in the development of gambling problems, particularly as it relates to early childhood exposure to gambling (Reith & Dobbie, 2011). Complicating this issue somewhat is the research that has implicated inherited genetic factors in the development of gambling addiction (Slutske, Zhu, Meier, & Martin, 2010). Therefore, it is likely a combination of exposure and genetic predispositions that contribute to the development of gambling problems.

Various psychological variables have been associated with gambling addiction, including impulsivity, interpersonal sensitivity, interpersonal distrust, poor coping skills, lack of self-discipline, and a tendency to make decisions hastily (Reid et al., 2011). Impulsivity, in particular, may play a moderating and a mediating role in the relationship between life stress and gambling problems (Tang & Wu, 2012). Other psychological variables associated with gambling problems are neuroticism, low esteem from others, and low family support (Taormina, 2009). Anger problems are also commonly seen amongst pathological gamblers (Korman et al., 2008).

Gambling behavior has also been strongly associated with incarcerated individuals, which may be an issue during incarceration as well as upon re-entry into society (Williams & Walker, 2009). A study of incarcerated individuals found that nearly 16% had moderate to severe gambling problems (Turner, Preston, Saunders, McAvoy,
Another study of adolescents in youth detention centers found that 18% met the criteria for pathological gambling (Magoon, Gupta, & Derevensky, 2007).

Religion may also play a role in the development, or prevention of gambling problems. Individuals that frequently attend religious services have been found to less often develop gambling problems (Hoffman, 2000). Therefore, active religious affiliation may be a protective factor for developing gambling problems.

Co-occurring disorders, both psychological and substance-related, have been strongly associated with gambling addiction. The relationship between the two might be very complicated, which is to say that co-occurring disorders might influence the development and course, as well as be a product of gambling addiction. It has been estimated that as many as 96.3% of pathological gamblers will develop at least one World Health Organization Composite International Diagnostic Interview (CIDI)/DSM-IV-TR disorder(s) during their lifetime (e.g., Chan et al., 2009; Gill et al., 2006; Kessler et al., 2008; Lorains et al., 2011). Further, individuals who engage in non-pathological, or social gambling have been found to have higher rates of psychiatric disorders compared to non-gamblers (Westermeyer, 2008).

There is debate over whether mental health issues are generally a cause or a consequence of problem gambling; however, the literature seems to suggest that the mental health issues are likely more of a precipitant to the gambling (O’Brien, 2011). From this perspective, the gambling is used as a coping mechanism for the preexisting psychological difficulties. It has been estimated that for individuals who experience pathological levels of gambling along with another lifetime mental health disorder,
74.3% had at least one disorder that preceded the onset of the pathological gambling (Kessler et al., 2008). This supports the idea that the gambling disorder is developed as an attempt to cope with a mental health condition.

There is a broad range of disorders that are found to co-occur with gambling addiction. A meta-analysis of research on gambling and co-occurring disorders found that the following were the most commonly observed co-occurring mental health conditions: any type of mood disorder (37.9%) and any type of anxiety disorder (37.4%) (Lorains, Cowlishaw, & Thomas, 2011). In the National Comorbidity Study replication, Kessler et al. (2008) found that among pathological gamblers, 38.6% also had major depressive disorder or dysthymia, 55.6% also had a type of mood disorder, 60.3% also had a type of anxiety disorder, and 14.8% also had Posttraumatic Stress Disorder (PTSD). These are compared to rates of 19.1%, 20.8%, 28.8%, and 6.8%, respectively, for these four psychological disorders found within the general population (Kessler et al., 2005). This indicates that for the disorders identified by Kessler (2005; 2008), rates may be approximately twice as high for pathological gamblers compared to the general population. Among anxiety disorders, obsessive-compulsive disorder, in particular, has been found to commonly co-occur with gambling addiction (Gonzalez-Ibanez et al., 2003). Personality disorders have also been linked to pathological gambling. It has been estimated that over 60% of pathological gamblers have at least one personality disorder, with obsessive-compulsive, paranoid, and antisocial personality disorders being the most common (Petry, Stinson, & Grant, 2005). Attention Deficit/ Hyperactivity Disorder (ADHD) is another disorder that is gaining increasing amounts of attention for its
relationship to gambling addiction, in part, because of the impulsive characteristics commonly associated with both ADHD and pathological gambling (Duven, Unterrainer, & Wölfling, 2012). Multiple studies have linked disordered gambling and ADHD, with lifetime rates of ADHD have been estimated between 13% and 20% among those with lifetime pathological gambling (Crockford & el-Guebaly, 1998; Kessler et al., 2008; Rugle & Melamed, 1993; Specker et al., 2005). Furthermore, individuals experiencing pathological levels of gambling are at a significantly increased risk for experiencing subclinical levels of psychiatric symptoms other than those included in the diagnostic criteria for pathological gambling (Boudreau, Labrie, & Shaffer, 2009).

These co-occurring mental health conditions have implications for clinical presentation and treatment. Research supports that the rate of psychiatric disorders increase with severity of gambling problems (Park et al., 2010). Similarly, among individuals seeking psychiatric treatment, those exhibiting pathological gambling presented with greater numbers of psychiatric disorders (Zimmerman, Chelminski, & Young, 2006). It has also been found that as the number of co-occurring disorders with which a client presents for treatment increases so too does the severity of gambling problems increase (Soberay et al., 2013). Consequently, clinical assessment that focuses on comorbid psychiatric conditions may lead to a better understanding of the gambling problems with which a client presents.

In terms of substance use and misuse with gambling addiction, drug, alcohol, and tobacco use have all been associated with the gambling disorder (Walther, Morgenstern, & Hanewinkel, 2012). Lorains, Cowlishaw, and Thomas (2011) found that approximately
58% of pathological gamblers also have a co-occurring substance use disorder. Similar to mental health issues, substance use disorders may play a significant role in the development as well as the perpetuation of gambling. In terms of treatment issues, individuals in treatment for gambling problems who were also currently using alcohol frequently presented with more severe gambling behavior, yet were equally responsive to treatment (Stinchfield, Kushner, & Winters, 2005). Similarly, within treatment, substance dependence is a major factor in differentiating those who present with pathological levels of gambling from those presenting with sub-clinical levels (Namrata & Oei, 2009). Even among substance abusers, those with gambling problems had greater rates of psychiatric distress (Petry, 2000). Therefore, the presence of multiple addictive disorders may have strong implications for problem gambling treatment.

Studies have repeatedly shown that a relatively large proportion of individuals exhibiting gambling addiction have considered, attempted, or completed suicide, which makes this a critical clinical issue for this population. Lifetime rates of suicidal ideation and attempts among pathological gamblers have been estimated between 24.7% to 81.4% and 6.3% to 32.7%, respectively (e.g., Battersby et al., 2006; Bu & Skutle, 2012; Grall-Bronnec et al., 2012; Ledgerwood & Petry, 2004; Park et al., 2010). Therefore, as many as nearly one in three pathological gamblers will attempt suicide during their lifetime. When looking only at the past twelve months, a study of problem gamblers found that 10.2% reported suicidal ideation and 2.4% reported suicide attempts during this time period (Afifi et al., 2010). Research also supports that increased gambling severity is
associated with increased risk for suicide (Brooker et al., 2009; Hodgins et al., 2006; Langhinrichsen-Rohling et al., 2004; Ledgerwood et al., 2005).

Psychotherapy is the most widely studied and most commonly applied approach to gambling treatment; moreover, psychotherapy has demonstrated effectiveness in the treatment of gambling addiction on a number of gambling, psychological, and interpersonal measures (e.g., Pallensen et al., 2005; Sharma & Sharma, 2012; Sylvain et al., 1997). Also, studies have shown that individuals who have received treatment for their gambling behavior have a decreased risk for suicide (Kennedy et al., 2005). A meta-analysis of studies examining the outcome effect of psychological treatment of pathological gambling found that of the 22 targeted studies, the meta-analysis found psychological treatments were far more effective than no treatment at termination as well as at follow-up periods beyond the termination of treatment (Pallesen et al., 2005). This meta-analysis also found that cognitive-behavioral therapy was the most commonly applied form of psychotherapy. Furthermore, Crisp et al. (2001) found that psychotherapeutic treatment of problem gamblers can demonstrate effectiveness after as few as five treatment sessions. In fact, one study found that a single-session psycho-education group was associated with reductions in gambling problems (Petry, Weinstock, Ledgerwood, & Morasco, 2008). However, it has been found that longer retention in treatment is associated with more positive outcomes and greater satisfaction with treatment (Toneatto & Dragonetti, 2008).

Delivery methods of therapy may be varied. Online-based support for problem gambling has various advantages, including cost-effectiveness and easy access, and can
range from psycho-education, peer-support networks, to professionally delivered intervention (Griffiths & Cooper, 2003). For instance, cognitive-behavioral treatments utilizing an internet-based format have demonstrated effectiveness in the treatment of gambling problems (Carlbring, Degerman, Jonsson, Andersson, 2012; Castren et al., 2013). Another study demonstrated that telephone counseling might be equally as effective as face-to-face counseling interventions (Tse et al., 2013). Self-help workbooks have also demonstrated effectiveness, but not as effective as face-to-face therapy (Petry et al., 2006). These findings may be particularly relevant for those who live in geographically isolated areas where treatment may not be available or for those who for any other reason may not be able to utilize in-person treatment.

Psychotherapy is not the only form of treatment for gambling addiction. Various medications have been identified as possible treatments for pathological gambling, particularly those that target the neurotransmitters serotonin and dopamine (Potenza, 2008). However, in a study of treatment preferences among pathological gamblers, medication was found to be the least desired of the available treatments, which included numerous forms of psychotherapy (Najavits, 2011).

Other, informal methods may also be helpful for individuals to overcome gambling problems. For instance, journaling has been found to be beneficial to problem gamblers (Dwyer, Piquette, Buckle, & McCaslin, 2013). Self-exclusion is a non-therapeutic approach whereby a gambler can voluntarily bar him or herself from entering casinos. Self-exclusion has been associated with reducing gambling problems,
particularly when coupled with formal treatment or self-help groups (Nelson, Kleschinsky, LaBrie, Kaplan, & Shaffer, 2010).

However, it is unclear whether varying forms of psychotherapy are more or less effective. Studies comparing a cognitive-behavioral treatment to a twelve-step facilitation treatment found that each was effective compared to a control, yet neither treatment was significantly more effective than the other (Marceaux & Melville, 2011; Toneatto & Dragonetti, 2008). Another study comparing cognitive-behavioral and motivational interviewing based group treatments found each to be superior to no treatment but equivalent to each other (Carlbring, Jonsson, Josephson, & Forsberg, 2010).

Not all people that overcome gambling problems do so as a result of any identifiable treatment steps. So-called natural, or spontaneous recovery may be a relatively common phenomenon among individuals who experience gambling problems at some point during their lifetime (Hodgins & El-Guebaly, 2000). In fact, it has been estimated that approximately one third of those that develop pathological gambling during their lifetime will overcome their gambling problems without seeking any form of treatment (Slutske, 2006).

In the treatment of gambling addiction, both cognitive and behavioral counseling strategies are commonly utilized forms of treatment; additionally, these two approaches are often used in conjunction, in what is referred to as Cognitive-Behavioral Therapy (CBT) (Ladouceur et al., 2002). Cognitive models of addiction treatment are grounded on the idea that gambling addiction is grounded in, and perpetuated by erroneous beliefs and maladaptive patterns of thinking, particularly regarding the ultimate outcomes of
gambling behavior, conceptions of personal luck, and erroneously perceived control over outcomes of random events (Johnson & Dixon, 2009; Ladoceur et al., 2002; Wohl, Young, & Hart, 2007). Moreover, irrational beliefs regarding the outcomes of gambling behavior have been linked to the development and perpetuation of gambling problems (Tang & Wu, 2010). Irrational thoughts about gambling behavior have also been positively associated with severity of gambling problems with which individuals present for treatment, as well as an increase in the number of co-occurring psychiatric disorders (Houndslow, Smith, Battersby, & Morefield, 2011; Xian et al., 2008). Metacognitions, or the ways we think about our thoughts, have also been implicated in gambling addiction, particularly negative beliefs regarding thoughts related to the uncontrollability or need to control gambling have been associated with problem gambling behavior (Lindberg, Fernie, & Spada, 2011). Therefore, cognitive interventions for gambling problems, through a process termed cognitive reappraisal, are designed to help challenge and, ultimately, alter these faulty and erroneous thinking patterns (Tolchard & Battersby, 2013). This process of cognitive correction can focus on various common gambling misconceptions, including: concepts of randomness, understanding of independence of events, and illusions of control over gambling outcomes (Ladouceur et al., 2003).

Behavioral models of addiction treatment focus on the degree to which addicted behaviors are learned and perpetuated through a process of reinforcement of the behavior (Ladouceur et al., 2002). Therefore, behavioral treatments place importance on the extent to which an addiction is a learned behavior and attempt to identify and modify sources of reinforcement and punishment in the development, maintenance, and treatment of
addictions. Behavioral intervention techniques, such as activity scheduling, contingency management, and in-vivo or imaginal desensitization, have also been found to be effective interventions for the treatment of pathological gambling (Dowling, Jackson, & Thomas, 2008; Rawson, McCann, Flammino, Shoptaw, Miotto, Reiber & Ling, 2006; Tavares, Zilberman, & el-Guebaly, 2003). Exposure therapy, a behavioral technique whereby a problem gambler is exposed to gambling cues without initiating gambling behavior, has also demonstrated effectiveness for gambling addiction (Riley, Smith, & Oakes, 2011).

CBT treatments often consist of the following aspects: (1) correcting negative beliefs regarding randomness and outcomes of gambling behavior, (2) development of problem-solving skills, (3) social skills training, and (4) relapse prevention, including ways of coping with “high risk” situations (Sylvain, Ladouceur, & Boisvert, 1997). The development of coping skills may be a critical component to the success of CBT treatment. The attainment of adequate coping skills has been shown to mediate the relationship between treatment and outcomes (Petry, Litt, Kadden, & Ledgerwood, 2007). CBT interventions have also demonstrated effectiveness in reducing gambling behavior among at-risk, sub-clinical gamblers (Larimer et al., 2012). Further, a meta-analysis of CBT interventions for gambling addiction found effect sizes to be significant as far out as 24 months from cessation of treatment (Gooding & Tarrier, 2009). It seems likely that therapeutic gains may last even longer; however, studies that follow clients for greater than two years following treatment are lacking.
CBT interventions are commonly delivered in both individual and group formats. Both group and individual delivery methods of CBT interventions have demonstrated effectiveness, although individual CBT may have greater therapeutic effectiveness, particularly at follow-up (Dowling, Smith, & Thomas, 2007). Further, combining individual and group CBT interventions has also demonstrated effectiveness in reducing gambling problems and improving overall functioning (Oakes, Laughlin, McLaughlin, & Battersby, 2012).

Dialectical Behavior Therapy (DBT), an offshoot of CBT, uses a focus on developing such skills as mindfulness, distress tolerance, and emotional regulation, and it has also been demonstrated to be helpful in the treatment of gambling problems (Christensen et al., 2013). In particular, mindfulness based interventions may be helpful in coping with the distressing thoughts and emotional states associated with gambling addiction (de Lisle, Dowling, & Sabura, 2011). Similarly, Acceptance and Commitment Therapy (ACT), another emerging form of CBT that focuses on accepting current experiences and committing to value-driven action has also shown initial evidence that it may be helpful in the reduction of gambling problems (Hayes, Luoma, Bond, Masuda, & Lillis, 2006; Nastally & Dixon, 2012; Weatherly, Montes, Peters, & Wilson, 2012).

