UNDERSTANDING THE UTILIZATION AND EFFECTIVENESS OF COMBINED

(INDIVIDUAL PLUS GROUP) MENTAL HEALTH TREATMENT

IN EVERYDAY PRACTICE

by

Halley J. Brown

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Educational Psychology

The University of Utah

August 2017

The University of Utah Graduate School

STATEMENT OF DISSERTATION APPROVAL

The dissertation of **Halley J. Brown**

has been approved by the following supervisory committee members:

Lois Huebner	, Chair	**May 1, 2017**
		Date Approved
Zac Imel	, Member	**May 1, 2017**
		Date Approved
David Robert Davies	, Member	**May 1, 2017**
		Date Approved
Anne Cook	, Member	**May 1, 2017**
		Date Approved
Brad Lundahl	, Member	**May 1, 2017**
		Date Approved

and by **Anne Cook** , Chair/Dean of

the Department/College/School of **Educational Psychology**

and by David B. Kieda, Dean of The Graduate School.

ABSTRACT

This mixed-methods study is a preliminary response to the shortage of empirical investigation into the combination of nonmanualized individual and group talk therapy in everyday outpatient mental health treatment. The current study addresses this shortage in two ways: (1) by exploring common patterns of co-utilizing individual and group psychotherapy services (i.e., combined treatment) in a naturalistic setting; and (2) by comparing the absolute and differential effectiveness of these service utilization patterns. Archival data included 508 combined treatment episodes collected between 1998 – 2012 at a college counseling center. A discussion group and interrater agreement analysis procedure suggests the presence of four discrete and identifiable combined treatment service utilization pattern categories: *concurrent*, *intermittent*, *segmented*, and *semi-overlapping*. Therapeutic outcomes were measured using the Outcome Questionnaire–45. Pre-post change and predicted recovery curves resemble previously published findings: specifically, that combined treatment demonstrates reliable absolute effectiveness. No particular combined treatment category outperformed the other identified combined treatment categories. Clinical implications and future directions are discussed.

TABLE OF CONTENTS

LIST OF FIGURES

Figures

CHAPTER 1

INTRODUCTION

The demand for empirical support of psychotherapy treatments has been

increasing over the past several decades. A working definition for *psychotherapy*, as

endorsed by the American Psychological Association (APA; 2012), reads as follows:

"Psychotherapy is the informed and intentional application of clinical methods and

interpersonal stances derived from established psychological principles for the purpose of

assisting people to modify their behaviors, cognitions, emotions, and/or other personal

characteristics in directions that the participants deem desirable" (Norcross, 1990, pp.

218-220). Furthermore, the APA (2012) defined *treatments* when used in the context of

health care, as:

> Any process in which a trained healthcare provider offers assistance based upon
> his or her professional expertise to a person who has a problem that is defined as
> related to "health" or "illness." In the case of "mental" or "behavioral" health, the
> conditions for which one may seek "treatment" include problems in living,
> conditions with discrete symptoms that are identified as or as related to illness or
> disease, and problems of interpersonal adjustment. The treatment consists of any
> act or services provided by a bona fide health provider intended to correct, change
> or ameliorate these conditions or problems (Beutler, 1983; Frank, 1973).

Clinicians, researchers, clients, and managed care organizations alike are eager for

evidence regarding which mental health treatment modalities are most successful, in what

context, and for whom.

There is much debate regarding the standards and methods by which treatment

outcomes should be operationalized, reported, and ultimately compared. Broadly, outcomes tend to be considered in terms of efficacy, effectiveness, and/or efficiency. *Efficacy*, defined as the capacity of a given treatment to produce a desired result or effect, is established in highly precise research settings such as randomized controlled trials (RCTs). Treatment delivery in RCTs is typically tightly controlled with ongoing supervision and training of therapists that is not characteristic of most naturalistic settings. For example, clinical trials use patients who have been carefully selected with multiple exclusion criteria so as to isolate the effect of treatments in question on target diagnoses and limit the interference of comorbidities. Efficacy trials tend to be fewer in number, relatively more expensive to execute, and because of their rigid study parameters, often afford limited generalizability. Nevertheless, efficacy studies offer important data pertaining to the maximum level of improvement one can expect from a particular treatment. In contrast, *effectiveness* refers to the impact of treatment when delivered under real-world conditions (e.g., Chambless et al., 1996; Elkin et al., 1989). While the strength of effectiveness research lies in its ability to produce estimates of the expected pre-to-post treatment change in actual clinical practice, such investigations often compromise the opportunity for causal inference or conclusive results. Lastly, *efficiency* refers to a therapy treatment's anticipated rate of effect change across sessions. This metric has risen into favor in parallel with managed care's increasing emphasis on cost containment. That is, a given psychotherapy treatment modality must now be concerned with the 'bang for your buck' it provides.

This study will explore the effectiveness and efficiency of combining individual and group-based mental heath therapies in naturalistic settings. The literature review will

begin with a historical perspective on the use of group-based psychotherapy treatments and the research that has been conducted on it. Next, an overview of integrated psychotherapy approaches, including group-plus-individual (combined) treatment, is provided. In addition, this review will outline the literature regarding mental health service utilization at large, as well as on college campuses. Lastly, the rationale for examining the implementation of combined individual and group psychotherapies in everyday practice is provided and this study's specific research aims are presented.

Review of Literature

Although the written history of therapy groups did not begin until the end of the 19th century (Ruitenbeek, 1969), many philosophers and social scientists have long acknowledged the power of the group. Marx and Engels (1848), for instance, wrote of the incredible power of groups to move human history forward. A consequence of such varied origins and interdisciplinary influences is that agreement over time and across specialties on a single, concise definition of group psychotherapy has been difficult to achieve. A current working definition of *group psychotherapy* can be expressed in the following way: "Group psychotherapy is the treatment of emotional or psychological disorders or problems of adjustment through the medium of a group setting, the focal point being the interpersonal (social), intrapersonal (psychological), or behavioral change of the participating clients or group members" (Burlingame & Baldwin, 2011, p. 505). Today, group psychotherapy is a primary treatment modality in many mental health settings. Yet, for a variety of reasons, our understanding of group treatment in clinical settings is still quite limited. This section will review several of the key contributors and

trends related to the practice and research of group psychotherapy over the past century.

Historical Trends in Group Psychotherapy Treatment

Our understanding of the mechanisms and effects of group therapy has evolved as a result of various scientific and economic factors. Irrespective of the point in history one chooses to consider, however, it has remained important for mental health researchers and clinicians to evaluate the status of group psychotherapy as a sound treatment option. The terms *group therapy*, *group psychotherapy*, and *group treatment* will be used interchangeably herein.