Given that CBT interventions are the most studied form of psychotherapy for the treatment of gambling problems, its effectiveness has been researched among special populations, particularly those that may be most susceptible. CBT interventions have demonstrated effectiveness in treating pathological gamblers with chronic schizophrenia, those with Parkinson’s disease, as well as those with acquired brain injury (Echeburua,
CBT has also shown to be adaptable for congruence with various cultural beliefs and behaviors (Okuda, Balan, Petry, Oquendo, & Blanco, 2009).

Solution-Focused Therapy (SFT) is a strengths-based, goal-directed approach to psychotherapy that originated in part as a response to the problem-focused treatments that characterize most of psychotherapy (de Shazer & Isebaert, 2004). SFT focuses more on helping clients work towards an idealized goal state that represents what they would like their lives to be like, rather than focusing primarily on their current difficulties. Common SFT techniques include the “miracle question”, whereby clients are asked to describe their lives if all their problems were miraculously solved. SFT also commonly uses scaling questions to assess client progress towards their goal state (de Shazer & Isebaert, 2004). Since the client is in control of determining their own goal-state, SFT is well suited to those seeking abstinence as well as those seeking to moderate their problem behavior (Nelle, 2005).

SFT has recently been applied more to, and has shown effectiveness in, the treatment of addictive disorders. Specifically it has demonstrated effectiveness in the treatment of substance abuse (Smock et al., 2008). However, the literature does not address whether this effectiveness will generalize to the treatment of problem gambling (Pallesen et al., 2005).

From the psychodynamic perspective, addictive disorders are an issue of difficulty regulating behaviors, relationships, and emotions (Khantzian, 2012). The addictive behavior then is an attempt to cope with this dysregulation. From this
perspective, addiction is also often viewed as a relational problem where the addict develops an attachment to their addicted behavior in an attempt to cope with previous attachment problems (Khantzian & Weegmann, 2009). Psychodynamic treatments of addiction often focus on the psychological transference that occurs in the therapeutic relationship (Sweet, 2012). This relationship can help serve to develop adaptive relational patterns, which in turn help to decrease the dependence on the addiction (Potik, Adelson, & Schreiber, 2007).

Psychodynamic models of therapy have demonstrated effectiveness in the treatment of various substance addictions, including nicotine, opiate, alcohol, and cocaine dependences (Crits-Cristoph et al., 1999; Gregory et al., 2008; Sandahl, Herlitz, Ahlin, & Ronnberg, 1998; Woody, McLellan, Luborsky, & O’Brien, 1995; Zernig et al., 2008). As is also the case with SFT, large-scale, methodologically sound studies specifically looking at the effectiveness of psychodynamic treatments for pathological gambling are virtually non-existent (Pallesen et al., 2005; Raylu, Loo, & Oei, 2013). However, particularly given the effectiveness of psychodynamic therapy on the treatment of personality disorders, it may be a viable treatment option for pathological gambling given the large proportion of pathological gamblers with comorbid personality disorders (Leichsenring & Leibing, 2003; Petry, Stinson, & Grant, 2005).

In contrast to strictly psychodynamic treatments, studies investigating the interpersonal elements related to problem gambling treatment have yielded more promising results. Given the interpersonal and familial antecedents and consequents of gambling behavior, an increasing number of clinicians conceive of treatment as an
interpersonal process, which focuses on improving interpersonal functioning and, thereby, improving overall functioning. This, in turn, can contribute to controlling gambling problems (McComb, Lee, & Sprenkle, 2009). Such approaches to addiction treatment have demonstrated increased motivation to change as well as personal empowerment (Wood, Englander-Golden, Golden, & Pillai, 2010). Interpersonal models of therapy that focus specifically on working with gamblers and their significant others have also demonstrated effectiveness (Lee, 2007). Further, research indicates that gambling treatment outcomes are more positive and retention is greater for those whose significant others are involved in the treatment process (Ingle, Marotta, McMillan, & Wisdom, 2008).

The working relationship between the psychotherapy client and their counselor has been demonstrated to be a moderate predictor of outcomes across treatment settings (Martin, Garske, & Davis, 2000). Research has also demonstrated the importance that the therapeutic alliance plays in the treatment of a wide range of addictions. Through a retrospective study of individuals having received a problem gambling treatment, client ratings of the working alliance were found to be a significant predictor of problem resolution (Smith, Thomas, & Jackson, 2010). The findings related to what source of ratings of the therapeutic alliance, whether the client, counselor, or an outside observer, best predicts outcomes have been mixed. Research has demonstrated that client, observer, and therapist perceptions and ratings of the therapeutic alliance may vary (Fenton, Cecero, Nich, Frankforter, & Carroll, 2001). Some findings suggest that therapist ratings of the alliance may be a better predictor of outcomes (Meier, Donmall, McElduff,
Barrowclough, & Heller, 2006). However, in a study of problem gamblers, client-rated alliance predicted both gambling and general functioning outcomes, while therapist-rated alliance measures only predicted general functioning outcomes; moreover, client satisfaction with treatment was found to have a mediating relationship between alliance and outcomes (Dowling & Cosic, 2011). Several factors have been implicated in whether or not an addiction client forms a strong alliance with their therapist, including: social support, attachment style, readiness to change, external pressure to change, counselor qualifications, and whether the counselors were themselves “ex-addicts” (Meier, Donmall, Barrowclough, McElduff, & Heller, 2005).

Current understandings of behavioral change conceptualize it as a cyclical process that traverses a series of stages, which range from a denial of any problem and an absence of any motivation for change to maintaining behavioral changes that have already been achieved (DiClemente, Schlundt, & Gemmell, 2004; Prochaska & Norcross, 2001). Each stage of change presents unique characteristics and clinical challenges, and an understanding of a client’s current stage of change may have implications for appropriate interventions. A problem gambler’s readiness for change has been found to increase with higher levels of emotional awareness, previous experience with GA, and higher levels of depression (Gomes & Pascual-Leone, 2008).

Models differ on exactly how many stages of change exist that a client may go through. Precontemplation, Contemplation, Preparation, Action, Maintenance, and Relapse/Termination are generally considered an exhaustive list of the stages; however, not all models include all six stages (Norcross, Krebs, & Prochaska, 2011; Prochaska &
Norcross, 2001; Prochaska, DiClemente, & Norcross, 1992). Precontemplation is characterized by a lack of awareness or complete denial that a problem exists; therefore, there is no motivation to change at this stage. At the contemplation stage, the individual has become aware that a problem exists but are still very much dealing with their ambivalence to change the behavior. At the preparation stage, the individual has resolved to change their problem behavior, and they may have begun to make minor modifications to their behaviors. In the action stage, the individual is actively changing their problem behavior. An individual is generally considered to enter the maintenance stage when they have sustained successful change of their problem behavior for at least six months. Relapse is generally considered a normal part of the addiction change process, and it signifies the point at which an individual begins to cycle back through the previous stages. If an individual does not experience relapse, the termination stage is viewed as being achieved when the individual has sustained change long enough that they no longer have to actively work at modifying their behavior.

An individual’s stage of change at the time of entering treatment has been implicated in the severity of problems with which they present. Awareness that a problem exists has been found to be associated with more severe clinical presentations, while those that have begun to modify their behaviors have been associated with less severe presentations (Gossop, Stewart, & Marsden, 2007). This finding indicates that the individuals associated with the contemplation stage may enter treatment reporting the most severe problems. Petry (2005), in a sample of pathological gamblers, found that individuals associated with the precontemplation stage reported less severe gambling
problems, which may be indicative of the denial associated with this stage. Petry (2005) also found that those associated with the action and maintenance stages reported more severe gambling problems, which, again, may be indicative of the increased awareness of the problems of their behaviors at these stages.

Viewing a client in terms of stages of change may also aid in understanding how long they may remain in treatment and/or whether they are at risk for dropping out of treatment prematurely. Clients exhibiting traits consistent with the precontemplation stage have been found to prematurely terminate from therapy (Callaghan et al., 2005). Similarly, those not exhibiting traits of the contemplation stage have also been found to prematurely terminate from therapy (Derisley & Reynolds, 2000). These findings suggest the role that acknowledgement of a problem plays in remaining in treatment. Conversely, clients and those with higher levels of readiness to change, such as those exhibiting characteristics of the maintenance stage, have been associated with greater treatment retention (George et al., 1998; Henderson, Saules, & Galen, 2004). In a sample of pathological gamblers, stages of change as measured by the University of Rhode Island Change Assessment (URICA) were a moderate predictor of treatment dropout; however this study found that no single stage of change was significantly related to retention (Gomez-Pena et al., 2012). Therefore, stages of change have been implicated in client retention in, and dropout from treatment. However, like many other areas of treatment, the research specific to gambling addiction remains mixed and incomplete.

Stages of change have also been associated with the extent to which an individual improves through the course treatment. Within addiction treatment, the precontemplation
stage has been associated with less symptom improvement through treatment, the contemplation stage has been associated with moderate levels of symptom improvement, and the action and maintenance stages have been associated with relatively higher levels of symptom improvement (Henderson, Saules, & Galen, 2004; Norcross, Krebs, and Prochaska, 2011; Rochlen, Rude, & Baron, 2005). These findings mirror those related to treatment retention in that individuals more characterized by denial improve less while those characterized by the actively working on change benefit more from treatment. Likewise, a study of the relationship between stages of change and improvement through treatment for a sample of pathological gamblers found similar results: precontemplation was associated with lower levels of therapeutic change, while action and maintenance were associated with higher observed levels of change (Petry, 2005). However, another study using a sample of pathological gamblers failed to find an association between stages of change and clinical improvement post-treatment or at treatment follow-up (Gomez-Pena, et al., 2012).

Assessing a client’s readiness for change may play a central role in the treatment of gambling addiction. However, relatively few studies have been conducted using problem gambling samples. Consequently, these relationships need to be further investigated using problem and pathological gambling samples.

As previously discussed, co-occurring mental health and substance use disorders are common, if not typical among those experiencing gambling addiction (e.g. Kessler et al., 2008). Research has shown that the assessment and treatment of any co-occurring conditions may be beneficial in a number of ways. Treatment outcomes have been found
to be more favorable when the treatment included a focus on co-occurring disorders compared to treatment with a sole focus on the addictive behavior (Kim et al., 2006; Penney et al., 2012). Also, individual who have undergone treatment that addressed co-occurring conditions have been shown to be more resilient to relapse (Brown et al., 1997). Moreover, addressing co-occurring disorders has been associated with increased client satisfaction with treatment (Schulte et al., 2011). The body of research indicates that it is best practices for gambling treatment to expand its scope beyond solely gambling behavior to include other disorders that the client is likely experiencing that may be playing a role in perpetuating the gambling.

Certain personological characteristics may also be related to treatment outcomes. Sensation seeking-traits have been associated with dropping out of treatment and, consequently, less favorable treatment outcomes (Smith et al., 2010). High levels of impulsiveness have also been associated with treatment dropout (Alvarez-Moya et al., 2011). Similarly, individuals exhibiting neurotic personality traits and those exhibiting low levels of conscientiousness have been found to be at risk for treatment drop out and relapse (Ramos-Grille, Aragay, Gomá-i-Freixanet, Valero, & Vallès, 2013).

Another element related to treatment outcomes is the severity of the symptoms with which the client presents for treatment. In psychotherapy in general, greater symptom severity at the beginning of treatment has been associated with less improvement through treatment (Lambert & Anderson, 1996). Accordingly, in studies of pathological gamblers, those presenting for treatment with more severe psychological distress were found to have less positive outcomes following treatment (Dowling, 2009;
Jimenez-Murcia et al., 2007). Previous experience with treatment may also be related to outcomes. Individuals re-entering treatment for gambling problems have been found to have more positive outcomes compared to those seeking treatment for the first time (Jackson et al., 2008).

There seem to be two general approaches to the measurement of gambling treatment outcomes. The first approach is to explicitly measure gambling behavior. However, this approach may be incomplete, particularly in the short term. Considering poor overall functioning as an indicator of risk for relapse, it may be more telling to consider taking the second approach, which is to measure the person’s psychosocial functioning (Sander & Peters, 2009). Moreover, a study of treatment outcomes found a very strong correlation between gambling-related outcomes and general functioning outcomes (Dowling & Cosic, 2011).

Summary

In summary, propensity score analyses are a set of procedures that were designed to address the selection bias that often exists within social science research. Traditional propensity score techniques were developed to balance two groups across a set of observed covariates, typically a treatment and a control. Recently, propensity scores have been extended to allow for the comparison of multiple treatment groups. However, examples of the practical implementation of these multi-group procedures are relatively scarce. This study seeks to add to the current literature through implementing one of these strategies, TriMatch, in the comparison of three different therapeutic approaches to the treatment of problem gambling.
Problem gambling was selected as the subject matter of this analysis for several reasons. Problem gambling is a behavioral disorder that has a profound negative impact on a significant portion of the general population. It is also relatively understudied compared with other addictive behaviors, such as alcohol and substance use disorders. Also, with the expansion of legalized gambling in the United States, problem gambling is likely to receive an increased amount of attention over the coming years. Finally, the current body of literature deals very little with two of the three therapeutic approaches contained in this study: solution-focused brief therapy and time-limited dynamic psychotherapy.
Chapter Two: Method

Participants

Participants in this study were drawn from consecutive admissions of individuals seeking outpatient counseling services for problems related to their gambling behaviors. Inclusion criteria for participation included age: all were adults, at least 18 years of age with no upper bound for age. Participants also must have been appropriate recipients of outpatient psychotherapy. In particular, individuals were referred out of outpatient treatment and removed from the study if the severity of their psychopathology warranted more intensive treatment. A detailed description of the sample is contained in the results section (Table 2; Table 4; Table 5).

Procedure

This study utilized archival data from a university-based counseling clinic, which provided both gambling-specific and general counseling services to members of the community. Consecutive admissions for outpatient counseling services for gambling problems were invited to allow their clinical information to be used for research purposes.

Prior to treatment, participants were administered a series of psychological instruments. Each participant was administered the University of Rhode Island Change Assessment (URICA) to gather data on pre-treatment stages of change. At this time, The
Hands Depression Screen, The Mood Disorder Questionnaire, the Carroll-Davidson Generalized Anxiety Disorder Screen, and the Sprint-4 PTSD Screen were also administered to assess for the presence of four commonly observed co-occurring disorders. Pre-treatment severity of gambling problems was assessed through the NORC Diagnostic Screen for Gambling Problems (NODS). The NODS also allowed for classification of participants as pathological, problem, or at-risk gamblers. Therefore, prior to treatment data were gathered related to the severity of gambling problems, the presence of common co-occurring disorders, and participant’s readiness to change.

Also prior as a part of the pre-treatment paperwork, clients provided basic demographic information, including age and gender. Clients were also asked for racial/ethnic identity. However, this was not included in the analyses due the relatively homogeneity of the clients at this treatment setting. Specifically, the vast majority, approximately 85 percent, of the clients at this setting was non-hispanic/white, which fits with the tendency for ethnic and racial minorities to be underrepresented in gambling treatment settings (Volberg, 1994). Therefore, all other racial and ethnic groups composed only about one seventh of the total group, which would be insufficient for inferential testing of this factor.

Prior to each therapy session, including the first, each participant was administered the Outcome Questionnaire 45 (OQ-45) to track psychosocial functioning through the course of therapy. Therefore, the general functioning of the participants was continually assessed during treatment.
Treatment was delivered at a university-based outpatient counseling clinic, which has a focus on the treatment of problem and pathological gambling. However, treatment is available to all community members, not only members of the university. The clinic is a research and training facility for masters and doctoral level graduate students within the Counseling Psychology program at the university. Therefore, treatment for this study was provided by graduate students. These students received direct supervision related to the therapeutic approaches being administered from licensed psychologists. Further, the students were directly observed during therapy sessions through the use of closed circuit television systems.

Treatment consisted of one hour of individual psychotherapy per week, except when the university was closed. Treatment began with a semi-structured intake interview, which typically took one to two sessions to complete. Immediately following the intake interview, participants were allowed to select which therapeutic approach they desired to receive. The participants could choose between Cognitive-Behavioral Therapy (CBT), Time-Limited Dynamic Psychotherapy (TLDP), and Solution-Focused Brief Therapy (SFBT). Prior to making their selection, participants were provided with a description of the basic tenets of each form of therapy (Table 1). However, for a period, it had been the practice of the treatment center to also randomly assign clients to a form of therapy. As such, a smaller proportion of the clients in this study were not given the opportunity to select their treatment.

Following the intake interview and therapy selection, psychotherapy began in accordance with the form of therapy the participant chose or were assigned to. Treatment
plans were personalized in accordance with the therapeutic approach and the participants’ particular goals for therapy, as no standardized treatment protocols were utilized. There was no limit on the number of sessions of therapy a participant could attend. Similarly, there was no expectation or mandate regarding how many sessions a participant was to attend. Treatment continued until there was a mutual agreement between the client and therapist that therapy was no longer required or until the client discontinued attending therapy sessions.

Table 1

Description of Therapy Choices

<table>
<thead>
<tr>
<th>Therapy</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognitive-Behavioral Therapy</td>
<td>Cognitive Behavioral Therapy (CBT) is a goal oriented therapy that is active and directive in nature. The purpose of this therapy is to explore thoughts and behaviors that may cause you to engage in problematic behaviors. You and your counselor work together to develop new ways of thinking about problems, and you will learn new skills to deal with them. To help identify patterns, thought logs are used frequently in your session and between sessions. Your counselor will ask you to complete assignments and try change techniques that may be practiced throughout your week.</td>
</tr>
<tr>
<td>Time-Limited Dynamic Psychotherapy</td>
<td>Dynamic Therapy’s goals including improved relationships, attunement to feelings, and/or a resolution of a conflict. Your therapist focuses on the expression of emotions, and explores wishes, attitudes, and behaviors. Your therapist will help you to talk about yourself and your relationships to...</td>
</tr>
</tbody>
</table>
identify your expectations and repetitive patterns in your life and your relationships. The focus is often on resolving past experiences and prior traumas, and identifying expectations you have for yourself and others. You will be asked to think about yourself and relationships between sessions.

| Solution-Focused Brief Therapy | Solution-Focused Therapy is a goal-oriented therapy that focuses on helping you to clarify what is important to you, changes you would like to have in your life, and steps you might take to achieve your goals. This is an active therapy where your counselor and you will be working to identify your strengths and successes and will search with you for solutions to your present dilemma. There often is discussion on what small changes and steps will improve your life, and what to pay attention to and what to think about doing differently between sessions. You will be asked to notice any progress. |

**Instruments**

**URICA**

The University of Rhode Island Change Assessment (URICA) was developed as a scale to measure an individual’s readiness to, or stage of change (McConnaughy, Prochaska, & Velicer, 1983). The URICA comprises 32-items, with 8 items relating to each of the four measured stages of change: precontemplation, contemplation, action, and maintenance. Items include: “As far as I’m concerned, I don’t have any problems that need changing” and “I am finally doing some work on my problems.” The participant assigns a rating of 1 to 5 (1 = strong disagreement, 5 = strong agreement) to each item.
based on the relevance of the statement to their particular situation. Therefore, scores for each stage can range from 8 to 40, with higher scores reflecting greater endorsement of the characteristics of that stage. The instrument then assigns each participant to one of the four stages of change.

Confirmatory factor analysis has indicated that the questions load onto the theorized four-factor structure, which supports the validity of the instrument (Pantalon et al., 2002). The URICA has also demonstrated concurrent validity, as it has been found to correlate strongly with the Contemplation Ladder, another instrument that measures readiness to change \( (r=0.41) \) (Amodei & Lamb, 2004). The URICA has demonstrated acceptable to strong internal consistency estimates for each of the four subscales \( (\alpha's = 0.74-0.88) \) during administrations with treatment seeking pathological gamblers (Petry, 2005).

**NODS**

The National Opinion Research Center (NORC) DSM-IV Diagnostic Screen for Gambling Problems (NODS) is a 17-item instrument designed to measure gambling-related problems. The NODS directly assesses the ten diagnostic criteria for Pathological Gambling found in the Diagnostic and Statistical Manual of Mental Disorders, fourth edition with text revisions (DSM-IV-TR) (American Psychiatric Association, 2000; Gerstein et al., 1999). The NODS is comprised of yes-no questions such as: “Have you, during the past year, tried to stop, cut down or control your gambling?” and “Have there been periods during the past year when you needed to raise your bets in order to get the same feeling of excitement?” Consequently, scores on the NODS range from 0 to 10,
depending on how many criteria are endorsed. Higher scores indicate greater gambling-related problems. A score of 5 or higher indicates pathological gambling in accordance with the DSM-IV-TR. The NODS has demonstrated the ability to identify subclinical levels of disordered gambling behavior (Hodgins, 2004; Volberg, 2002). Scores of 3 or 4 indicate problem gambling.

Fager (2006) in a study of people seeking treatment for gambling problems found strong evidence of test-retest reliability of \( r = .77 \). Wickwire (2008) found the NODS to have strong internal consistency estimates (\( \alpha = .88 \)). The NODS has also demonstrated high concurrent validity \( (r = .85) \) with the South Oaks Gambling Screen, another common instrument designed to measure problematic gambling behavior (Wickwire, 2008).

**OQ-45**

The Outcome Questionnaire 45 (OQ-45) was designed as a measure of mental healthcare outcomes (Lambert et al., 1996). The instrument measures across three domains of psychosocial functioning: subjective discomfort, interpersonal relations, and social role performance. The OQ-45 provides a score for each of the three domains, in addition to a total score representing overall psychosocial functioning. The Outcome Questionnaire is a 45 item instrument, which uses a Likert-scale of response options. The instrument contains items such as: “I feel no interest in things,” “I feel irritated,” and “I feel lonely.” Total scores on the OQ-45 can range from 0 to 180. Lower scores on the OQ-45 indicate higher levels of psychosocial functioning. Conversely, higher scores indicate greater levels of distress. Scores at or above 63 are considered to indicate psychosocial distress that is of clinical significance.
Lambert et al. (1996) found the OQ-45 to have sound psychometric properties, including: high test-retest reliability estimate ($r = 0.84$), strong overall internal consistency ($\alpha = 0.93$), and moderate to strong concurrent validity estimates ($r = 0.60$ to $r = 0.88$) across a number of measures of psychosocial distress. Further analysis of the OQ-45 by Vermeersch, Lambert, and Burlingame (2000) has demonstrated the OQ-45 to be sensitive to an individual’s change in level of psychosocial functioning.

**WAI-S**

The Working Alliance Inventory- Short (WAI-S) is an abbreviated version of the Working Alliance Inventory, which was created to measure the three specific aspects of the therapeutic alliance: agreement on tasks, agreement on goals, and interpersonal bond between client and therapist (Horvath & Greenberg, 1989). The 12 items on the WAI-S are designed to measure how well the client and therapist are working together in therapy, including: “We agree on what is important for me to work on”, “I believe the way we are working with my problem is correct”, and “I believe _____ likes me”. The WAI-S is a 12-item instrument, the items on which are answered on a 7-point rating scale (1= never, 7 = always). A higher score on the WAI-S, therefore, indicates a stronger therapeutic alliance.

Evaluation of implementations of the WAI-S used to measure the client’s perceptions of the therapeutic alliance has demonstrated moderate to strong internal consistency estimates ($\alpha’s = .74–.90$; Busseri & Tyler, 2003). The WAI-S has also demonstrated evidence for concurrent validity, as it correlates highly with previously established instruments for measuring the therapeutic alliance, such as the Penn Helping
Alliance Questionnaire \((r = 0.75)\) and the California Psychotherapy Alliance Scale \((r = 0.80)\) (Hatcher & Gillaspy, 2006).

**HANDS**

The Harvard Department of Psychiatry/National Depression Screening Day Scale (HANDS) was designed to assess for a major depressive episode in accordance with the DSM-IV-TR criteria (Baer et al., 1999). The instrument enquires into the functioning of the individual during the preceding two weeks. Items on the instrument assess how often an individual is experiencing certain depressive symptoms, including: poor appetite, decreased interest, and suicidal ideation. This 10-item scale uses a 4-point response scale, with response options including: none or little of the time, some of the time, most of the time, and all of the time. Possible scores can range of 0 to 30. Higher scores indicate a greater number of symptoms and/or higher frequency of symptom occurrence. A score of 9 points or higher indicates a positive screen for a major depressive episode.

Through psychometric evaluation, the HANDS has demonstrated strong internal consistency \((\alpha = 0.87; \) Baer et al., 1999). Validity investigation found that the HANDS has a sensitivity of 0.95 and a specificity of 0.94, which was found in that study to be superior to the Beck Depression Inventory-II (Baer et al., 1999).

**MDQ**

The Mood Disorder Questionnaire (MDQ) assesses lifetime history of bipolar disorder (Hirschfeld et al., 2000). It contains 13 yes/no questions related to manic and hypomanic symptoms associated with bipolar disorder I and II. The instrument assesses if there has ever been a period during the individual’s life when they have exhibited
hypomanic symptoms, including: hyperactivity, racing thoughts, and decreased need for sleep. Possible scores range from 0 to 13, depending on number of endorsed symptoms. Scores of 7 points or higher constitute a positive screen for a mood disorder.

The MDQ has demonstrated high internal consistency ($\alpha=0.90$; Hirschfield et al., 2000). The MDQ has also demonstrated acceptable levels of sensitivity (0.73) and strong specificity (0.90) in the original validation study (Hirschfeld, 2010). In a review of numerous studies implementing the MDQ, it was found that the overall sensitivity was 0.61 and overall specificity was 0.88; however, the sensitivity was found to be significantly higher in studies that implemented the MDQ to clinical samples as opposed to the general population (Zimmerman & Galione, 2011).

**CDGAD**

The Carroll-Davidson Generalized Anxiety Disorder Screen (CDGAD) was designed to detect symptoms of generalized anxiety from the previous six months in accordance the DSM-IV-TR criteria for Generalized Anxiety Disorder (Carroll & Davidson, 2000). This 12-item instrument asks individuals if for most days during the past six months they have experienced anxious symptoms, including: nervous feeling, excessive worry, irritability, and difficulty sleeping. Each item has yes/no response options. Therefore, total scores range from 0 to 12, with higher scores indicating greater symptom endorsement. Scores of 6 or higher constitute a positive screen for Generalized Anxiety Disorder. An evaluation of this instrument found it to exhibit moderate internal consistency ($\alpha = 0.82$; Leyton-Armakan et al., 2012). This instrument has also
demonstrated a sensitivity of 0.64 and specificity of 0.90 (Carroll & Davidson, 2000; Screening for Mental Health, 2014).

**SPRINT-4**

The SPRINT-4 is an abbreviated version of the Short Posttraumatic Stress Disorder Rating Interview (SPRINT; Connor & Davidson, 2001). The instrument begins by asking if the participant has experienced or witnessed a traumatic event that involved serious injury, loss of life, or significant threat of injury. Then, four yes/no questions assess whether the individual, as a consequence of that event, is currently experiencing symptoms of Posttraumatic Stress Disorder (PTSD), such as intrusive memories of the event. Moreover, each item corresponds to one of the four PTSD symptom clusters (intrusive, avoidance, numbing, and hyperarousal). Affirmative responses to 2 or more of the items qualify as a positive screen for PTSD. The SPRINT has demonstrated strong concurrent validity with the Clinician-Administered Scale for DSM-IV (CAPS) with correlations between the two instruments found to range between 0.64 and 0.79 (Vaishnavi, Payne, Connor, & Davidson, 2006). A Romanian version of the SPRINT demonstrated very strong internal consistency within a clinical sample ($r = 0.90$), as well as strong sensitivity (.88) and specificity (.90; Herta, Nemes, & Cozman, 2013).

**Analytical Strategy**

R version 3.0.2 and SPSS 22 were used for the statistical analyses.

**Missing Values Analysis**

Given the nature of this study as an implementation of propensity score analysis with a relatively small sample, missing values were imputed rather than listwise deleting
incomplete cases. Prior to imputation, missing values had to be tested for whether or not they were missing completely at random. Little’s Missing Completely At Random (MCAR) Test was used to determine if missing values were occurring completely at random (Little, 1988). When data is MCAR the missing values have no relationship with any of the observed values within the dataset, which means that the missing values can be imputed without introducing bias to the eventual analyses (McKnight, McKnight, Sidani, Figueredo, 2007).

Expectation Maximization (EM) was utilized to replace the missing values. EM is an iterative procedure that utilizes maximum likelihood (ML) estimation. This process consists of estimating missing values using ML, generating parameter estimates, and this process iterates until finally converging upon a solution (McKnight et al., 2007). EM techniques have been shown effective in calculating relatively unbiased estimates for missing values (Gold & Bentler, 2000; Liu & Brown, 2013)

**Research Question One**

The first component of this research sought to explore which variables were contributing to the participants’ particular therapeutic selections. Particularly given the exploratory nature of this research question, the following variables were analyzed in individual multinomial logistic regression models to determine their relationship with therapeutic selection: age, gender, stages of change, co-occurring psychological disorders (depression, post-traumatic stress disorder, anxiety, and mood disorder), and severity of gambling problems. To explore which client characteristics might be related to treatment selected groups based on treatment selected were compared for equivalence across these
variables. For the continuous variables, one-way ANOVAs were conducted to test whether means of these variables were equivalent between groups. For the categorical variables, chi-square tests of independence were run to determine if proportions varied between groups. Variables whose p-value was less than 0.25 were retained and included in the propensity score estimation model.

**Research Question Two**

This component of the research sought to determine if the participants as a whole were selecting particular forms of therapy disproportionately to the other forms. Therefore, a chi-square goodness of fit test was run to assess whether or not the three therapies were selected in equal proportions. A chi-square of 5.991 with 2 degrees of freedom (p<0.05) was used as the critical value at which differences in proportions of therapeutic selections was to be considered statistically significantly different.

**Research Question Three**

For the second component of the research, to determine whether or not the three therapeutic approaches demonstrate equally efficacious short-term outcomes, change in psychosocial functioning, as measured by the OQ-45, from the first to the fifth treatment session was compared across the three treatment groups. Since the majority of participants were allowed to choose the type of treatment they receive, propensity score matching was implemented to correct for any potential selection bias within the sample. The TriMatch statistical package within R was designed to allow for matching of three groups, as opposed to traditional propensity score matching techniques that are designed to only match participants from two groups (Bryer, 2013). Similar to other propensity
score matching methods, the TriMatch algorithm uses logistic regression to estimate propensity scores. In the estimation of propensity scores, three logistic regression models are calculated. The first model estimates the probability of being in the CBT group, as opposed to SFBT. The second model estimates the probability of being in the CBT group, as opposed to TLDP. The third model estimates the probability of being in the SFBT group, as opposed to TLDP. The matching system then finds trios of participants, one from each group, based on minimizing the within-trio differences in propensity scores.