The Turn of the Twentieth Century

There are several figures credited with initiating group psychotherapy. Beginning in 1902, Sigmund Freud hosted a group of well-known medical doctors, psychoanalysts, and laymen, including Alfred Adler, Wilhelm Stekel, Max Kahane, Rudolf Reitler, Max Graf, Carl Jung, and Otto Rank, to name a few, for informal meetings at his apartment in Vienna, Austria (Gay, 1988). Originally referred to as the Wednesday Psychological Society, the meetings typically included paper presentations or case histories with discussion and a final summary by Freud, with some members at times even presenting a detailed account of their own psychosexual histories (Gay, 1988). Freud (1922) would later speculate on group dynamics in a paper titled "Group Psychology and the Analysis of the Ego." Meanwhile, in the United States, an internist at Massachusetts General Hospital in Boston named Joseph Pratt began treating patients with tuberculosis in small group "classes." Tuberculosis was the second leading cause of death in 1900, which

meant that there were scores of patients interfacing with medical professionals for intervention. More closely aligned with what is considered group therapy today, Pratt's "thought control" classes began in 1905 as a cost-effective attempt at getting a large number of patients to commit to the medical regimen deemed crucial to curing their disease (Burlingame & Baldwin, 2011). Pratt noted several ingredients tended to produce more successful "classes," including patients' ability to identify with one another, establish hope for recovery, and develop faith in the class (Burlingame & Baldwin, 2011). Pratt went on to publish his observations in a seminal article in 1907 and gradually transformed his classes into formal therapy groups, ultimately culminating in an article in 1945 on the use of group therapy in treating psychosomatic illness. He gave informational talks to psychiatric patients at the Boston Dispensary throughout the 1920s and 1930s that placed less emphasis on the disease and more emphasis on the emotions that accompanied it as well as their effect on the illness (Pratt, 1945). In many ways, Pratt's work regarding group gave rise to several of the key themes seen in the group therapy literature over the next century, namely group therapy's cost efficiency, unique therapeutic properties, and success with focused disorders (Burlingame & Baldwin, 2011).

There was a noticeable increase with respect to the use of groups for treating psychiatric patients in the subsequent decades. E. W. Lazell (1921) documented his use of group educational interventions to treat World War I veterans with schizophrenia at St. Elizabeth's Hospital in Washington, DC. Beginning in 1919, Lazell called his lectures to the patients on psychoanalytic dynamics "group analysis" (Lazell, 1921). Lazell's work corroborated Pratt's in that he noticed mentally ill patients also gained a sense of support

and hope from one another when provided with the opportunity to share their common experiences. Many exciting developments followed, including work on topics like here-and-now focus (Syz, 1928), group analysis (Burrow, 1927, 1928), milieu therapy (Marsh, 1931, 1933), the re-creation of the primary group (Schilder, 1939; Wender, 1936), and group treatment for children (Slavson, 1943). It was in 1932 that J. L. Moreno first coined the terms "group therapy" and "group psychotherapy" at a conference of the American Psychiatric Association in Philadelphia, after doing basic research on prison populations. Moreno's truly interactional, group-centered approach was in contrast to earlier group methods that were often lecture classes in mental health (Moreno & Whitin, 1932). Rudolph Dreikurs, the protégé of Alfred Adler, is credited with running the first private therapy groups in Vienna in the 1930s. Driekurs (1959) later described group treatment as a powerful forum for a patient's family of origin and as a positive influence for change. The beauty of these early and simultaneous efforts is that they set the stage for future contributions from diverse theories such as psychoanalytic and psychodrama, as well as varied methodologies including empirical to anecdotal.

Post World War II

The modern practice and research of group psychotherapy launched in the 1940s. When an overwhelming number of World War II service members returning from overseas combat were struggling to handle the negative emotional effects of war, mental health care providers responded by delivering therapy to groups of individuals rather than in the traditional one-on-one format. Many clinicians saw this method of therapy as advantageous not only for its ability to serve more than one client at a time so as to

address the growing need for services, but it became clear that group therapy provided a unique environment for emotional and/or behavioral change that was impossible to duplicate in the individual treatment context (e.g., Klapman, 1946; Slavson, 1940). Notably, the American Group Psychotherapy Association (AGPA) was founded in 1942 to provide professional, education, and social support for group psychotherapists both in the United States and abroad. Still, skeptics wanted proof that this innovative format was in fact easing its participants' troubles.

While case studies and anecdotal reports on the therapeutic value of group therapy appeared ever more abundantly in the 1940s, a systematic and objective means of measuring outcomes, determining client progress, and understanding therapeutic factors had not yet been established (Cotton, 1948; Luchins, 1947). Inconsistent findings and confusing conclusions, as a result, left many mental health practitioners doubting the effectiveness of group treatment. Luchins (1947), while acknowledging reports of positive outcomes, urged group therapy researchers to pursue more efficient, reliable, and empirical methods of defining treatment outcomes. He suggested that future investigatory efforts apply a comparative format, utilize control groups when possible, and administer objective measures to determine outcomes so that results could be reliably compared to other modalities, such as individual therapy.

1950s-1970s

Several key contextual factors shaped the practice and research of group psychotherapy over the next 30 years. One trend, which precedes 1950 but greatly gained steam thereafter, pertains to the expansion and differentiation of psychotherapy models.

Not only was the psychoanalytic community bifurcating internally with regard to the emphasis placed on individual versus group dynamics, there was also a steady advent of new theories of psychotherapy, including the gestalt, humanistic, behavioral, interpersonal, and cognitive. Each of these contending approaches incorporated the group format to varying degrees and in varying ways. Another contextual factor was the emergence of group psychology, which is the social-psychological study of small groups. While scholars and researchers in this area primarily studied nonclinical populations, many of the published theories and findings came to influence group-based psychotherapy and intervention methods with clinical populations. Examples include training- or T-groups, sensitivity groups, and encounter groups. Lastly, as was seen in the years immediately following World War II, group therapy was regarded as a rare and viable solution for addressing the stark imbalances in mental health supply and demand. For instance, pressures emerging in the 1970s from publically funded mental health services meant that state hospitals and community mental health centers came to be heavily reliant on the delivery of therapeutic services via the small group format.

Several studies published in the early part of this epoch demonstrated researchers' attempts at addressing Luchins' (1947) concerns, particularly comparing formats and further defining therapeutic factors specific to group psychotherapy (Baehr, 1954; Fairweather et al., 1960; Lieberman, Lakin, & Whitaker, 1968). However, these attempts were not adequate to provide the type of methodologically sound empirical evidence needed to establish the efficacy of group therapy (Luborsky, Singer, & Luborsky, 1975). Comparative studies from the first half of the 1960s, at best, established group therapy as a helpful supplementary treatment (Barlow, Burlingame, & Fuhriman, 2000; Pattison,

1965; Rickard, 1962; Stotsky & Zolik, 1965). Research in the second half of the decade only marginally enhanced our corpus of knowledge, suggesting group therapy as a stand-alone treatment to be capable of producing measurably positive outcomes (Anderson, 1968; Mann, 1966). A modest increase in independent and comparative studies in the early to mid- 1970s painted a somewhat sharper picture. Results from these efficacy studies concluded that group-based treatments could consistently produce outcomes superior to those observed in a control condition, and they began to provide evidence supporting the long-held assumption that group therapy was about as effective as other forms of psychotherapy (Bednar & Kaul, 1978; Emrick, 1975; Lieberman, 1976; Luborsky, Singer, & Luborsky, 1975; Meltzoff & Kornreich, 1970).

That said, research during the 1960s and 70s generally failed to solidify the efficacy of group-based therapies, as it tended still to have problems akin to those of past eras. While the increasing numbers of comparative studies and the production of positive results were promising, they failed to meet the need for evidence in volume and quality. For example, Luborsky et al. (1975) found only 13 studies to include in their review of comparative studies, while Parloff and Dies (1977) reported that studies directly comparing group to other treatments were not numerous enough to make any definite conclusions about efficacy. Specifically, in a review of the outcome research literature from 1966 to 1975, Parloff and Dies (1977) echoed sentiments put forth 30 years prior by Luchins (1947) and Cotton (1948), stating that: (a) many methodological problems were apparent; (b) group therapy was being performed by poorly trained therapists; (c) clear statements about underlying assumptions, postulates, and hypothesizes in published work were scarce; and (d) much was still not known about how client factors, therapist skills,

specific techniques used, and duration of treatment may affect outcome.