The Trimatch statistical package includes various matching options (Bryer, 2013). The first procedure is analogous to matching without replacement in two-group propensity score matching. This procedure, which is referred to Maximum Treat matching within Trimatch, will only match a treatment 1 unit for a second time if a treatment 2 unit would be otherwise left unmatched. The second procedure, the caliper matching, allows for matching of all possible matched triplets within a specified caliper. In other words, the program will create all matches that meet a certain criteria for closeness of match. The default caliper is 0.25 standard deviations of the propensity scores. The third and fourth procedures are analogous to the 1:n procedures in two-group propensity score matching. These procedures utilize 2:1:n and 3:2:n matching, respectively. These final two procedures specify how many times each of the first two groups may be reused within the matching. It should be noted that the default in Trimatch is to order the groups in terms of sample size. Therefore, the 3:2:n allowed participants in the largest group to be matched up to three times and participants in the second largest group to be matched up to two times. Each of these matching procedures were utilized
and compared in terms of level of effectiveness in balancing the groups across the identified covariates.

After the matching has occurred, this package conducted a repeated measures ANOVA comparing mean OQ-45 scores across groups. If the main effect of time was found to be significant at \(p<0.05\), the package automatically allowed for dependent samples \(t\)-tests to explore specific group differences. When significant pairwise differences were detected, sensitivity analyses were conducted to test the robustness of these findings to potential confounding influences of unobserved covariates (Rosenbaum, 1987).

To compare matching strategies, three criteria were assessed: proportion of subjects used in the matching, number of matched triplets created, and balance of the covariates from the propensity score estimation models. Balance will be assessed through calculating statistical significance of covariate differences across groups following matching. Balance will also be assessed through an inspection of the reduction in bias through matching (Cochran & Rubin, 1973). Bias, in this context, was defined as the sum of the absolute values of the pairwise comparisons between groups on a given covariate. Percent bias reduced is defined as the percentage of the pre-match bias that is reduced through the matching procedure. For instance, a bias reduced value of 0.75 indicates that the post-matching bias is 75 percent less than the pre-match bias. It was calculated by subtracting the post-matching bias from the pre-match bias, then dividing that amount by the pre-match bias. The final way in which balance was assessed was through the calculation of standardized post-matching bias (Caliendo & Kopeinig, 2008). Following
matching, the pairwise standardized differences were calculated by dividing each pairwise difference for a given covariate divided by the pooled standard deviation for the groups being compared. These standardized pairwise differences were then averaged to achieve an overall standardized bias for each covariate. These values were then averaged across each covariate within a given matching strategy. These overall standardized biases were also used to compare the matching procedures.
Chapter Three: Results

Missing Values Analysis

Prior to data imputation, missing values were assessed for appropriateness for replacement. Table 2 contains the proportions of missing values for each variable included in subsequent analyses. Proportions of missing values ranged from 0 to 6.9 percent. The most commonly missing variables were drawn from the questionnaire related to the co-occurring mental health disorders. Data were also imputed using OQ-45 values from the first six sessions. Each of the three cases missing the OQ-45 score from the fifth session provided an OQ-45 score from a sixth session.

The results of the Little’s MCAR test were not statistically significant ($X^2(175) = 183.04, p = 0.323$; Little, 1988). This indicates that the missing values are likely to be occurring completely at random.
Table 2

**Missing Values**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean (SD)</th>
<th>N</th>
<th>% Missing</th>
<th>Imputed Values</th>
<th>Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>45.91 (11.90)</td>
<td>101</td>
<td>0</td>
<td>45.91 (11.90)</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>59.8% Male</td>
<td>101</td>
<td>0</td>
<td>59.8% Male</td>
<td></td>
</tr>
<tr>
<td>URICA&lt;sup&gt;1&lt;/sup&gt;</td>
<td>10.16 (1.71)</td>
<td>96</td>
<td>5.0</td>
<td>10.12 (1.70)</td>
<td></td>
</tr>
<tr>
<td>OQ first session&lt;sup&gt;2&lt;/sup&gt;</td>
<td>68.25 (20.27)</td>
<td>100</td>
<td>1.0</td>
<td>68.10 (20.22)</td>
<td></td>
</tr>
<tr>
<td>OQ fifth session&lt;sup&gt;2&lt;/sup&gt;</td>
<td>57.51 (22.36)</td>
<td>98</td>
<td>3.0</td>
<td>58.23 (22.60)</td>
<td></td>
</tr>
<tr>
<td>WAI-S&lt;sup&gt;3&lt;/sup&gt;</td>
<td>10.02 (2.23)</td>
<td>98</td>
<td>3.0</td>
<td>10.00 (2.20)</td>
<td></td>
</tr>
<tr>
<td>HANDS&lt;sup&gt;4&lt;/sup&gt;</td>
<td>55.32% positive</td>
<td>94</td>
<td>6.9</td>
<td>56.44% positive</td>
<td></td>
</tr>
<tr>
<td>CD-GAD&lt;sup&gt;5&lt;/sup&gt;</td>
<td>55.32% positive</td>
<td>94</td>
<td>6.9</td>
<td>56.44% positive</td>
<td></td>
</tr>
<tr>
<td>SPRINT-4&lt;sup&gt;6&lt;/sup&gt;</td>
<td>47.87% positive</td>
<td>94</td>
<td>6.9</td>
<td>49.50% positive</td>
<td></td>
</tr>
<tr>
<td>MDQ&lt;sup&gt;7&lt;/sup&gt;</td>
<td>35.11% positive</td>
<td>94</td>
<td>6.9</td>
<td>34.65% positive</td>
<td></td>
</tr>
<tr>
<td>NODS&lt;sup&gt;8&lt;/sup&gt;</td>
<td>8.09 (2.12)</td>
<td>101</td>
<td>0</td>
<td>8.09 (2.12)</td>
<td></td>
</tr>
</tbody>
</table>

<sup>1</sup> University of Rhode Island Change Assessment; <sup>2</sup> Outcome Questionnaire-45; <sup>3</sup> Working Alliance Inventory- Short Form; <sup>4</sup> Harvard Department of Psychiatry National Depression Day Screening Scale(% positive screen); <sup>5</sup> Carroll-Davidson Generalized Anxiety Disorder Screen (% positive screen); <sup>6</sup> Short Post Traumatic Rating Interview(% positive screen); <sup>7</sup> Mood Disorder Questionnaire(% positive screen); <sup>8</sup> NORC DSM-IV Diagnostic Screen for Gambling Problems.
Characteristics of the Sample

The characteristics of the sample, following the expectation maximization imputation are provided in Table 2. Of the 101 individuals included in the analyses, 61 identified as male. The average age for the sample was 45.9 years. The mean NODS score (8.09) indicates that overall this sample consisted of individuals with relatively severe gambling problems. As measured by the NODS, 93.1 percent of the sample met the DSM-IV-TR criteria for pathological gambling, and each participant in the sample endorsed at least one criterion for pathological gambling.

When comparing gender differences within the sample, the female participants were significantly older than their male counterparts (t(96.56)=3.28, p=0.001). Females presented with significantly more severe gambling problems, as measured by the NODS (t(96.01)=3.17, p=0.002). Females also presented with more severe psychosocial difficulties, as measured by their first session OQ-45 score (t(99)=2.20, p=0.03).

Also notable is the prevalence of co-occurring disorders in this sample. Only 17.8 percent of the sample did not screen positively for any of the co-occurring mental health disorders. In fact, 60.4 percent of the sample screen positively for at least two of the co-occurring disorders. Depression and Generalized Anxiety were the two most commonly observed co-occurring mental health disorders.

Assessing Treatment Selection

The first two research questions sought to identify characteristics of the problem gambling participants that may be influencing their selection of treatments, as well as identify which, if any, of the treatments are being selected disproportionately to the
others. To address this latter question, a chi-square goodness of fit test was run to determine if the treatments were being selected in equal proportions.

Of the 101 participants included in the analysis of outcomes, 79 (78.2%) were allowed to select their own form of therapy as described within the method section of this paper. The following set of analyses was conducted on this subset of the sample in order to assess treatment preferences, as well as the factors contributing to treatment selection.

Table 3 contains the observed proportions of therapies that were selected. Both the cognitive-behavioral therapy and the solution-focused brief therapy were selected more often than would be expected, while the time-limited dynamic psychotherapy was selected less often than expected. However, the results of the chi-square goodness of fit test was not statistically significant ($X^2(2)=4.13$, $p=0.13$). Therefore, the data indicate that the therapies were selected in statistically equivalent proportions.

Table 3

<table>
<thead>
<tr>
<th>Therapy Type</th>
<th>Observed proportion</th>
<th>Observed Frequency</th>
<th>Expected Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>TLDP</td>
<td>22.8%</td>
<td>18</td>
<td>26.3</td>
</tr>
<tr>
<td>SFBT</td>
<td>36.7%</td>
<td>29</td>
<td>26.3</td>
</tr>
<tr>
<td>CBT</td>
<td>40.5%</td>
<td>32</td>
<td>26.3</td>
</tr>
</tbody>
</table>

Comparing Treatment Selection Groups

The comparisons between groups based on treatment selection are summarized in Table 4. As the table indicates, none of the individual observed baseline characteristics statistically varied statistically significantly between groups at the $p<0.05$ level.
However, when viewed as a whole the results suggest that groups may vary in terms of overall clinical presentation. Specifically, the SFBT group had the lowest mean scores on initial OQ-45 and NODS, as well as had the lowest proportion screening positively for depression, generalized anxiety, and mood disorder. The selection of TLDP appears to be favored by younger, male therapy clients.

Table 4

*Baseline Covariates Across Treatment Selection Groups*

<table>
<thead>
<tr>
<th></th>
<th>CBT n=32</th>
<th>SFBT n=29</th>
<th>TLDP n=18</th>
<th>Test Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>45.51</td>
<td>47.17</td>
<td>39.61</td>
<td>F(2,76) = 2.24, p = 0.11</td>
</tr>
<tr>
<td>Gender (% Male)</td>
<td>56.25</td>
<td>58.62</td>
<td>77.78</td>
<td>X²(2) = 2.49, p = 0.29</td>
</tr>
<tr>
<td>URICA</td>
<td>9.88</td>
<td>10.27</td>
<td>10.38</td>
<td>F(2,76) = 0.61, p = 0.55</td>
</tr>
<tr>
<td>OQ-45 first session</td>
<td>70.16</td>
<td>62.24</td>
<td>68.72</td>
<td>F(2,76) = 1.24, p = 0.30</td>
</tr>
<tr>
<td>WAI-S</td>
<td>9.63</td>
<td>9.92</td>
<td>10.85</td>
<td>F(2,76) = 1.72, p = 0.19</td>
</tr>
<tr>
<td>HANDS (% positive)</td>
<td>62.50</td>
<td>51.72</td>
<td>55.56</td>
<td>X²(2) = 0.74, p = 0.69</td>
</tr>
<tr>
<td>CD-GAD (% positive)</td>
<td>62.50</td>
<td>37.93</td>
<td>61.11</td>
<td>X²(2) = 4.28, p = 0.12</td>
</tr>
<tr>
<td>SPRINT-4 (% positive)</td>
<td>56.25</td>
<td>44.83</td>
<td>44.44</td>
<td>X²(2) = 1.02, p = 0.60</td>
</tr>
<tr>
<td>MDQ (% positive)</td>
<td>34.38</td>
<td>20.69</td>
<td>44.44</td>
<td>X²(2) = 3.08, p = 0.21</td>
</tr>
<tr>
<td>NODS</td>
<td>8.44</td>
<td>7.55</td>
<td>8.11</td>
<td>F(2,76) = 1.16, p = 0.32</td>
</tr>
</tbody>
</table>
Comparing Overall Treatment Groups

Selecting variables for the propensity score estimation model involved two steps. The first step involved comparisons of baseline characteristics between the overall treatment groups. Similar to the previous analysis, one-way ANOVAs were used to compare continuous variables, and chi-square tests of independence were used for categorical covariates. The clinical and demographic covariates were also tested for their relationship with the outcome variable. These analyses include both those that were randomly assigned to a form of treatment (n=22) and those that selected their own form of treatment (n=79).

The results of these comparisons are summarized in Table 5. Given that the majority of participants selected their own treatment, it is not surprising that the results are similar to those contained in Table 4. Once again, no covariates were found to significantly vary between groups at the p<0.05 level. Considered in totality, the Solution-Focused Brief Therapy group again had a less severe clinical presentation compared to the other treatments. Also, the Time-Limited Dynamic Psychotherapy group was again found to be marginally different in both age and gender.

The largest difference between the comparison of selection groups and the comparison of overall treatment groups is with the Working Alliance Inventory measure. The analysis of selection groups found this variable to marginally significantly differ across groups (p = 0.19; Table 4). However, in the comparison of overall treatment groups, this difference becomes virtually non-existent (p = 0.60; Table 5).
### Table 5

*Comparing Covariate Balance Across Treatment Groups*

<table>
<thead>
<tr>
<th></th>
<th>CBT n=45</th>
<th>SFBT n=32</th>
<th>TLDP n=24</th>
<th>Test Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>46.53</td>
<td>47.81</td>
<td>42.25</td>
<td>F(2,98) = 1.61, p = 0.21</td>
</tr>
<tr>
<td>Gender (% Male)</td>
<td>53.33</td>
<td>59.38</td>
<td>75.00</td>
<td>X^2(2) = 3.09, p = 0.21</td>
</tr>
<tr>
<td>URICA</td>
<td>10.06</td>
<td>10.25</td>
<td>10.06</td>
<td>F(2,98) = 0.13, p = 0.88</td>
</tr>
<tr>
<td>OQ first session</td>
<td>70.67</td>
<td>63.03</td>
<td>70.04</td>
<td>F(2,98) = 1.49, p = 0.23</td>
</tr>
<tr>
<td>WAI-S</td>
<td>9.90</td>
<td>9.85</td>
<td>10.41</td>
<td>F(2,98) = 0.52, p = 0.60</td>
</tr>
<tr>
<td>HANDS (% positive)</td>
<td>60.00</td>
<td>50.00</td>
<td>58.33</td>
<td>X^2(2) = 0.81, p = 0.67</td>
</tr>
<tr>
<td>CD-GAD (% positive)</td>
<td>62.22</td>
<td>43.75</td>
<td>62.50</td>
<td>X^2(2) = 3.07, p = 0.22</td>
</tr>
<tr>
<td>SPRINT-4 (% positive)</td>
<td>53.33</td>
<td>46.88</td>
<td>45.83</td>
<td>X^2(2) = 0.48, p = 0.79</td>
</tr>
<tr>
<td>MDQ (% positive)</td>
<td>33.33</td>
<td>28.13</td>
<td>45.83</td>
<td>X^2(2) = 1.96, p = 0.38</td>
</tr>
<tr>
<td>NODS</td>
<td>8.38</td>
<td>7.56</td>
<td>8.25</td>
<td>F(2,98) = 1.49, p = 0.23</td>
</tr>
</tbody>
</table>

### Relationships Between Covariates and Outcome

The second phase of identifying covariates for the propensity score model was to identify baseline characteristics that are related to the outcome. Table 6 contains the bivariate correlations between each covariate and the outcome variable (the difference between OQ-45 scores from the first and fifth sessions) for the total sample (N=101). None of the ten observed characteristics was significantly related to the outcome variable.
at the p<0.05 level. Only one variable, OQ-45 score from the first session, was related at p<0.25.