1980s-Present

 Assorted changes in the social, economic, and political landscapes of the 1980s revived the determination to distinguish cost-effective methods for treating psychological difficulties (McRoberts, Burlingame, & Hoag, 1998; Pilkonis, Imber, Lewis, & Rubinsky, 1984). Both government-funded health programs (e.g., Medicare and Medicaid) and privately run health maintenance organizations (HMOs) began utilizing managed care techniques and cost-controlling health care practices, leading to the proliferation of more brief and inexpensive approaches to psychotherapy (Budman et al., 1988). Group-based interventions ignited particular interest due to their ability to provide psychological services to multiple clients at once, thus reducing the number of clinicians needed, hours spent, and ultimately the cost per person for treatment. This area (i.e., at the juncture of treatment effectiveness and efficiency) continues to be of interest within the mental health care field today.

 With respect to research, investigators of the 1980s largely moved away from broad treatment comparisons in favor of studies focused on identifying process variables and strengths of group therapy, as well as the necessary and sufficient components for stable client improvement in small group treatments. For instance, Fuhriman, Drescher, and Burlingame (1984) issued a conceptualization of small group processes, and Erickson (1982) contributed a review of small group treatments in inpatient settings. Kaul and Bednar (1986) challenged the field to supply a conceptual model elucidating the essential elements integral to effective group treatment. Fuhriman and Burlingame (1990)

responded to the task with an empirically derived enumeration of the commonalities and distinctions between the individual and group-based treatment formats across the therapeutic dimensions of relationship, interventions, and factors.

The group literature of the 1990s, by and large, centered on articulating conceptual frameworks and practical implementation considerations vis-à-vis specific patient diagnoses, clinical settings, and therapist orientations (e.g., Fettes & Peters, 1992; Hoag & Burlingame, 1997). In a review of 400 studies between 1980-1992, Fuhrman and Burlingame (1994) noticed that 30 distinct client populations were being treated by group therapies. Findings from this same review suggested that cognitive-behavioral therapies were being applied five times more frequently across client populations than other types of groups (e.g., client centered, psychodrama, gestalt). Burlingame, MacKenzie, and Strauss (2004) summarized 107 studies and 14 meta-analyses published from 1990 to 2001 across six disorders (mood, anxiety, eating, substance abuse, personality, and psychotic disorders) and four patient populations (older people, domestic violence, sexual abuse, and medical illness) and found a similar pattern favoring the frequency with which cognitive, behavioral, and cognitive-behavioral group therapies were being studied. The same 5:1 ratio was seen vis-à-vis the application and investigation of cognitive-behavioral group approaches to treat mood, anxiety, and eating disorders. Alternative group models (e.g., process, interpersonal) were studied more frequently with other patient populations (e.g., trauma, substance abuse); however, cognitive-behavioral therapies still maintained a presence with these populations. General conclusions on effectiveness suggested that the magnitude of improvement varied across patient populations (Burlingame, MacKenzie, & Strauss, 2004). For example, cognitive-

behavioral group therapies performed best in treating social phobia, whereas multiple group treatment approaches appeared to produce similar gains for mood and eating disorders.

The amount of evidence amassed by the end of the 1990s, along with key advancements in statistical and methodological understanding, provided the field an unprecedented ability to address several of its lingering controversies. For instance, McRoberts, Burlingame, and Hoag (1998) examined five decades of research in order to test the influence of process effects and confounding moderator variables that had been poorly attended to in the past. The authors noted that the meta-analyses published by Dush, Hirt, and Schroeder (1983) and Nietzel, Russell, Hemmings, and Gretter (1987), which suggested the superiority of individual therapy over group therapy under some circumstances, included studies that operationalized group therapy in nontraditional ways (e.g., delivering treatment packages originally designed for use in individual therapy to more than one person at the same time; Fuhriman & Burlingame, 1994). Additionally, they noted that meta-analytic investigations such as those published by Smith, Glass, and Miller (1980) and Miller and Berman (1983), both of which concluded that the individual and group therapy formats were equally effective, used a between-study rather than within-study comparison methodology. Comparative meta-analyses that calculate differential efficacy using between-study designs are often accompanied by a host of possible confounds (e.g., nonequivalence of client, setting, methodology, and therapist variables), which cannot be controlled for or remedied through the meta-analytic process (Robinson, Berman, & Neimeyer, 1990; Shadish, 1992). Thus, the meta-analysis from McRoberts, Burlingame, and Hoag (1998) was important for its within-study comparison

design and its attention to including only studies that delivered group therapy in its traditional sense. Results indicated no measurable difference in individual and group therapy outcomes, even when accounting for client, therapist, methodology, treatment, and group variables, and subsequent meta-analyses corroborated their findings (e.g., Burlingame, Fuhriman, & Johnson, 2004; Burlingame, Fuhriman, & Mosier, 2003; Burlingame & Krogel, 2005). Moreover, empirical evidence published during the 1990s supported group treatment as comparatively effective to individual treatment for demographically diverse populations (see DeLucia-Waack, Kalodner, & Riva, 2013).

Some were still unconvinced, though. Namely, contemporary researchers pointed out that many of the statistical approaches previously employed to evaluate group-administered therapies did not empirically account for the unique environment in which group clients are receiving treatment—that is, in the presence of other clients. When not properly considered, within-group dependencies (nested effects) such as this can overestimate treatment effects and dramatically increase Type I (false-positive) error rates. To address this, Baldwin, Murray, and Shadish (2005) re-analyzed 33 studies of group-administered treatment from the American Psychological Association's empirically supported treatments list (Task Force on Promotion and Dissemination of Psychological Procedures, 1998). Results indicated that 12.4% to 68.2% of tests originally reported as significant remained significant after corrections, depending on the assumptions made about how much dependency occurs; and, among all tests (not just those that were originally reported as significant), 7.3% to 40.2% remained significant after correction. These findings suggest that a majority of researchers were not accounting for the nested data structure inherent to group-based treatments; hence, any

and all conclusions drawn from previous research that failed to analytically address this were now suspect. Though this study focused on comparisons between treatments and wait-list or placebo control groups, the authors point out that the lack of power is likely to be even more problematic in studies that compare two active treatments (Kazdin & Bass, 1989).

Over the past 20 years, the greater specificity of treatments being applied to distinct patient populations has provided opportunities for investigators to test the effects of pretreatment and in-session process variables on overall improvement. There has been a resurrection of sorts in the study of the unique and defining features of group therapy using programmatic and sophisticated research as well. Topics such as therapeutic factors (e.g., cohesion), group development, and member interaction have found refined measurement and application as treatments and populations become more distinct. Indeed, in a recent review (Burlingame, Fuhriman, & Johnson, 2004), the strength of evidence for 11 group properties was summarized using a four-fold classification scheme (group structure, verbal interaction, therapeutic relationship, and therapeutic factors). Group properties with very good to excellent research support (backed by evidence from two or more randomized clinical trials or meta-analytic studies) include the systematic use of member-to-member interpersonal feedback and the therapeutic alliance; additionally, factors with promising to good empirical support (defined as limited evidence from randomized clinical trials or uncontrolled pre-post treatment improvement) are pregroup preparation (induction), early group structure, leader verbal style, group climate, and the differential effect of therapeutic factors on the basis of the treatment setting.

In sum, while there is still much that can (and should) be done to advance the group therapy literature, significant accomplishments have been made since Pratt's work with tuberculosis patients 100 years ago: (a) the birth and proliferation of effectiveness studies (i.e., examining how specific treatments and treatment formats translate from the laboratory to real-world settings); (b) greater sophistication of process-oriented research; (c) increased use of advanced statistical analyses, complex hypotheses, and innovative study designs; and (d) development of psychometrically supported measurement tools designed specifically for use in group therapy practice.

Historical Trends in Integrated Psychotherapy Treatment

Formal ideas on integrating psychotherapies appeared in the literature as early as the 1930s (Goldfried, Pachankis, & Bell, 2005). In 1932, Thomas French stood before the American Psychiatric Association at their annual meeting and asserted parallels between Freud's psychoanalytic and Pavlov's behavioral approaches (French, 1933). As one might imagine, his contention that operant conditioning was part of the psychoanalytic process was highly controversial, for analysts and behaviorists alike. Sol Rosenzweig's 1936 article also noted commonalities among various systems of psychotherapy. By 1970, about half of APA's Division of Clinical Psychology identified themselves as "eclectic" psychotherapists (Garfield & Kurtz, 1975), and by 5 years later, that figure rose to 64% (Patterson, 1980). This trend persisted in the subsequent decades, as both the number of publications (Arkowitz, 1992) and the proliferation of relevant organizations, textbooks, and journals (Norcross & Goldfried, 2005) have continued to rise. Several explanations have been put forth to explain this movement, including that

psychotherapists slowly began to be confronted with realizations that their treatment techniques were clinically inadequate or functionally incomplete for the sheer variety of patients, contexts, and problems they were encountering in day-to-day practice (Norcross & Halgin, 2005).

A zeitgeist of psychotherapy eclecticism and integration persists today (Norcross & Goldfried, 2005). Psychotherapy *integration* can be characterized as "a general desire to increase therapeutic efficacy, efficiency, and applicability by looking beyond the confines of single theories and the restricted techniques traditionally associated with those theories" (Norcross, 2005, p. 8). Henceforth, *integration* will be used as an all-encompassing term to reflect any and all manners by which psychotherapy theories, interventions, or delivery formats are synthesized. Four general routes to integration have been proposed, including technical eclecticism, theoretical integration, common factors, and assimilative integration (Castonguay et al., 2015, Norcross, 2005). Technical eclectics seek to combine methods, strategies, and techniques from existing theories, with little regard for the creation of or adherence to a new theory or model. It is "actuarial rather than theoretical" (Norcross, 2005, p. 8), often guided by personal experience or pooled data on what has worked best for others in the past with similar problems and characteristics. Multimodal Therapy (MMT; Lazarus, 1989, 1997) and Systematic Treatment Selection (STS; Beutler, 1983; Beutler & Clarkin, 1990) serve as examples of technical eclecticism. While eclectics use procedures drawn from different sources without necessarily subscribing to the theories that spawned them, theoretical integrationists synthesize diverse systems in order to create a theory or therapy that is better than the constituent therapies alone (Norcross, 2005). That is, theoretical

integration emphasizes the synthesis of underlying theories of psychotherapy in addition to the therapy techniques from each so as to produce an entirely new and cohesive approach (Norcross, 2005; Palmer & Woolfe, 1999). One prominent example of theoretical integration is Paul Wachtel's model of Cyclical Psychodynamics, which synthesizes psychodynamic, behavioral, and family systems theories (Wachtel, 1977; Wachtel, 1987; Wachtel, Kruk, & McKinney, 2005). Another example is Anthony Ryle's model of Cognitive Analytic Therapy, integrating ideas from psychoanalytic object-relations theory and cognitive psychotherapy (Ryle, 1990; Ryle, 2005). The most notable common factors model is Prochaska and DiClemente's Transtheoretical Model (Prochaska & DiClemente, 1984; Prochaska & DiClemente, 2005), which emphasizes therapeutic actions that have been demonstrated to be effective, while overlooking specific techniques that have been developed within particular theories. Common factors theory is based in the empirical literature suggesting it is the factors that are common to the most psychotherapies that make any psychotherapy successful (Frank & Frank, 1991; Imel & Wampold, 2008; Miller, Duncan, & Hubble, 2005; Wampold, 2001). Lastly, assimilative integration combines the advantages of a single, coherent theoretical system with the flexibility of a broader range of technical interventions from multiple systems. This route acknowledges that most psychotherapists select a theoretical orientation that serves as their foundation but, with experience, incorporate ideas and strategies from other sources into their practice. Formal models of assimilative integration have been described based on a psychodynamic foundation (Frank, 1999; Stricker & Gold, 2005), cognitive-behavioral therapy (Castonguay, Newman, Borkovec, Holtforth, & Maramba, 2005), and interpersonal and cognitive therapies (Safran, 1998; Safran & Segal, 1990).

To summarize, none of these four strategies on integration are mutually exclusive. Technical eclectics cannot totally disregard theory, and no theoretical integrationist can ignore technique. Assimilative integrationists and technical eclectics both believe that synthesis should occur at the level of practice, as opposed to theory, and even the most ardent proponent of common factors cannot practice without applying specific techniques.

Sophisticated integrative practice is obviously complex, with few clinicians being formally and rigorously trained in specific and efficacious integrated practices. As a result, psychotherapy integration (particularly the eclectic and assimilative styles) has garnered rather negative connotations, largely for its "alleged disorganized and indecisive nature" (Norcross, 2005, p. 15). In response, proponents of integrated methods argue that such criticisms should actually be redirected to *syncretism*—the uncritical and unsystematic combination of theories and techniques (Norcross, 1990; Patterson, 1990). Sometimes thought of as the result of inadequate training or rogue decision making, Eynsenck (1970) characterized this version of psychotherapy synthesis as a "mish-mash of theories, a hugger-mugger of procedures, a gallimaufry of therapies, and a charivaria of activities having no proper rationale, and incapable of being tested or evaluated" (p. 145). Modern naysayers still rather concur, suggesting that any method of integration needs to have some criteria in place for guiding the decision-making process on the part of the therapist or psychologist (McLeod, 2009). Decision making is, therefore, particularly pertinent in integration; however, there is still a paucity of research regarding the clinical decision-making practices and outcomes associated with integrated psychotherapy (Cutts, 2011; Lazarus, 2005; Patterson, 1989; Schottenbauer et al., 2007).

In sum, integrationist scholars, clinicians, and psychotherapy researchers are being encouraged to join together to bridge the practice-research divide via empirical evidence. Castonguay and colleagues (2015) put forth several research directions aimed at strengthening and supporting integrative practice (and, perhaps more importantly, examining any potentially harmful effects, which are heretofore unknown). In short, they called for identify and testing factors that are related to unskillful and/or inappropriate integration of various interventions, relational and technical processes that are toxic within and across integrated orientations, as well as inadequate matching of clients with particular integrated approaches (Castonguay et al., 2015).