Table 6

*Bivariate Correlations Between Baseline Characteristics and Outcome Measure*

<table>
<thead>
<tr>
<th></th>
<th>correlation (r)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.09</td>
<td>0.38</td>
</tr>
<tr>
<td>Gender</td>
<td>0.07</td>
<td>0.48</td>
</tr>
<tr>
<td>URICA</td>
<td>0.09</td>
<td>0.38</td>
</tr>
<tr>
<td>OQ first session</td>
<td>0.16</td>
<td>0.11</td>
</tr>
<tr>
<td>WAI-S</td>
<td>-0.02</td>
<td>0.85</td>
</tr>
<tr>
<td>HANDS (positive screen)</td>
<td>-0.01</td>
<td>0.89</td>
</tr>
<tr>
<td>CD-GAD (positive screen)</td>
<td>0.04</td>
<td>0.67</td>
</tr>
<tr>
<td>SPRINT-4 (positive screen)</td>
<td>-0.04</td>
<td>0.66</td>
</tr>
<tr>
<td>MDQ (positive screen)</td>
<td>-0.03</td>
<td>0.73</td>
</tr>
<tr>
<td>NODS</td>
<td>0.09</td>
<td>0.39</td>
</tr>
</tbody>
</table>

**Unmatched Analysis**

For comparison, an analysis was run comparing the three treatment groups on the outcome variable without matching or covariate adjustment. A one-way ANOVA was used to compare the means of the dependent variable, amount of change in OQ-45 scores from the first to the fifth treatment sessions, across treatment groups. The results indicate that this analysis was unable to detect any differences between groups (F(2,98) = 0.93, p = 0.40).
Propensity Score Analyses

The following sections include the tests involved in specifying the propensity score estimation models, as well as the results from each of the four matching procedures.

Specifying the Propensity Score Model

In accordance with the *a priori* criteria set in the proposed analytic strategy, covariates that differed between groups at the p<0.25 level were retained for the propensity score estimation model. Additionally, covariates correlated with the outcome at the p<0.25 level were included in the propensity score model. From the comparison of covariate balance across treatment groups the following covariates were identified: age, gender, first session OQ-45 score, CD-GAD diagnostic classification, and NODS score. The only variable identified for its correlation with the outcome, first session OQ-45 score, was already selected for inclusion based on the comparison of treatment groups. Therefore, propensity scores were estimated using these five covariates.

The propensity score matching procedures would then seek to balance the treatment groups, specifically as related to these five covariates. In order to assess the effectiveness of the balance created by each matching method, a baseline level of bias was calculated for each covariate. Bias, in accordance with the method utilized by Bai (2013), was calculated by summing the absolute value of the pairwise differences between each group for each covariate (Table 7).
Table 7

Pairwise Differences of Covariates Prior to Matching

<table>
<thead>
<tr>
<th></th>
<th>CBT-SFBT</th>
<th>CBT-TLDP</th>
<th>SFBT-TLDP</th>
<th>Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>-1.28</td>
<td>4.28</td>
<td>5.56</td>
<td>11.12</td>
</tr>
<tr>
<td>Gender</td>
<td>-0.06</td>
<td>-0.22</td>
<td>-0.16</td>
<td>0.44</td>
</tr>
<tr>
<td>Initial OQ-45</td>
<td>7.64</td>
<td>0.63</td>
<td>-7.01</td>
<td>15.28</td>
</tr>
<tr>
<td>Anxiety</td>
<td>-0.18</td>
<td>0.00</td>
<td>-0.19</td>
<td>0.37</td>
</tr>
<tr>
<td>NODS</td>
<td>0.82</td>
<td>0.13</td>
<td>-0.69</td>
<td>1.64</td>
</tr>
</tbody>
</table>

Covariate Balance Across Propensity Estimation Models

In the estimation of propensity scores, three logistic regression models are calculated. The first model estimates the probability of being in the CBT group, as opposed to SFBT. The second model estimates the probability of being in the CBT group, as opposed to TLDP. The third model estimates the probability of being in the SFBT group, as opposed to TLDP.

Prior to any matching procedures, covariate balance was tested using the propensity scores from these three models stratified across quintiles. Figure 1 illustrates the post matching covariance balance for all three propensity score models. Column 1 on Figure 1 illustrates the logistic regression model created between CBT and SFBT, column 2 of Figure 1 illustrates the model created between the CBT and TLDP groups, and column 3 illustrates the balance created from the model containing SFBT and TLDP. The red line illustrates the un-adjusted effect sizes of the covariate imbalance, whereas the blue line illustrates the adjusted biases resulting from the stratified comparison. As can be seen, the balance is improved for all covariates across each treatment combination,
except in one instance. The proportions of positive anxiety screens were marginally less balanced when stratified by propensity scores, but this covariate was already well balanced between the CBT and TLDP groups prior to matching (Table 1, column 2).

Figure 1

*Covariate Balance From the Three Propensity Score Estimation Models*

<table>
<thead>
<tr>
<th>Column 1: the model containing CBT and SFBT clients</th>
<th>Column 2: the model containing CBT and TLDP clients</th>
<th>Column 3: the model containing SFBT and TLDP clients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum Treat Matching</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Following the matching procedure, 28 matched triplets were created. Overall, 61 (60.4%) of the subjects were utilized with this matching procedure. Table 8 summarizes
the unmatched individuals by treatment group. Only those cases involved in the matching process were included in the following set of analyses.

Table 8

*Unmatched Participants Within Maximum Treat Matching*

<table>
<thead>
<tr>
<th></th>
<th>Number unmatched</th>
<th>% unmatched</th>
</tr>
</thead>
<tbody>
<tr>
<td>CBT</td>
<td>26</td>
<td>57.8</td>
</tr>
<tr>
<td>SFBT</td>
<td>9</td>
<td>28.1</td>
</tr>
<tr>
<td>TLDP</td>
<td>5</td>
<td>20.8</td>
</tr>
<tr>
<td>Total</td>
<td>40</td>
<td>39.6</td>
</tr>
</tbody>
</table>

The values for each covariate by group are contained in Table 9. The results of the inferential tests of covariate balance are contained in Table 10. As was also the case prior to matching, none of the covariates significantly differed at p<0.05 after matching.

Table 9

*Covariate Values Following Maximum Treat Matching*

<table>
<thead>
<tr>
<th></th>
<th>CBT</th>
<th>SFBT</th>
<th>TLDP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>44.21</td>
<td>44.79</td>
<td>45.64</td>
</tr>
<tr>
<td>Gender¹</td>
<td>0.43</td>
<td>0.36</td>
<td>0.43</td>
</tr>
<tr>
<td>Initial OQ-45</td>
<td>67.25</td>
<td>66.04</td>
<td>72.50</td>
</tr>
<tr>
<td>Anxiety²</td>
<td>0.64</td>
<td>0.54</td>
<td>0.64</td>
</tr>
<tr>
<td>NODS</td>
<td>7.82</td>
<td>8.54</td>
<td>8.43</td>
</tr>
</tbody>
</table>

¹: proportion male; ²: proportion screen positively for generalized anxiety by the CD-GAD.
The reduction in bias was calculated for each covariate following the Maximum Treat matching (Table 11). The largest bias reduction was observed for age and gender with 74% and 68%, respectively. The matching was less effective for the other covariates, particularly initial OQ-45 scores and NODS scores, with 9% and 12% bias reductions, respectively. The statistical significance of the between-group differences of NODS scores remained marginally significant despite a modest reduction in bias (Table 10). Boxplots of each covariate by group, as well as boxplots of the pairwise comparisons following Maximum Treat matching are contained in Appendix A.

Table 10

Inferential Testing of Maximum Treat Post-Matching Covariate Balance

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Test Statistic</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>F(2,54) = 0.20</td>
<td>p = 0.82</td>
</tr>
<tr>
<td>Gender</td>
<td>X^2(2) = 0.80</td>
<td>p = 0.67</td>
</tr>
<tr>
<td>OQ</td>
<td>F(2,54) = 0.85</td>
<td>p = 0.43</td>
</tr>
<tr>
<td>Anxiety</td>
<td>X^2(2) = 1.06</td>
<td>p = .59</td>
</tr>
<tr>
<td>NODS</td>
<td>F(2,54) = 1.70</td>
<td>p = 0.19</td>
</tr>
</tbody>
</table>
Table 11

*Pairwise Differences of Covariate Balance Following Maximum Treat Matching*

<table>
<thead>
<tr>
<th>Variable</th>
<th>CBT-SFBT</th>
<th>CBT-TLDP</th>
<th>SFBT-TLDP</th>
<th>Bias</th>
<th>Stand. Bias</th>
<th>% Bias reduced</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>-0.58</td>
<td>-1.43</td>
<td>-0.85</td>
<td>2.86</td>
<td>0.09</td>
<td>0.74</td>
</tr>
<tr>
<td>Gender</td>
<td>0.07</td>
<td>0.00</td>
<td>-0.07</td>
<td>0.14</td>
<td>0.09</td>
<td>0.68</td>
</tr>
<tr>
<td>Initial OQ-45</td>
<td>1.21</td>
<td>-5.25</td>
<td>-6.46</td>
<td>13.92</td>
<td>0.20</td>
<td>0.09</td>
</tr>
<tr>
<td>Anxiety</td>
<td>0.1</td>
<td>0.00</td>
<td>-0.1</td>
<td>0.20</td>
<td>0.14</td>
<td>0.46</td>
</tr>
<tr>
<td>NODS</td>
<td>-0.72</td>
<td>-0.61</td>
<td>0.11</td>
<td>1.44</td>
<td>0.28</td>
<td>0.12</td>
</tr>
</tbody>
</table>

Following the matching procedure, a repeated measures ANOVA was run comparing the three treatment groups on the dependent variable, the amount of change in OQ-45 scores between the first and fifth sessions. The results of the test were statistically significant (F(2,54) = 3.37, p = 0.042. Given the statistical significance, a series of dependent sample t-tests were automatically conducted through Trimatch to assess pairwise comparisons of treatment effectiveness (Table 12). These results suggest that the CBT treatment was significantly more effective than the TLDP treatment. The outcome values for each treatment group are presented in Figure 2. The boxplots in Figure 2 contain a box and whisker plot containing the median and quartile ranges, as well as a red circle indicating the mean and green bars illustrating the 95% confidence interval about the mean. The pairwise comparisons are illustrated in Figure 3 in the same fashion as the plots in Figure 2, except the standard error is illustrated by a solid green box and the pairwise differences are listed on the graph.
Table 12

Post-hoc Pairwise Comparisons of Outcome Following Maximum Treat Matching

<table>
<thead>
<tr>
<th></th>
<th>Value</th>
<th>Test Statistic</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SFBT-TLDP</td>
<td>4.9</td>
<td>t(27) = 1.28</td>
<td>0.21</td>
</tr>
<tr>
<td>SFBT-CBT</td>
<td>-4.3</td>
<td>t(27) = -1.31</td>
<td>0.20</td>
</tr>
<tr>
<td>TLDP-CBT</td>
<td>-9.2</td>
<td>t(27) = -2.62</td>
<td>0.01</td>
</tr>
</tbody>
</table>

Figure 2

Boxplots of OQ-45 Change for Each Group After Maximum Treat Matching
A sensitivity analysis was also conducted to test the robustness of the CBT-TLDP pairwise comparison against bias introduced by unobserved variables to the propensity score estimation models (Rosenbaum, 1987; Table 13). The results suggest the findings remain robust to Gamma < 1.4. In other words, an unobserved variable(s) would have to change the odds of assignment by a factor of 1.4 before the finding would become non-significant.
Table 13

*Sensitivity Analysis for CBT-TLDP Comparison Following Maximum Treat Matching*

<table>
<thead>
<tr>
<th>Gamma</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>0.01</td>
</tr>
<tr>
<td>1.1</td>
<td>0.02</td>
</tr>
<tr>
<td>1.2</td>
<td>0.03</td>
</tr>
<tr>
<td>1.3</td>
<td>0.04</td>
</tr>
<tr>
<td>1.4</td>
<td>0.06</td>
</tr>
</tbody>
</table>

**Caliper Matching**

The caliper matching resulted in the largest number of matched triplets (126). It also resulted in the largest portion of the overall sample being utilized in the matching procedure (79.2%). The unmatched participants by group are summarized in Table 14. Only those cases involved in the matching process were included in the following set of analyses.

Table 14

*Unmatched Participants Within Caliper Matching*

<table>
<thead>
<tr>
<th></th>
<th>Number unmatched</th>
<th>% unmatched</th>
</tr>
</thead>
<tbody>
<tr>
<td>CBT</td>
<td>7</td>
<td>15.6</td>
</tr>
<tr>
<td>SFBT</td>
<td>9</td>
<td>28.1</td>
</tr>
<tr>
<td>TLDP</td>
<td>5</td>
<td>20.8</td>
</tr>
<tr>
<td>Total</td>
<td>21</td>
<td>20.8</td>
</tr>
</tbody>
</table>

The values for each covariate by treatment group following the caliper matching are contained in Table 15. None of the covariates differed significantly across groups at
p<0.05; however, the covariates for positive anxiety screens and NODS scores approached statistical significance (Table 16). The fact that between-group differences became closer to statistically significant following a matching procedure that reduced bias for the anxiety and NODS covariates by 35% and 51%, respectively, is likely a result of the greatly increased statistical power resulting from the large number of matched triplets created with this procedure (Table 17). Boxplots of each covariate by group, as well as boxplots of the pairwise comparisons following caliper matching are contained in Appendix B.

Table 15

*Covariate Values Following Caliper Matching*

<table>
<thead>
<tr>
<th></th>
<th>CBT</th>
<th>SFBT</th>
<th>TLDP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>45.47</td>
<td>44.47</td>
<td>46.13</td>
</tr>
<tr>
<td>Gender</td>
<td>0.49</td>
<td>0.49</td>
<td>0.53</td>
</tr>
<tr>
<td>Initial OQ-45</td>
<td>71.74</td>
<td>70.71</td>
<td>73.62</td>
</tr>
<tr>
<td>Anxiety</td>
<td>0.78</td>
<td>0.66</td>
<td>0.69</td>
</tr>
<tr>
<td>NODS</td>
<td>8.51</td>
<td>8.76</td>
<td>8.36</td>
</tr>
</tbody>
</table>

Table 16

*Inferential Testing of Caliper Post-Matching Covariate Balance*

<table>
<thead>
<tr>
<th></th>
<th>Test Statistic</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>F(2,250) = 1.35</td>
<td>p = 0.26</td>
</tr>
<tr>
<td>Gender</td>
<td>X²(2) = 1.43</td>
<td>p = 0.49</td>
</tr>
<tr>
<td>OQ</td>
<td>F(2,250) = 0.85</td>
<td>p = 0.43</td>
</tr>
<tr>
<td>Anxiety</td>
<td>X²(2) = 5.66</td>
<td>p = 0.06</td>
</tr>
<tr>
<td>NODS</td>
<td>F(2,250) = 2.83</td>
<td>p = 0.06</td>
</tr>
</tbody>
</table>
Bias reduced for each covariate following the caliper matching is outlined in Table 17. For four of the five covariates, bias was reduced by more than 50%. The only exception was proportion of positive anxiety screens, as measured by the CD-GAD, which had a bias reduction of 35%.