Unfortunately, this paucity of research on integrated approaches cannot be attributed to a lack of encouragement or guidance (Castonguay, 2015). The National Institute of Mental Health (NIMH), anticipating that integration would be a major focus of future empirical research and funding, sponsored a Task Force that brought together a large number of influential researchers to delineate recommendations for future research (Wolfe & Goldfried, 1988). Sadly, more than a quarter century later, it does not appear that these recommendations have inspired researchers (nor those who financially support their research) to prioritize the study of integrated mental health treatment. The imperative of evidence-based practice is unavoidable now, though, and does not look to be dissipating anytime soon. As a result, proponents and practitioners of integration may well be up against a wall, so to speak, in order to survive. In an effort to emphasize this imperative even more, the main professional organization at the helm of the integrative movement (the Society for the Exploration of Psychotherapy Integration, SEPI) has recently adopted a new goal: building stronger links between science and practice

(Goldfried, 2013).

Individual Plus Group (Combined) Psychotherapy Treatment

The combination of therapy formats (e.g., individual, couples, family, group) has also come to be considered within the legitimate boundaries of integration (Norcross & Napolitano, 1986). Today, inpatient psychiatric hospitals, residential and intensive day treatment programs, private practitioners, community mental health clinics, and college counseling centers all (to a greater or lesser extent) implement some combination of group and individual treatment; and, whether it be theoretically designed or just pragmatically implemented, the combining of individual and group therapy has come to be associated with its own proposed mechanisms of change, indications and contraindications, and operative approaches (Porter, 1993). Henceforth, the term *combined* is used to refer to the integration of individual and group psychotherapy in mental health treatment.

When used separately, the individual and group formats each provide a different therapeutic environment with its own distinct advantages. Some proponents argue that the individual therapy format serves as a place to address intrapsychic issues while the group format functions to confront behavioral and interpersonal matters (Porter, 1993; Yalom, 2005). Group therapy can be structured in a variety of ways to achieve this function, including general support, interpersonal/psychodynamic-oriented processing, cognitive-behavioral, and psycho-education/skills-based interventions. For such proponents, the administration of both is intended to capture the advantages of each format thus "capitaliz[ing] on the presence…of multiple settings, multiple transferences, multiple

observers, multiple interpreters, and multiple maturational agents" (Yalom, 2005, p. 432). Combining individual and group may also be instituted to help maintain a client in treatment who might otherwise terminate, or to bolster service engagement in a client whose treatment needs are judged to be especially severe or complex (Yalom, 2005).

There are a number of potential complications associated with combined treatment. For example, when a client is receiving individual and group counseling from two different therapists, there is the potential for therapists to either intentionally or unintentionally undermine the other's therapy format, resulting in client confusion, dropout, and potentially poorer outcomes (Lipsius, 1991; Schermer, 2009; Yalom, 2005). Combined therapy with multiple therapists works best when the client provides informed consent for communication between the group and individual therapist, recognizes the importance of working in good faith in both formats, and accepts the responsibility of bringing clinical material appropriately to each setting (Yalom, 2005). Mutual respect and open dialogue between therapists, although time-consuming, facilitates the treatment's effectiveness (Yalom, 2005). Furthermore, in a situation where the same therapist is present in both formats, it is possible for clients to become frustrated during group sessions, where the therapist's attention is now being shared (Yalom, 2005). Regardless of whether the therapists for each format are distinct or the same, these complications suggest that if combined treatment is to be effective, therapists and clients must work toward the creation of a nonredundant and cohesive therapeutic experience (Yalom, 2005). Clarity about the reason(s) for combining formats and agreement about the overall objectives of treatment between the individual therapist, group therapist, and client increases the likelihood of success (Yalom, 2005). That is, simply adding a second

therapy is unlikely to remedy a resistance to the first therapy, and may encourage an avoidance of working on treatment goals in earnest (Yalom, 2005).

While individual therapy is usually considered the main treatment format with group therapy supplementing it, some advocate for a model of care wherein group therapy is primary with individual therapy serving as an adjunct (Golden, Corazzini, & Grady, 1993; Rutan & Alonso, 1982). A framework of treatment delivery using this latter approach has demonstrated therapeutic value and cost-effectiveness in several settings, including in student mental health (Amaranto & Bender, 1990). In their model, Amaranto and Bender employ monthly individual psychotherapy sessions as an adjunct to weekly psychodynamically oriented group therapy, arguing in favor of utilizing individual sessions to help group members clarify and make optimal use of the group sessions, thereby making the group treatment process more focused, better directed, and easier to manage. Supplementing group therapy with individual sessions at the beginning of a course of group-based treatment may also be helpful in reducing early dropout rates (Staats, 2010). Internationally, there exists a small but continuous stream of publications reporting on combined psychotherapy approaches (e.g., Amaranto & Bender, 1990; Cunningham & Matthews, 1982; Lipsius, 1991; Rutan & Alonso, 1982; Schwartz, 2004; Ulman, 2002); however, that stream has slowed substantially in recent years (Staats, 2010). Altogether, the corpus of literature in this area is still relatively nascent and imprecise.

<u>Review of Research Findings</u>

The research on combined treatment has long disseminated conflicting messages. Baehr's (1954) comparative effectiveness study found in favor of the combined treatment approach above and beyond either method in isolation. In contrast, Kadis and Markowitz (1958) declared individual and group treatments to be in opposition to one another; ultimately, their summative review concluded that the two formats do not conform, and in effect, called into question the prospect of a truly combined form of therapy. Such conclusions, however, should be seen in the context of the times (and the authors' psychoanalytic assumptions). That is, psychoanalysts' historical eschewing of group psychotherapy has been attributed to an impression by its adherents that the group format does not lend itself well to the psychoanalytic principles of treatment; particularly, "transference is considered to be so diluted in group psychotherapy that effective transference interpretations are not possible" (Schachter, 1987, p. 456). Sager (1960), while acknowledging in his symposium the clear dearth of literature on the topic, eventually concurred with Baehr's (1954) research and advocated in favor of the broad utility and increasing popularity of combined individual and group approaches. Several subsequent studies from the 1980s (e.g., Bostwick, 1987; Freeman & Munro, 1988) point toward superior outcomes for outpatient clientele receiving combined individual and group treatment when compared to patients who received either individual- or group-only treatment. The Pittsburgh Psychotherapy Project (Pilkonis et al., 1984), which evaluated individual, group, and combined modes of "insight psychotherapy," reported that differences between therapists accounted for a greater proportion of the variance in patient outcomes than did observed differences between treatments.

Today, practice guidelines and published theory (e.g., Lipsius, 1991; Porter, 1993; Schermer, 2009; Yalom, 2005) regarding the simultaneous delivery of individual and group therapy remain in need of an enhanced empirical basis. Much of the research on the combination of individual and group therapy formats has focused on establishing its effectiveness in a given setting (e.g., psychiatric inpatient hospital, residential facility) or demonstrating its efficacy for specific presenting concerns or psychiatric disorders (e.g., substance abuse, psychosis). What's more, a large proportion of the recent research conducted in this area pertains to testing outcomes of *unified treatment models*, wherein one-on-one and group-based interventions are conceptually integrated constituents of a cohesive or manualized therapeutic protocol.

The most prominent manualized combined therapy approach is dialectical behavior therapy (DBT; Linehan, 1993). DBT combines basic behavioral procedures of skills training; exposure-based procedures; cognitive modification; contingency management; and problem solving with validation, mindfulness practices, reciprocity, and a focus on the patient-therapist relationship (Koerner & Linehan, 2000; Linehan, 1993). DBT's protocol explicitly integrates the individual and group formats to address treatment goals in an organized and interrelated way; namely, it espouses a skills training group (2-2.5 hours/week for the usual year of treatment) in tandem with twice-weekly individual therapy sessions and telephone coaching to deal with issues associated with emotion regulation, distress tolerance, and interpersonal behavior. The protocol has been shown to be beneficial in the treatment of borderline personality disorder (e.g., Linehan et al., 1991; Linehan et al., 2006; Turner, 2000), co-occurring substance abuse and borderline personality disorder (e.g., Linehan et al., 1999; Linehan et al., 2002), binge

eating disorder (e.g., Telch, Agras, & Linehan, 2001), bulimia nervosa (e.g., Safer, Telch, & Agras, 2001), and depression in older adults (e.g., Lynch et al., 2003), to name a few. DBT's efficacy is supported by meta-analysis (Kliem, Kroger, & Kosfelder, 2010), four RCTs (Koons et al., 2001; McMain et al., 2009; Turner, 2000; Verheul et al., 2003), two pre-post studies conducted by independent researchers (Ben-Porath, Peterson, & Smee, 2004; McQuillan et al., 2005), as well as numerous internal investigations put forth by its developers. It is important to note that many of the studies conducted on DBT were set in intensive outpatient/day treatment programs or residential/inpatient locations and/or with individuals who meet stringent psycho-diagnostic criteria.

The value of an integrated treatment protocol such as DBT's is that it provides researchers with a rich venue for isolating and investigating the contributions made by each of the various treatment components, rather than merely relying on one overall observed treatment effect. To directly evaluate the relative importance of DBT's elements, Linehan and colleagues (2015) performed a single-blind randomized clinical trial involving 1 year of treatment and 1 year of follow-up with 99 women diagnosed with borderline personality disorder. The study compared Standard DBT (consisting of weekly individual therapy, group skills training, between-session telephone coaching, and a therapist consultation team), DBT Skills Training (DBT-S; which replaced individual therapy with a manualized strengths-based case management intervention), and DBT Individual Therapy (DBT-I; which replaced group skills training with an activity-based support group and prohibited individual therapists from teaching DBT skills). All three versions were shown to significantly reduce suicidality, while those that included group skills training (Standard DBT and DBT-S) were more effective in reducing depression,

anxiety, and the frequency of acts of nonsuicidal self-injury (NSSI) than the one that did not (DBT-I). In addition, compared to DBT-I, Standard DBT resulted in lower treatment dropout rates and fewer emergency department visits and psychiatric hospitalizations after treatment. Results from a previously conducted dismantling study indicated that the DBT skills group component did not incrementally increase efficacy when added to ongoing, non-DBT individual therapy (Koerner & Linehan, 2000), which suggests that the delivery of a theoretically integrated treatment protocol provides several salient advantages over approaches that are either organizationally and/or thematically patched together.

Whereas fully integrated combined psychotherapy is widely considered to be the optimal approach for clients with trauma and personality disorder diagnoses (Karterud, Johansen, & Wilberg, 2007; Stoffers, Vollm, Rucker, Timmer, Huband, & Lieb, 2012), outcomes are mixed with respect to its efficacy to address substance abuse disorders (Weiss, Jaffee, de Menil, & Cogley, 2004). For instance, Crits-Christoph and colleagues' (1999) multisite clinical trial randomly assigned 487 cocaine-dependent patients to one of four manual-guided psychotherapy treatments: individual drug counseling plus group drug counseling (GDC), cognitive therapy plus GDC, supportive-expressive therapy plus GDC, or GDC alone. Individual drug counseling plus GDC produced the greatest improvement in patients' Addiction Severity Index-Drug Use Composite scores and was superior with respect to the number of days of reported cocaine use in the past month. However, in a different study, Weinstein and colleagues (1997) reported no relative differences in drug use or psychological functioning outcomes for cocaine abusers receiving either individual-only or individual-plus-group treatment. Panas, Caspi,

Fournier, and McCarty's (2003) archival study of 7,815 clients treated in 63 publically funded outpatient substance use programs in Massachusetts indirectly supports the superiority of the combined treatment model. Half of the patients in this study received no group treatment (individual-only), 18% received "light" group treatment (less than 2/3 of treatment were group sessions), and 32% received "heavy" group treatment (more than 2/3 of their treatment were group sessions). Patients in heavy group treatment attained better outcomes than the other two conditions; there was no outcome difference between the individual-only and light group treatment conditions.

Outside of DBT and substance abuse treatments, there is anecdotal and empirical support in favor of combining the individual and group therapy formats to treat bulimia nervosa (Halmi, 2005), domestic violence/battery (Babcock, Green, & Robie, 2004), acute psychosis (Drury et al., 1996), pathological gambling (Lesieur & Blume, 1991), and compulsive hoarding (Steketee et al., 2000).

All in all, while the findings from efficacy studies are promising, the empirical investigation of combined therapy's implementation and effectiveness in actual naturalistic contexts is fairly robust for theoretically integrated models but remains severely lacking and/or riddled by mixed findings in terms of less integrated or general practice models (Mickelson, 2008). For instance, in a recent retrospective analysis by Burlingame and colleagues (2015), the differential effectiveness of combined individual plus group treatment in a college counseling center varied depending on the way in which it was defined. A period of service utilization was identified as combined treatment if the proportion of group sessions to total sessions exceed 0% but was less than 100%, and, it included both individual and group treatment in the same course of therapy. When

outcomes were computed based on absolute rates of change (i.e., pre-post symptom reduction), combined treatment posted a higher percentage of "Improved" and lower percentage of "No change" compared to the individual-only and group-only formats. On the other hand, when combined treatment was operationalized as a proportional variable (proportion of group sessions a client attended) and as a categorical variable (individual-only, group-only, and combined), combined treatment produced less change than the individual- and group-only formats. These analyses examined change trajectories as the measure of outcome, which is different from the aforementioned comparison of absolute rates. The authors note that the differential effectiveness observed across treatment formats may be a result of influences including length of treatment, client variables, and treatment-related effects. The combined treatment sample utilized, on average, almost three times as many total services as those in individual-only and about twice as many as those in group-only. Additionally, compared to individual-only, the combined treatment sample was comprised of a higher percentage of personality, eating, and anxiety disorders and a lower percentage of adjustment disorders. The combined treatment sample reported the greatest initial distress at intake and was more likely to indicate having prior mental health treatment compared to individual-only. Results from a separate study found that combined treatment clients reported lower levels of engagement and cohesion with their groups than their group-only counterparts (Davies, Burlingame, Johnson, Gleave, & Barlow, 2008). In sum, the outcomes produced by tightly controlled efficacy studies, which are designed to test specific treatments for specific populations and/or in specific settings, are often in stark contrast to those which are produced when the treatment is delivered in the messiness of everyday clinical work where the unique needs of patients,

therapists, systems, payers, and so forth all must be taken into account to best serve clients in a sustainable way.

Utilization and Effectiveness of Psychotherapy Services

Research on the amount of therapy necessary to achieve positive and lasting therapeutic outcomes has influenced administrative policy and treatment planning decisions. Kadera and colleagues (1996) presented a mathematical model that generated a linear function illustrating client change during therapy. Kadera's work led to procedures now commonly used for determining relationships between a single unit of treatment, the *dose*, and its effect on therapy outcome, the *response*, which can be plotted on a session-by-session basis. Commonly referred to as dose-effect or dose-response modeling, the graphical curves produced by this analysis, called *trajectories*, provide information for understanding the rate by which clients recover in individual therapy (Kadera et al., 1996). The relevance of dose-effect modeling rests in the specified information it yields, which can be used as a common language for clinicians to evaluate treatment progress while also providing feedback for trainees, supervisors, and experienced therapists (Lutz, Martinovich, Howard, & Leon 2002).