Table 17

*Pairwise Differences of Covariate Balance Following Caliper Matching*

<table>
<thead>
<tr>
<th></th>
<th>CBT-SFBT</th>
<th>CBT-TLDP</th>
<th>SFBT-TLDP</th>
<th>Bias</th>
<th>Stand. Bias</th>
<th>Bias reduced</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1.00</td>
<td>-0.66</td>
<td>-1.66</td>
<td>3.32</td>
<td>0.11</td>
<td>0.70</td>
</tr>
<tr>
<td>Gender</td>
<td>-0.04</td>
<td>0.00</td>
<td>0.04</td>
<td>0.08</td>
<td>0.05</td>
<td>0.81</td>
</tr>
<tr>
<td>Initial OQ-45</td>
<td>1.03</td>
<td>-1.88</td>
<td>-2.91</td>
<td>5.82</td>
<td>0.09</td>
<td>0.62</td>
</tr>
<tr>
<td>Anxiety</td>
<td>0.12</td>
<td>0.09</td>
<td>-0.03</td>
<td>0.24</td>
<td>0.18</td>
<td>0.35</td>
</tr>
<tr>
<td>NODS</td>
<td>-0.25</td>
<td>0.15</td>
<td>0.40</td>
<td>0.80</td>
<td>0.18</td>
<td>0.51</td>
</tr>
</tbody>
</table>

The repeated-measures ANOVA main effect of treatment on the outcome variable was statistically significant (F(2, 250) = 9.37, p < 0.001), indicating significant variability in treatment effectiveness. The post-hoc comparisons are contained in Table 18. As with the Maximum Treat matching, CBT was found to be more effective than TLDP. Unlike the previous analysis, SFBT was also found to be more effective than TLDP. Figure 4 illustrates the differences in outcome between groups, while Figure 5 illustrates the pairwise comparisons.
Table 18

*Post-hoc Pairwise Comparisons Following Caliper Matching*

<table>
<thead>
<tr>
<th></th>
<th>Value</th>
<th>Test Statistic</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SFBT-TLDP</td>
<td>6</td>
<td>t(125) = 3.69</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>SFBT-CBT</td>
<td>-1.4</td>
<td>t(125) = -0.74</td>
<td>0.46</td>
</tr>
<tr>
<td>TLDP-CBT</td>
<td>-7.3</td>
<td>t(125) = -3.89</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>

Figure 4

*Boxplots of OQ-45 Change for Each Group After Caliper Matching*
Two sensitivity analyses were run for the caliper matching data, one for each of the significant pairwise comparisons of the outcome (Table 19; Table 20). For each of these sensitivity analyses, the results were found to be robust against hidden biases at Gamma < 1.5. This suggests that these findings are slightly more robust than that found through the Maximum Treat matching.
Table 19

*Sensitivity Analysis for SFBT-TLDP Comparison Following Caliper Matching*

<table>
<thead>
<tr>
<th>Gamma</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>1.1</td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td>1.2</td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td>1.3</td>
<td>0.02</td>
</tr>
<tr>
<td>1.4</td>
<td>0.04</td>
</tr>
<tr>
<td>1.5</td>
<td>0.08</td>
</tr>
</tbody>
</table>

Table 20

*Sensitivity Analysis for CBT-TLDP Comparison Following Caliper Matching*

<table>
<thead>
<tr>
<th>Gamma</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>1.1</td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td>1.2</td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td>1.3</td>
<td>0.02</td>
</tr>
<tr>
<td>1.4</td>
<td>0.04</td>
</tr>
<tr>
<td>1.5</td>
<td>0.07</td>
</tr>
</tbody>
</table>

2:1:n Matching

Following the 2:1:n matching, 44 matched triplets were created. A total of 62, or 61.4\% of the total sample was utilized by this matching procedure. The unmatched participants are summarized by group in Table 21. Only those cases involved in the matching process were included in the following set of analyses.
Table 21

*Unmatched Participants Within 2:1:n Matching*

<table>
<thead>
<tr>
<th></th>
<th>Number unmatched</th>
<th>% unmatched</th>
</tr>
</thead>
<tbody>
<tr>
<td>CBT</td>
<td>22</td>
<td>48.9</td>
</tr>
<tr>
<td>SFBT</td>
<td>9</td>
<td>28.1</td>
</tr>
<tr>
<td>TLDP</td>
<td>8</td>
<td>33.3</td>
</tr>
<tr>
<td>Total</td>
<td>39</td>
<td>38.6</td>
</tr>
</tbody>
</table>

The values for each covariate by treatment group following the 2:1:n matching are summarized in Table 22. Boxplots of each covariate by group, as well as boxplots of the pairwise comparisons following 2:1:n matching are contained in Appendix C. The covariates were each statistically balanced across groups at p<0.05 (Table 23).

Table 22

*Covariate Values Following 2:1:n Matching*

<table>
<thead>
<tr>
<th></th>
<th>CBT</th>
<th>SFBT</th>
<th>TLDP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>44.16</td>
<td>44.73</td>
<td>44.57</td>
</tr>
<tr>
<td>Gender</td>
<td>0.50</td>
<td>0.39</td>
<td>0.50</td>
</tr>
<tr>
<td>Initial OQ-45</td>
<td>66.55</td>
<td>65.70</td>
<td>69.05</td>
</tr>
<tr>
<td>Anxiety</td>
<td>0.61</td>
<td>0.50</td>
<td>0.64</td>
</tr>
<tr>
<td>NODS</td>
<td>7.82</td>
<td>8.36</td>
<td>8.18</td>
</tr>
</tbody>
</table>
Table 23

Inferential Testing of 2:1:n Post-Matching Covariate Balance

<table>
<thead>
<tr>
<th></th>
<th>Test Statistic</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>F(2,86) = 0.06</td>
<td>p = 0.95</td>
</tr>
<tr>
<td>Gender</td>
<td>X^2(2) = 2.94</td>
<td>p = 0.23</td>
</tr>
<tr>
<td>OQ</td>
<td>F(2,86) = 0.32</td>
<td>p = 0.73</td>
</tr>
<tr>
<td>Anxiety</td>
<td>X^2(2) = 2.30</td>
<td>p = 0.32</td>
</tr>
<tr>
<td>NODS</td>
<td>F(2,86) = 1.39</td>
<td>p = 0.25</td>
</tr>
</tbody>
</table>

Further inspection of the balance created by the 2:1:n matching indicates that for three of the covariates (age, gender, and initial OQ-45 score), at least 50% of the bias was reduced through this procedure (Table 24). The bias in age is largely eliminated, with 90% of the overall bias reduced.

Table 24

Pairwise Differences of Covariate Balance Following 2:1:n Matching

<table>
<thead>
<tr>
<th></th>
<th>CBT-SFBT</th>
<th>CBT-TLDP</th>
<th>SFBT-TLDP</th>
<th>Bias</th>
<th>Stand. Bias</th>
<th>Bias reduced</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>-0.57</td>
<td>-0.41</td>
<td>0.16</td>
<td>1.14</td>
<td>0.04</td>
<td>0.90</td>
</tr>
<tr>
<td>Gender</td>
<td>0.11</td>
<td>0.00</td>
<td>-0.11</td>
<td>0.22</td>
<td>0.15</td>
<td>0.50</td>
</tr>
<tr>
<td>Initial OQ-45</td>
<td>0.85</td>
<td>-2.50</td>
<td>-3.35</td>
<td>6.70</td>
<td>0.10</td>
<td>0.56</td>
</tr>
<tr>
<td>Anxiety</td>
<td>0.11</td>
<td>-0.03</td>
<td>-0.14</td>
<td>0.28</td>
<td>0.18</td>
<td>0.24</td>
</tr>
<tr>
<td>NODS</td>
<td>-0.54</td>
<td>-0.36</td>
<td>0.18</td>
<td>1.08</td>
<td>0.21</td>
<td>0.34</td>
</tr>
</tbody>
</table>

The results of the repeated-measures ANOVA on the outcome variable returned nonsignificant results (F(2,86) = 2.81, p = 0.06). Given that this test was nonsignificant at
p < 0.05, no post-hoc dependent sample t-tests were returned. Neither was sensitivity analysis warranted. Boxplots of the dependent variable by treatment group, as well as boxplots of the pairwise comparisons of the dependent variable are contained in Figure 6 and Figure 7, respectively.

Figure 6

*Boxplots of OQ-45 Change for Each Group After 2:1:n Matching*
Figure 7

Boxplots of Pairwise Differences of OQ-45 Change After 2:1:n Matching

3:2:n Matching

The final matching procedure, the 3:2:n matching, resulted in 67 matched triplets. This procedure utilized 69.3% of the total sample. A summary of the unmatched participants by treatment group is contained in Table 25. Only those cases involved in the matching process were included in the following set of analyses.
### Table 25

*Unmatched Participants Within 3:2:n Matching*

<table>
<thead>
<tr>
<th></th>
<th>Number unmatched</th>
<th>% unmatched</th>
</tr>
</thead>
<tbody>
<tr>
<td>CBT</td>
<td>16</td>
<td>35.6</td>
</tr>
<tr>
<td>SFBT</td>
<td>9</td>
<td>28.1</td>
</tr>
<tr>
<td>TLDP</td>
<td>6</td>
<td>25</td>
</tr>
<tr>
<td>Total</td>
<td>31</td>
<td>30.7</td>
</tr>
</tbody>
</table>

Values for each covariate by treatment group following the 3:2:n matching are contained in Table 26. Though no covariate significantly differed between groups at \( p < 0.05 \), NODS score was only marginally equivalent across groups (Table 27). Similar to what was observed with the caliper matching, the between-group difference on the NODS covariate was closer to statistically significant than prior to matching despite having bias reduced by 23% with the 3:2:n matching (Table 28). This is likely due to the power increase associated with the large number of triplets created through this procedure. Boxplots of each covariate by group, as well as boxplots of the pairwise comparisons following 3:2:n matching are contained in Appendix D.
Table 26

*Covariate Values Following 3:2:n Matching*

<table>
<thead>
<tr>
<th></th>
<th>CBT</th>
<th>SFBT</th>
<th>TLDP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>44.23</td>
<td>44.69</td>
<td>44.90</td>
</tr>
<tr>
<td>Gender</td>
<td>0.42</td>
<td>0.40</td>
<td>0.42</td>
</tr>
<tr>
<td>Initial OQ</td>
<td>-45</td>
<td>67.97</td>
<td>66.58</td>
</tr>
<tr>
<td>Anxiety</td>
<td>0.66</td>
<td>0.53</td>
<td>0.58</td>
</tr>
<tr>
<td>NODS</td>
<td>7.94</td>
<td>8.57</td>
<td>8.16</td>
</tr>
</tbody>
</table>

Table 27

*Inferential Testing of 3:2:n Post-Matching Covariate Balance*

<table>
<thead>
<tr>
<th></th>
<th>Test Statistic</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>F(2,122) = 0.11</td>
<td>p = 0.89</td>
</tr>
<tr>
<td>Gender</td>
<td>$X^2(2) = 0.09$</td>
<td>p = 0.96</td>
</tr>
<tr>
<td>OQ</td>
<td>F(2,122) = 0.24</td>
<td>p = 0.79</td>
</tr>
<tr>
<td>Anxiety</td>
<td>$X^2(2) = 2.65$</td>
<td>p = 0.26</td>
</tr>
<tr>
<td>NODS</td>
<td>F(2,122) = 2.74</td>
<td>p = 0.07</td>
</tr>
</tbody>
</table>

Pairwise comparisons and bias reductions for each covariate are contained in Table 28. For two of the covariates, age and gender, bias was greatly reduced, with 88% and 91% respectively. However, the 3:2:n matching also had two covariates where the bias reduced was at or less than 30%.
Table 28

*Pairwise Differences of Covariate Balance Following 3:2:n Matching*

<table>
<thead>
<tr>
<th></th>
<th>CBT-SFBT</th>
<th>CBT-TLDP</th>
<th>SFBT-TLDP</th>
<th>Bias</th>
<th>Stand. Bias</th>
<th>Bias reduced</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>-0.46</td>
<td>-0.67</td>
<td>-0.21</td>
<td>1.34</td>
<td>0.04</td>
<td>0.88</td>
</tr>
<tr>
<td>Gender</td>
<td>0.02</td>
<td>0</td>
<td>-0.02</td>
<td>0.04</td>
<td>0.02</td>
<td>0.91</td>
</tr>
<tr>
<td>Initial OQ-45</td>
<td>1.39</td>
<td>-0.98</td>
<td>-2.37</td>
<td>4.74</td>
<td>0.07</td>
<td>0.69</td>
</tr>
<tr>
<td>Anxiety</td>
<td>0.13</td>
<td>0.08</td>
<td>-0.05</td>
<td>0.26</td>
<td>0.18</td>
<td>0.30</td>
</tr>
<tr>
<td>NODS</td>
<td>-0.63</td>
<td>-0.22</td>
<td>0.41</td>
<td>1.26</td>
<td>0.25</td>
<td>0.23</td>
</tr>
</tbody>
</table>

The results of the repeated-measure ANOVA on the 3:2:n matched groups was not statistically significant (F(2,122) = 2.98, p = 0.055). Given this nonsignificance, post-hoc analyses, as well as sensitivity analyses were not conducted. Boxplots for the outcome variable for each group and boxplots for the pairwise comparisons of the 3:2:n matched groups are contained in Figure 8 and Figure 9, respectively.
Figure 8

Boxplots of OQ-45 Change for Each Group After 3:2:n Matching

Figure 9

Boxplots of Pairwise Differences of OQ-45 Change After 3:2:n Matching
Comparing Matching Procedures

Each matching procedure is compared across four elements: number of matched triplets created, proportion of the overall sample that were involved in the matching, proportion of bias reduced through matching, and the p-value of the overall comparison of treatment effectiveness (Table 29).

In terms of creating matches, the caliper matching procedure far outperformed the other three methods. The caliper matching produced nearly twice as many matches as the next highest, the 3:2:n matching, and over four times as many matches as the lowest, the Maximum Treat matching. The Maximum Treat was the only procedure utilizing fewer data points than the unmatched analysis, which used the total of 101 participants. The Maximum Treat utilized only 28 matched triplets, or 84 total data points.

For proportion of the overall sample utilized in the matching, three of the four procedures, the Maximum Treat, the 2:1:n, and the 3:2:n, all performed relatively equivalently. Each of these three procedures matched between 60 and 70% of the total sample. The caliper matching again outperformed the other methods, matching 79.2% of the total sample.

In order to compare the effectiveness of each matching procedure at reducing covariate imbalance, total proportion of bias reduced was calculated by averaging the percentage of bias reduced for each covariate within each matching procedure (Table 29). Two of the procedures, the caliper and the 3:2:n matching, had virtually identical proportions of total bias reduced through matching, at 60 percent. The 2:1:n matching reduced 51 percent of the pre-matching bias, and the Maximum Treat matching reduced
the least amount of bias, at 42 percent. Balance achieved was also calculated as average standardized bias for each covariate within each matching procedure (Table 29). The caliper and the 3:2:n matching again outperformed the other procedures, though the 3:2:n matching slightly outperformed the caliper matching on this metric.

The caliper matching had the lowest p-value (p < 0.001) among the post-matching repeated-measures ANOVAs of the outcome variable, which may at least in part relate to the larger sample size utilized in that analysis. The Maximum Treat matching also returned a significant repeated-measures ANOVA (p = 0.04) despite utilizing the smallest sample size of all the analyses. The remaining, two procedures, the 2:1:n and 3:2:n, returned p-values that were close to but not statistically significant at p<0.05. Therefore, taken from a strict p-value approach, only two of the procedures lead to the conclusion of differential treatment effect.

Table 29

Comparing Matching Procedures

<table>
<thead>
<tr>
<th></th>
<th>Max Treat</th>
<th>Caliper</th>
<th>2:1:n</th>
<th>3:2:n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Triplets</td>
<td>28</td>
<td>126</td>
<td>44</td>
<td>67</td>
</tr>
<tr>
<td>% Matched</td>
<td>60.4</td>
<td>79.2</td>
<td>61.4</td>
<td>69.3</td>
</tr>
<tr>
<td>Stand. Bias</td>
<td>0.16</td>
<td>0.12</td>
<td>0.14</td>
<td>0.11</td>
</tr>
<tr>
<td>PBR</td>
<td>0.42</td>
<td>0.60</td>
<td>0.51</td>
<td>0.60</td>
</tr>
<tr>
<td>p-value</td>
<td>0.04</td>
<td>&lt; 0.001</td>
<td>0.06</td>
<td>0.055</td>
</tr>
</tbody>
</table>
Chapter Four: Discussion

Overview

This study sought to expand current knowledge on several statistical and clinical issues, on which the current literature base is as of yet insufficient. Propensity score analyses are a relatively new set of procedures designed to counteract selection bias in non-randomized studies (Rosenbaum & Rubin, 1983). These procedures have only recently been adapted to studies involving more than two treatment groups (e.g. Bryer, 2013). One of the aims of this study was provide a practical application of propensity score matching across three treatment groups. Also, there is a relatively small amount of literature regarding the implementation of propensity score analyses with smaller sample sizes. This study utilized a sample that may well represent the number of individuals commonly available for analysis from a treatment setting.

In terms of clinical contributions, this study sought to expand the current understanding related to problem gambling treatment. This was primarily achieved through the implementation of two forms of therapy that are not represented in the current problem gambling literature: Solution-Focused Brief Therapy and Time-Limited Dynamic Psychotherapy. These therapies were compared with Cognitive-Behavioral Therapy to answer three questions: 1) are the problem gambling clients differentially selecting these treatments, 2) what factors contribute to which therapy is selected, 3) do
these three treatments exhibit comparable short-term effects. Each of these three questions was formulated to yield substantive contributions to the field of problem gambling treatment.

Prior to the primary analyses, the missing data were screened for appropriateness for imputation. Given that this study already involved a relatively small sample size for propensity score analysis, data imputation was considered a superior strategy to listwise deletion of cases with missing values given the consequent loss of subjects and power associated with the removal of cases with missing data. Overall, the proportion of missing values for each variable was relatively small, ranging from 0 to 6.9%, which indicated that expectation maximization imputation could be appropriate (Gold & Bentler, 2000). Also prior to imputation, missing data were assessed for whether it was random occurrence. Little’s MCAR test was designed to test whether missing values within a multivariate dataset are occurring completely at random (Little, 1988). The results of this test suggested that data were missing completely at random and, therefore, further demonstrated that the missing data were eligible for imputation. Following imputation, not surprisingly given the relatively small proportion of missing values, post-imputation values did not differ substantially from pre-imputation values (Table 2, pp. 71).

Assessing the sample revealed that the individuals included in this study were relatively representative of treatment-seeking problem gamblers in general. The gender breakdown of this study was very similar to that found by Potenza et al. (2007) in a study of gambling helpline callers. The mean NODS score (8.09) indicates that, overall, this
sample represents individuals with relatively severe gambling problems. The high prevalence of co-occurring disorders in this sample is also indicative of what the literature suggests is typical among this population (e.g., Kessler et al., 2008). The gender differences observed in this study are also similar to those suggested in the literature. The female subjects in this study were found to be significantly older, which likely echoes previous findings that women begin gambling later in life (Ladd & Petry, 2002). The women in this study also presented with more severe gambling problems, which may help confirm the notion that there is a social stigma against women gambling that prevents them from seeking treatment until their problems become more severe (Ledgerwood et al., 2012). Overall, this sample seems to well reflect what is generally expected in a representative problem gambling sample, which helps to bolster the external validity of the conclusions drawn from this study.

**Research Question One**

The first research question focused on what clinical and demographic factors might be related to treatment selection. The reasons for this were twofold: identifying factors related to treatment selection is an important component in propensity score estimation, and identifying which clients might prefer a particular therapy may have implications for matching clients with an appropriate treatment (Brooks & Ohsfeldt, 2013). However, this issue became somewhat complicated for this study, because not all participants were given the opportunity to select their treatment. Of the total sample of 101, 22 participants were randomly assigned to treatment, so it would have been misleading to assess group differences of the overall sample and represent that as
assessing factors related to treatment selection. Consequently, separate analyses were run on only those individuals that selected a treatment so as to determine what types of problem gambling clients might be preferring particular therapies.

The current body of literature provided very little guidance in the formulation of hypotheses regarding treatment selection. In fact, only one study could be located that addressed the issue of treatment preferences among problem gambling therapy clients. Najavits (2011) studied the influence of co-occurring PTSD on treatment preferences and found that preferences did vary between gamblers with and those without PTSD. Interestingly, they found the group with PTSD ranked treatments as more likely to be helpful. However, this trend went across the board, as opposed to the comorbid group preferring only particular treatments more than the PG only group. Though this study did not provide any specific information regarding how the selection of the three treatments in this study might be influenced by a co-occurring disorder, it did provide evidence that treatment preference and selection might be related to comorbid mental health conditions. This study assessed the impact of four common co-occurring mental health disorders, including PTSD, on treatment selection and outcome. In addition this study also sought to explore the role of certain demographic and clinical indicators that have been implicated in problem gambling treatment outcomes.

For each covariate tested, the groups did not differ significantly at p < 0.05. However, some interesting overall trends emerged from these comparisons. The SFBT group had the least severe presentation. This group had the lowest initial OQ-45 scores, lowest proportion screening positively for three of the four co-occurring disorders
(depression, generalized anxiety, and mood disorder), as well as lowest pre-treatment NODS scores. These characteristics seem to logically fit with the strengths-based nature of SFBT. The description of SFBT that is provided to the clients mentions that SFBT ‘[works] to identify your strengths and successes...’ (Table 1, pp. 64). It seems that those with more severe presentations may be less willing to and/or capable of identifying strengths on which to build.

The TLDP group had the highest ratings on the WAI-S, was the youngest group, and had the highest proportion of males. The high ratings on the WAI-S seem to fit with TLDP’s focus on relationships. Clients with higher levels of interpersonal functioning may be more willing to select TLDP and may also form better working alliances with their counselor. The lower age found in the TLDP group may reflect a generation gap in the willingness of younger gamblers to work on the interpersonal aspects of disordered gambling. The gender imbalance found with TLDP came as a surprise. Given the gender stigma often experienced by males regarding psychotherapy, it seemed like men would be less likely to select a therapy that focuses on interpersonal relationships and expressing emotions (Schomerus, Matschinger, & Angermeyer, 2009). Yet nearly four-fifths of the clients that selected TLDP were male.

The CBT group appeared to have the most severe overall presentations. This group had the highest initial OQ-45 scores, highest proportion screening positively for three of the four co-occurring disorders (depression, generalized anxiety, and PTSD), as well as the highest pre-treatment NODS score. The CBT group also had the lowest scores on the URICA and WAI-S. These trends for the CBT group may best be understood by
considering what CBT is not. More severe clients, as discussed, may be less likely to select a strengths-based treatment (SFBT). More severe clients may also be such as a result of poor interpersonal functioning and social support, which may preclude them from preferring TLDP. This latter aspect may be supported by the low WAI-S scores associated with the CBT group.

Though none of the single covariates significantly differed between the treatment selection groups, the evidence suggests that the selection process is not random. In other words, there are meaningful differences between the groups related to overall clinical presentations. This would then lead to the necessity to balance groups across these characteristics associated with the treatment selection process.

**Research Question Two**

The second aspect of the research focused on whether the problem gamblers, as a whole, were differentially selecting from among the three treatment options. The results of the chi-square goodness of fit test returned a marginal p-value \((p = 0.13)\), which warranted a closer investigation of the specific proportions of the treatment selections.

Among those clients that were allowed to select their treatment \((n=79)\), 32 \((40.5\%)\) selected CBT. In contrast, only 18 \((22.8\%)\) selected TLDP, and 29 \((36.7\%)\) selected SFBT. The expected frequency for each group was 26.3.

An overall preference for CBT may fit with the traditional view of problem gambling as a brain disease that has its origins in 12-step approaches to recovery (Tangenberg, 2005). CBT might then be seen as directly addressing the thought dysfunction that this paradigm views as foundational to disordered gambling behavior.
CBT may also seem the most targeted intervention to individuals who are entering treatment control one particular set of behaviors. The opposite may be perceived of TLDP. This sample, which is likely indicative of typical treatment-seeking problem gamblers, presented for treatment with relatively severe gambling problems. Therefore, it is reasonable to assume a level of desperation to control their gambling. In the description of CBT that was provided to the clients there is a reference to addressing ‘problematic behaviors.’ Conversely, the description of TLDP describes the goals of treatment as ‘including improved relationships, attunement to feelings, and/or a resolution of a conflict.’ For these individuals seeking to control their gambling behavior, it may be difficult to perceive the indirect effect that problematic relationships have on the perpetuation of their gambling problems.

CBT may also be appealing to problem gamblers for not seeming personally invasive. The high rates of psychological disorders, including PTSD, among this population coupled with a desire to focus specifically on the gambling behavior may preclude many individuals for desiring a treatment that is perceived to focus on past experiences and underlying emotional content (Kessler et al., 2008). This may further explain the trend away from TLDP and toward selecting CBT.

These overall treatment selection trends may also be better understood in the context of comparing the preferences of problem gambling and non-problem gambling therapy clients. Soberay and Faragher (2011) made just such a comparison using the same three treatment options as presented in this study. Their results suggested that there was a significant difference in the proportions of these treatments selected by problem
gamblers compared to non-problem gambling clients. Specifically, the general therapy clients actually had a slight preference for the TLDP treatment while it was the least selected among the problem gambling clients. This may further strengthen the argument that problem gambling clients are particularly seeking treatment to change their gambling behaviors without exploring underlying relationships, feelings, and conflicts. However, these results suggest that there may be sub-groups within the problem gambling population that may prefer TLDP treatment, such as younger clients. Any comparison of outcomes between these various treatments would need to control for these client differences.

**Research Question Three**

The initial step to the propensity score matching was to ensure that the model estimating the propensity scores was adequately specified. This study chose a two-step approach: identify variables that differ between treatment groups and identify variables that are related to the dependent variable. Cuong (2013) suggested that also including the latter element would lead to more effective estimation of the propensity scores.

To identify variables to include in the propensity score estimation model the entire sample was utilized, including both those that selected their form of therapy and the smaller portion that were assigned to a treatment type. Though these analyses also included individuals randomly assigned to a treatment, the results were very similar to those found when assessing only those clients that selected their treatment. This was not surprising given that the majority of the participants (78.2 %) selected their treatment. The SFBT group was again differentiated by having the least severe overall clinical
presentation, and the TLDP group again tended to be younger and with a larger proportion of males. Though none of the individual variables differed at \( p < 0.05 \), multiple variables differed at the \( a \) priori inclusion criteria of \( p < 0.25 \). These variables were age, gender, initial OQ-45 score, positive screen for generalized anxiety (CD-GAD), and severity of gambling problems (NODS).

Perhaps the most unexpected results of this study were in assessing the relationships of the covariates with the dependent variable, the change in OQ-45 scores through treatment. Only one variable, initial OQ-45 score, was found to be related at the inclusion criteria of \( p < 0.25 \). However, previous studies would have suggested that this relationship would have been in the opposite direction. Dowling (2009) and Jimenez-Murcia (2007) both suggested that the more severe the problem gambler’s presentation, the less they improve through treatment. In this study, though, there was a positive relationship between severity of initial psychosocial distress and improvement through treatment. The existing literature would have also suggested that other variables, including gender, readiness to change, and strength of therapeutic alliance would be associated with treatment outcomes (Crisp et al., 2000; Martin et al., 2000; Petry, 2005). Consequently, the empirically derived propensity score estimation model used in this study excluded several variables that the literature would have suggested to include.

From the described approach, the following variables were selected for the propensity score estimation model: age, gender, initial OQ-45 score, whether the individual screened positive for generalized anxiety (CD-GAD), and NODS score. These variables seemed to provide good coverage for the differences in overall presentation.
observed between groups. OQ-45, the anxiety screen, and the NODS should have adequately captured the variability in severity between groups. Also, age and gender should have adequately captured the demographic differences that were primarily observed with the TLDP group. Therefore, achieving balance on these variables should go a long way in providing overall balance to the three treatment groups.

Analogous to how adding independent variables to an OLS regression equation will always increase the proportion of variance explained, the propensity scores would have been, at least marginally, more accurately estimated had all the available variables been included in the estimation model. However, the issue of sample size precluded this approach. Pirracchio (2012) demonstrated that propensity scores could be effectively estimated with group sizes as small as twenty utilizing four covariates. This ratio of one covariate per five participants was used as a general guideline for this study. Coincidentally, the criteria set for identification of variables to include in the propensity score estimation for this study resulted in a nearly identical ratio of one variable per every five participants in the smallest treatment group. Further consideration of this issue would have been warranted had a substantially larger number of covariates been identified as important for the propensity score estimation through the aforementioned inclusion process.

Following the propensity score estimation, four matching procedures were utilized that are available through the Trimatch statistical package: 1) a form of matching without replacement called Maximum Treat, 2) a with replacement caliper matching, 3) a 2:1:n matching, 4) a 3:2:n matching. The procedures were compared in terms of three
values: number of matched triplets created, proportion of overall sample utilized in the matching procedure, and proportion of bias reduced through matching.

As expected, the caliper matching, since it had no limitations of the reutilization of cases to make matches so long as they fit within the specified caliper, created the largest number of matches, 126. This number was nearly twice as many as any of the other procedures. It also utilized the largest proportion of the overall sample, 79.2%. This latter aspect may be particularly salient to small group propensity score analyses where excluding cases may be more problematic than when dealing with larger samples. Conversely, the Maximum Treat method, since it only reused a participant if in doing so it would create a match with another participant who would have otherwise been unmatched, resulted in the smallest amount of matched triplets, 28.

Two procedures, the caliper matching and the 3:2:n matching, outperformed the others in terms of reducing covariate imbalance. Both of these procedures reduced the pre-matching bias by 60%. However, the profiles of the individual balances achieved through these two procedures differed. The 3:2:n matching had two covariates that each had approximately 90% of the bias reduced through matching. On the other hand, the 3:2:n procedure also had two variables that had less bias reduced than the least balanced variable from the caliper matching. The caliper matching had a more consistent pattern of bias reduction across covariates. It may be important to remember that the proportion of bias reduced statistic that was utilized to compare matching procedures involved averaging the bias reduced for each individual variable. In essence, this metric assumes that balance on each covariate is equally important. Yet, as Cuong (2013) indicated, it
may be most important to balance across covariates that are also related to the outcome variable. Therefore, it may be preferable to have a matching procedure that creates good balance for all covariates, as opposed to a procedure that creates strong balance for some and leaves others less balanced. For this reason, the caliper matching may have more utility than the 3:2:n matching in balancing groups across all covariates. However, the 3:2:n matching did marginally outperform the caliper matching in terms of standardized bias post-matching.