In a review of 156 publications on the topic of a dose-response relationship between 1950 and 1992, Orlinsky, Grawe, and Parks (1994) found that 64% of studies showed a positive relationship between treatment length and outcome, 32% were unable to detect a statistically significant relationship, and only 4% of studies reported a negative relationship between treatment length and outcome. Hansen, Lambert, and Forman (2002) also reviewed the dose-response literature and concluded that, on average,

between 13 and 18 sessions of individual counseling are needed for clinically significant

psychiatric symptom alleviation across a variety of diagnoses. More recently, Baldwin

and colleagues (2009) have challenged the dose-effect approach by arguing in favor of a

good-enough level (GEL) model in which rate of therapeutic change varies across clients.

Thus, some individuals will require just a few sessions to achieve clinically significant

improvement while others may need many sessions to achieve the same therapeutic

effect. In the case of group therapy delivered in naturalistic settings, there are

comparatively fewer published studies than those relating to individual therapy. As a

result, less is known about the optimal dosage of group treatment (Burlingame,

Fuhriman, & Mosier, 2003). If the group process is considered broadly, Yalom (2005)

theorizes that generally eight or fewer group sessions are sufficient for many clients to

return to their precrisis levels.

I am not aware of an empirical analysis of dose-response or differential attrition

that focuses on combined treatment in a naturalistic setting. Similarly, I was not able to

find a direct examination of the sequences or frequency with which clients and/or

clinicians elect to utilize particular types of psychotherapy services within unstructured

outpatient settings. A better understanding how various mental health services are utilized

separately and together will undoubtedly yield greater insight into enhancing the overall

efficiency and effectiveness of psychotherapy treatment.

Utilization of Services at University Counseling Centers

According to the Association for University and College Counseling Center

Directors Annual Survey, approximately 10% of students nationwide received counseling

services from their campus counseling center during the 2015-2016 academic year (Reetz, Bershad, LeViness, & Whitlock, 2016). Findings of the National Survey of Counseling Center Directors were consistent with this estimate, specifying that close to 11% of students received individual or group counseling in 2014 (Gallagher, 2015). The mean number of services utilized (i.e., sessions attended) per treatment episode, regardless of service type, for a given client at a university counseling center is approximately five to seven (Minami et al., 2009; Snell et al., 2001; Stone, Vespia, & Kanz, 2000). Moreover, counseling center clients indicating higher severity at intake have been shown, on average, to utilize more services than those indicating lower severity at intake, suggesting that there is a subset of the overall treatment-seeking population whose concerns are both more severe and tend to take longer to treat (Melling, 2014). There do not appear to be any primary studies exploring the specific patterns of service utilization that comprise the combined approach to treatment, nor the comparable effectiveness of various service patterns.

Current Study

This section will put forth the rationale and purpose of the current study and present the study's central research questions and hypotheses.

Rationale

The ever-increasing emphasis on effectiveness, efficiency, and cost of mental health services by clinicians, researchers, and managed care organizations over the last 25 years is well documented. However, actual clinical decision making regarding referral to

the three standard forms of psychotherapeutic treatment delivery (i.e., individual therapy only, group therapy only, or combined individual plus group therapy) still tends to be determined by agency philosophy, agency resources, clinician inclination, or client preference, rather than by empirical evidence. As highlighted by Lazarus, Beutler, and Norcross (1992), "At the very least, a quest for improved therapeutic efficacy argues that therapists require particular organizing principles to guide them in determining under what circumstances a given procedure should be applied or withheld. The mishmash of divergent bits and pieces, and the muddle of idiosyncratic and ineffable clinical creations are the antithesis of effective and efficient psychotherapy" (p. 12). Whereas the comparative effectiveness of individual versus group-based psychotherapy has been sufficiently established via meta-analysis (see DeLucia-Waack, Gerrity, Kalodner, & Riva, 2004; Fuhriman & Burlingame, 1994; McRoberts, Burlingame, & Hoag, 1998), there are strikingly few published empirical analyses of combined psychotherapy treatment in naturalistic settings. As noted previously, the little research that has taken place predominantly focuses on the comparison of treatment outcomes from unified treatment models for clients with severe trauma, personality disorders, and/or substance abuse diagnoses. Research is lacking on the comparison of treatment outcomes for general outpatient therapy clients who engage in the combination of treatment formats.

Purpose

The current study developed as a response to the shortage of investigation into combining individual and group-based talk therapy in everyday practice. Not only has there been little research into the relative effectiveness of mixing nonmanualized

individual and group therapy in naturalistic settings, absent is a comprehensive account of the manner(s) by which the two are joined (quantity, frequency, order, or duration of sessions). Such a gap in the psychotherapy literature is alarming given the longstanding and widespread use of this treatment combination. Furthermore, when one considers that providing both requires more resources than either group or individual alone, clinicians and agencies alike are making decisions regarding resource allocation without sufficient empirical findings to guide or defend such choices. The current study aims to address this gap in the literature in two ways: first, by exploring common patterns of combining individual and group psychotherapy services in a naturalistic setting; and second, by comparing the differential effectiveness of these service utilization patterns. Clinical implications and future directions will also be discussed.

Research Questions and Hypotheses

Based on the two aims stated above, the following research questions will be addressed in this study:

Research Question 1: How does combined mental health treatment look in everyday clinical work? Specifically, in naturalistic settings (i.e., a university counseling center), are there discrete and identifiable patterns of individual and group psychotherapy service utilization; and, if so, what are they?

Hypothesis 1: The counseling center from which these data were collected offers an array of mental health service options, such as individual therapy, group therapy, case management, psychiatry, and crisis intervention. Similar to many other mental health treatment providers, this center does not have strictly enforced policies dictating how or

when particular services, like individual and group therapy sessions, can or should be attended.

It is hypothesized that several discrete and identifiable patterns of individual and group psychotherapy service utilization exist in naturalistic settings. The preceding literature review, in addition to anecdotal knowledge and prior clinical experience, alludes to individual and group psychotherapy services patterns like: a) *Individual-as-primary*, wherein group sessions are irregularly interspersed within a treatment episode consisting primarily of individual sessions (e.g., Panas et al., 2003); b) *Group-as-primary*, wherein individual sessions are irregularly interspersed within a treatment episode consisting primarily of group sessions (e.g., Amaranto & Bender, 1990; Staats, 2010); c) *Engaged*, wherein individual and group sessions are similar in number and attended contemporaneously within the treatment episode (e.g., Linehan, 1993; Schwartz, 2004); and d) *Stacked*, wherein individual and group sessions are similar in number but not attended contemporaneously within the treatment episode. While it is assumed that therapists are highly thoughtful about the recommendations and treatment plans they make with each of the clients they see, it is not assumed that these various patterns of combined treatment utilization were premeditated (either by the therapist/s or the client); rather, that they all unfolded over time in relatively unintended fashion.

Research Question 2: Is the use of combined psychotherapy in naturalistic settings effective; and, if so, are specific service utilization patterns more likely to be associated with superior outcomes than others?