It should also be noted that even though the 3:2:n and caliper matching performed the other procedures in terms of reducing bias, each of these two procedures had a covariate that approached statistical significance for differing between groups following matching. Since no covariates were near statistical significance at p < 0.05 prior to matching and all variables experienced a reduction in bias through matching, this would appear to be largely a power issue. These two procedures, in addition to creating the best balance, created the largest numbers of matched triplets, which contributed to the post-matching analyses having more power to find between-group differences. This would also suggest that had the treatment groups been larger in size, the pre-matching differences for at least some covariates might have been statistically significant.

Prior to any matching procedure, a one-way ANOVA was run to test whether an unadjusted analysis would indicate a significant difference in treatment effect. The results of this test indicated no difference (p = 0.40). In stark contrast, each of the four matching procedures returned a test of overall differential treatment effect that was either significant or marginally significant at p < 0.05. Initially, it may be assumed that the
significant results post-matching would be due primarily to increasing the sample size and statistical power through matching participants several times. However, the Maximum Treat matching created only 28 triplets, which is a total of 84 participants, yet it returned a significant differential overall effect. This is likely primarily due to the Maximum Treat matching resulting in the largest difference in mean OQ-45 improvement between any two groups (a 9.2 unit difference between CBT and TLDP).

Given that this study utilized actual clinical information, not simulated data, it cannot be said which method best represented the true differences in treatment effects. Though, given the purpose of propensity score analysis, these findings do lend support to the practical utility of these approaches. Due to the fact that this study involved participants who self-selected their treatment, it could be assumed that some level of bias was being introduced. The utilization of propensity score matching with this data would, in theory, reduce this bias, thereby resulting in a more accurate estimate of treatment effects. The fact that each matching procedure returned at least a marginally significant result, \( p \leq 0.06 \) for each matching procedure, suggests that there are in fact differences in treatment effects that could only be detected once this selection bias was accounted for.

Specifically, the TLDP treatment may have less utility in short term outcomes, particularly relative to CBT treatment. This finding would seem to fit with the overall purpose of these treatments. The TLDP treatment is designed to address relational difficulties and conflict that likely underlie the gambling problems. It may take an extended period of time before an individual is able to resolve these conflicts and, thereby, fully control their gambling behavior. The CBT and SFBT treatments, though
ultimately addressing underlying issues, begin with more of a focus on incremental and immediate changes. Therefore, TLDP may have more utility in a treatment setting where there is more of an expectation of prolonged treatment engagement, such as a residential treatment facility, though longer-term effectiveness of this treatment for this population has yet to be established.

One of the fundamental issues in propensity score analysis is the inclusion of all relevant covariates in the propensity score estimation (Brooks & Ohsfeldt, 2013). To address this issue, Rosenbaum (2002) devised a sensitivity analysis for use with propensity score analyses that would assess the robustness of the findings to the influence of unobserved covariates. The significant pairwise comparisons for the Maximum Treat and the caliper matching were found to remain significant at Gamma < 1.4 and Gamma < 1.5, respectively. This means that the significant results would remain robust unless an unobserved covariate changed the odds of treatment assignment by a factor of 1.4 or 1.5 respectively. In other words, this study could be considered sensitive to a hidden confounding bias. That, of course, is not to say that the results of this study should be disregarded. Rather, an element of caution should be exercised. The variables tested for inclusion into the propensity score estimation model reflect a broad range of clinical indicators, including several that are commonly associated with the dependent variable from this study. However, given that there is very little knowledge regarding what factors underlie treatment selection, particularly among problem gamblers, it is plausible that an influential covariate may have been omitted.
This study reflects the potential for conducting propensity score matching procedures with relatively small sample sizes. Previous research has suggested propensity score analyses utilizing matching procedures may not be as effective as other uses of propensity scores (e.g., as regression weights) with smaller samples (Holmes & Olsen, 2010). However, that study was attempting to balance large covariate differences and with a relatively restrictive caliper setting (0.05). The results of this study seem to suggest that with small to moderate pre-matching covariate imbalance, matching procedures may not result in a problematic loss of participants and, therefore, may be effective in reducing bias and isolating treatment effects.

One other issue, not explored by this study, was the utilization of bootstrapping procedures. Bai (2013) found evidence that the utilization of bootstrapping could help improve the accuracy of propensity score estimation. However, among the matching procedures implemented in that study, caliper matching was found to benefit the most from the bootstrapping. Given that this procedure might benefit only one of the four matching procedures in this study, it was not applied.

The results of these propensity score matching analyses contain implications for problem gambling treatment, as well as the ways in which the associated data are analyzed. Given the ubiquity of non-randomized studies, as well as studies including more than two treatment groups, multi-group propensity score analyses appear to be a viable option for reducing bias and estimating treatment effects. Specifically, the Trimatch package for R appears capable of balancing potentially confounding covariates across three treatment groups, as well as detecting treatment effects where unadjusted
analyses may fail to do so. Clinically, the results of these analyses suggest that particular treatments, particularly CBT, may be more effective in reducing psychosocial distress for short-term problem gambling interventions.

**Limitations**

In terms of the propensity score analysis, this study had several limitations. Perhaps most notably, propensity score analyses, as previously discussed, were designed to correct for selection biases found in observational studies, particularly as they relate to individuals selecting their own form of treatment. However, in this study not all the participants were given the opportunity to choose their form of therapy. Twenty-two of the participants in this sample were randomly assigned to a form of treatment. This may have limited the ability to detect systematic between-group differences when using the complete sample.

Another major limitation in terms of assessing the utility of this multi-group propensity score matching approach relates to the magnitude of the observed differences between groups. The between-group differences on the observed covariates were all below the threshold for statistical significance at $p < 0.05$, which may indicate one of several possibilities. There may be unobserved variables that explain treatment selection, which would result in a misspecification of the propensity score estimation models. Alternatively, treatment selection may be a complex process that is difficult to explain through the assessment of any single characteristic. The differences in overall clinical presentations, as previously discussed, may support this notion. In either case, the lack of large observed between-group differences may have hindered the ability to demonstrate
the effectiveness of these matching procedures. With relatively small group differences, this study did not have the opportunity to demonstrate that these matching procedures are effective in reducing relatively large covariate imbalances.

For the treatment aspects of the study, there are also multiple limitations. This study contained no formal means of validating treatment fidelity. Specific supervision in each of the treatment types was provided for the counselors, and closed-circuit television was used to observe treatment sessions live. Both of these measures served to increase treatment fidelity; however, there was no systematic review of the interventions utilized. More to this point, the therapists for this study were all graduate students, which introduced an element of variability in clinical experience, especially as it relates to the particular forms of treatment. These limitations impact the ability to attribute the outcomes to a strict view of what each treatment type should consist of.

Other limitations relate to the duration of treatment assessed by this study. Given that treatment selection occurred as a part of the treatment process, the specific treatments were not applied for the entirety of the observed timeframe. Often, the treatment selection occurred during the second session, so each client did not receive all sessions in the form of treatment they selected. This may have limited the ability to see differences between treatments. This time frame was selected as a practical representation of the amount of time that clients are likely to remain in treatment before attrition becomes a major factor. However, other, particularly residential treatment settings may be more interested in outcomes from a longer-term treatment.
Recommendations for future research

Future research would benefit from building upon the statistical and clinical ramifications of this study. Propensity score analyses for multiple treatment groups utilize statistical applications that have only recently been developed (e.g. Bryer, 2013; Ridgeway, McCaffrey, Morral, Burgette, & Griffen, 2014). Therefore, future research would benefit from a more in-depth study of these applications and the circumstances under which they may be more or less effective.

This study also uses a sample that may be of typical size for many treatment settings. As previously mentioned, Holmes and Olsen (2010) found that matching may not be the ideal approach with small sample sizes, multiple treatment groups, and large between-group covariate imbalance. This would then lead to the question of what sets of circumstances would contribute to a given usage of propensity scores being more or less effective when dealing with multiple treatment groups. Therefore, future research may benefit from further investigation of propensity score analyses with smaller sample sizes, particularly as it relates to multi-group propensity score techniques.

The propensity scores for this study were estimated using empirical criteria for covariate inclusion. The result was the exclusion of several variables, such as the strength of the working alliance, which the existing literature has implicated in treatment outcomes. It may be useful to compare the results of this empirically derived model to one that uses a theoretical approach to specifying the propensity score estimation model.

The specification of the propensity score estimation model was also limited by the sample size. Given the sample size requirements of the logistic regression models used to
estimate the propensity scores, the number of covariates selected had to be limited. Future research may want to focus on whether larger samples and the inclusion of more covariates would result in better matching in a similar study. Also, it may be valuable to look at alternative propensity score estimation techniques that don’t have the sample size requirements when using smaller samples, such as the predicted values from a multiple regression.

This study sought to explore factors related to treatment selection among problem gambling therapy clients, the results of which suggested that overall clinical presentation may be related to therapeutic preference. However, this study does not address the clinical utility of allowing problem gambling clients to select a form of treatment and/or trying to match a client with a particular form of therapy. Future research may want to assess whether the outcomes might be different for clients that are assigned to a treatment compared to those whose treatment is based on individual preference. Particularly if there is a utility found in tailoring treatment to the individual, future research may also want to further explore what demographic, clinical, and personological variables may be related to treatment selection.

The clients at this treatment facility were relatively homogenous in terms of racial and ethnic composition, which is commonly observed in problem gambling treatment settings (Volberg, 1994). Future research may want to take a more proactive approach to recruiting minority clients so as to better understand the role of diversity as it relates to the treatment of problem gambling.
This study also provided evidence that forms of psychotherapy, other than CBT, may be effective in the treatment of this population but that not all forms may be equally efficacious, at least not in the short term. Future research may benefit from exploring whether these forms of therapy may be differentially effective over a longer course of treatment and/or differentially effective at periods following the termination of treatment. Given that this study suggests that a broader range of therapies than previously researched may be effective in treating this population, additional forms of therapy may be tested with this population.
References


American Psychiatric Association. (2013). *Diagnostic and statistical manual of*


treatments. Retrieved from http://cran.r-project.org/web/packages/TriMatch/
vignettes/TriMatch.pdf


Gamblers Anonymous (2013). *Questions and answers about gamblers anonymous.*


Golinelli, D., Ridgeway, G., Rhoades, H., Tucker, J., & Wenzel, S. (2012). Bias and
variance trade-offs when combining propensity score weighting and regression: with an application to HIV status and homeless men. *Health Services & Outcomes Research Methodology, 12*(2/3), 104-118


Grall-Bronnec, M., Wainstein, L., Feuillet, F., Bouju, G., Rocher, B., Venisse, J.L., &


problem gambling: Factors associated with first treatment and treatment re-entry.

*Addiction Research & Theory, 16*(6), 618-632.


Jiménez-Murcia, S., Álvarez-Moya, E. M., Stinchfield, R., Fernández-Aranda, F.,


Kessler, R.C., Hwang, I., LaBrie, R., Petukhova, M., Sampson, N.A., Winters, K.C., &


Kuss, O. (2013). The z-difference can be used to measure covariate balance in matched propensity score analyses. *Journal Of Clinical Epidemiology, 66*(11), 1302-1307.


Lambert, M.J., Burlingame, G.M., Umphress, V., Hansen, N.B., Vermeersch, D.A.,


Lunt, M. (2014). Selecting an appropriate caliper can be essential for achieving good


Martin, G., Macdonald, S., & Ishiguro, S. (2013). The role of psychosocial characteristics
in criminal convictions among cocaine and gambling clients in treatment.


decade of self exclusion: Missouri casino self-excluders four to ten years after enrollment. *Journal Of Gambling Studies, 26*(1), 129-144.


Petry, N. M. (2003). Patterns and correlates of Gamblers Anonymous attendance in
pathological gamblers seeking professional treatment. *Addictive Behaviors, 28*(6), 1049.


Rosenbaum, P. R. (1984). From association to causation in observational studies: The


satisfaction- a systematic review. *BMC Psychiatry, 11*, 64-75.


Hong Kong SAR. *Journal of Applied Social Psychology, 37*(12), 2788-2799.


Appendix A

Boxplots of Age for Each Group After Maximum Treat Matching

Boxplots of Pairwise Differences of Age After Maximum Treat Matching
Bar Charts for Gender for Each Group After Maximum Treat Matching

Boxplots of Pairwise Differences of Gender After Maximum Treat Matching
Boxplots of Initial OQ-45 Score for Each Group After Maximum Treat Matching

Boxplots of Pairwise Differences of Initial OQ-45 Score After Maximum Treat Matching
Bar Charts for Positive Anxiety Screens for Each Group After Maximum Treat Matching

Boxplots of Pairwise Differences of Positive Anxiety Screens After Maximum Treat Matching
Boxplots of NODS Score for Each Group After Maximum Treat Matching

Boxplots of Pairwise Differences of NODS Score After Maximum Treat Matching
Appendix B

Boxplots of Age for Each Group After Caliper Matching

Boxplots of Pairwise Differences of Age After Caliper Matching
**Bar Charts for Gender for Each Group After Caliper Matching**

![Bar Charts](image)

**Boxplots of Pairwise Differences of Gender After Caliper Matching**

![Boxplots](image)
Boxplots of Initial OQ-45 Score for Each Group After Caliper Matching

Boxplots of Pairwise Differences of Initial OQ-45 Score After Caliper Matching
Bar Charts for Positive Anxiety Screens for Each Group After Caliper Matching

Boxplots of Pairwise Differences of Positive Anxiety Screens After Caliper Matching
Boxplots of NODS Score for Each Group After Caliper Matching

Boxplots of Pairwise Differences of NODS Score After Caliper Matching
Appendix C

Boxplots of Age for Each Group After 2:1:n Matching

Boxplots of Pairwise Differences of Age After 2:1:n Matching
Bar Charts for Gender for Each Group After 2:1:n Matching

Boxplots of Pairwise Differences of Gender After 2:1:n Matching
Boxplots of Initial OQ-45 Score for Each Group After 2:1:n Matching

Boxplots of Pairwise Differences of Initial OQ-45 Score After 2:1:n Matching
Bar Charts for Positive Anxiety Screens for Each Group After 2:1:n Matching

Boxplots of Pairwise Differences of Positive Anxiety Screens After 2:1:n Matching
Boxplots of NODS Score for Each Group After 2:1:n Matching

Boxplots of Pairwise Differences of NODS Score After 2:1:n Matching
Appendix D

Boxplots of Age for Each Group After 3:2:n Matching

Boxplots of Pairwise Differences of Age After 3:2:n Matching
Bar Charts for Gender for Each Group After 3:2:n Matching

Boxplots of Pairwise Differences of Gender After 3:2:n Matching
Boxplots of Initial OQ-45 Score for Each Group After 3:2:n Matching

Boxplots of Pairwise Differences of Initial OQ-45 Score After 3:2:n Matching
Bar Charts for Positive Anxiety Screens for Each Group After 3:2:n Matching

Boxplots of Pairwise Differences of Positive Anxiety Screens After 3:2:n Matching
Boxplots of NODS Score for Each Group After 3:2:n Matching

Boxplots of Pairwise Differences of NODS Score After 3:2:n Matching