Hypothesis 2: Vast empirical research indicates that psychotherapy treatment provided at a college counseling center is very effective (e.g., Minami et al., 2009; Vonk

& Thyer, 1999). In addition, comparative findings show equivalency in a wide range of outcomes from the individual and group psychotherapy formats in everyday practice within naturalistic settings (Burlingame, Strauss, & Joyce, 2013). Therefore, it is hypothesized that combined psychotherapy will demonstrate overall effectiveness when being evaluated using an established symptom-based clinical outcome measure. A typology of combined treatment utilization patterns has not yet been investigated or formulated. This means that it is difficult to predict which utilization type(s) are more likely to be associated with differentially superior effectiveness. In a study of individual therapy treatment at a college counseling center, Minami and colleagues (2009) found an inverse relationship between session frequency (i.e., the total number of sessions divided by the number of days in treatment) and treatment outcome, after controlling for initial client severity and the length of treatment (i.e., number of sessions attended; $r = -.039$, $p = .046$). Though statistically significant, the magnitude of this effect was less than one-sixth of a percent and thus is unlikely to have any practical relevance; however, this result may hint that a type like *Engaged*, where clients might be attending both an individual and a group session in the same week for multiple weeks, will be associated with less effective outcomes than a type such as *Stacked*, where there is little to no overlap in formats and the session frequency is likely to be lower. It would be predicted, then, that the *Individual-as-primary* and *Group-as-primary* types would be associated with outcomes that rank somewhere in between. Similar results were published by Burlingame and his colleagues (2015), who found that college counseling center clients who received either more group (greater than 80% of sessions for a given client in a given episode were in group format) or more individual (less than 40% of sessions for a given client in a

given episode were in group format) improved at a faster rate than those who received a

more equal mix of individual and group sessions in a given episode of treatment.

CHAPTER 2

METHOD

Setting

Clinical data from the University of Utah Counseling Center's (UCC) archival

database were used for this study. These data were originally collected between January

6, 1998 and January 31, 2012. The UCC operates as an outpatient agency providing

direct clinical services (e.g., individual, group, and couples counseling, psychological

assessment, crisis intervention, and psychiatry) as well as an assortment of nonclinical

services to the broader campus community (e.g., training, workshops, outreach

presentations, and consultation). In addition, the UCC is deeply committed to its role as a

training site for graduate students in clinical and counseling psychology, interns in social

work, predoctoral interns in psychology, and residents in psychiatry, meaning that a

significant number of the services being provided are delivered by supervised trainees.

Overall service utilization at the UCC has seen a consistent rise over the past

decade. The total number of unique clients seen (2003–2004 = ~900 clients; 2008–2009 =

~1,100 clients; 2013–2014 = ~1,300 clients) and the total number of sessions attended

(2003–2004 = ~5,700 sessions; 2008–2009 = ~6,400 sessions; 2013–2014 = ~8,000

sessions) has increased at a steady pace during this time frame (Melling, 2014; University

of Utah, 2014). Between 1998–2012 at the UCC, individual counseling services totaled

34,797 sessions (57% of all clinical appointments), while group counseling services equaled 7,415 sessions (12% of all clinical appointments). Previous explorations of clinical data from the University of Utah counseling center specify that the average treatment episode (defined by a 90-day break in attendance) is approximately seven sessions ($SD = 8$, Mdn = 4, Md = 1, range = 1-254; see Melling, 2014; Minami et al., 2009). The average number of treatment episodes per client in the 1998–2012 data set is 1.38 ($SD = .81$; Mdn = 1; range = 1-12).

Participation in UCC services is entirely voluntary. In most cases, the first meeting a client has is an intake interview. A common exception to this is when a person who is not already a client at the UCC comes in for crisis intervention services, in which case a formal intake interview may not occur until the second visit. A standard initial intake consists of 30 minutes of paperwork in addition to a 30–50-minute interview. During the time frame in which data for this study were collected (1998–2012), students were typically charged $10 per individual session and $5 per group session. If a student is unable to afford the fees, a fee reduction is arranged. All intake and crisis sessions are provided free of charge. Faculty and staff are charged on a sliding scale based on their self-reported income. Individual sessions generally last 50–60 minutes each and may occur weekly, biweekly, or as needed, while the group sessions are typically held for 90 minutes every week. The UCC has historically had a very active group psychotherapy program, including interpersonal process groups, support groups, as well as psycho-educational or skill-based groups. Between July 1, 2010 and June 30, 2011, the last full fiscal year included in this study's archival dataset, clients attended 1,292 group sessions (a 30% increase over the penultimate fiscal year), in 11 groups. One or two facilitators,

with at least one facilitator being a licensed clinician, are assigned to lead each group session. In principle, individual therapy is limited to 12 sessions per year although this is flexibly enforced. There is no session limit in place for participation in group therapy. Treatment planning, therapist assignment, and referral to various available therapy modalities are determined by the clinician(s) overseeing the client's care.

The UCC has utilized a team model since the fall of 1994. Under the team model, each clinical staff member is assigned to one of four teams. Initial assignment to a specific group or individual therapist (or both) is typically made at the Clinical Team Meeting (a client disposition meeting held at the beginning or the end of the day Monday – Thursday). Protocol dictates that clients referred to group meet with the group leader(s) for a pregroup screening/orientation appointment prior to their first group session; these meeting are typically 30 minutes in length. Due to the UCC's limited resources, clients are referred out to mental health services in the community if they are in need of inpatient psychiatric attention, are struggling with a severe personality or eating disorder, and/or have treatment needs that will likely cannot be met with available resources.

Participants

Discussion Group

To begin, I conducted a discussion with a small group of permanent clinical staff (PCS) at the University of Utah Counseling Center. The discussion took place on April 22, 2016. Only the seven licensed PCS who were working at the UCC between 1998– 2012 were invited, as policies, directives, and norms regarding internal referral practices and client involvement in combined treatment have been modified at the UCC since

2012. Four of the five PCS who participated in the discussion group self-identified as White/Caucasian, and four identified themselves as female. Four PCS participants indicated holding a license as a psychologist and one as a clinical social worker; the average length of licensure was 19.8 years ($SD = 10.23$; range = 7-31).

Coders

Current graduate students from mental health fields (i.e., Counseling Psychology, Clinical Psychology, Social Work) at the University of Utah were recruited to complete a card sort task. Use of graduate students in mental health fields as coders was intended to capitalize on their familiarity with psychotherapy. Eleven of the 15 coders who completed the card sort task self-identified as White/Caucasian, and eight self-identified as female. The sample included 11 Counseling Psychology doctoral students, one Clinical Psychology doctoral student, and three Clinical Mental Health master's students; the average length of graduate training in a mental health field was 4 years ($SD = 2.14$; range = 1-7).

Clients

For this study, I analyzed data from the archives of the University of Utah's Counseling Center. The data set provided to me by the UCC contained data collected from more than 6,400 unique clients who attended approximately 61,000 appointments and close to 8,600 treatment episodes at the UCC between January, 1998 and January, 2012. While the majority of UCC clients are University of Utah students, clinical services are also offered to faculty and staff employed for at least .75 FTE on campus. Clients

who participate in therapy at the UCC provide consent at intake to make de-identified information collected by the UCC available for research purposes. The demographics (e.g., age, race/ethnicity, gender, religious affiliation, and marital status) of the UCC's client population roughly approximate the demographics typical of the university's general student population. Across the 14-year period between 1998–2012, 58% of clients (who reported gender information; $n = 5,710$) self-identified as female/woman, and less than 1% identified as transgender or other self-identified gender. During the same time period, amongst clients who reported racial/ethnic information ($n = 5,865$), approximately 75% of clients identified as White or Caucasian, with 5% as Hispanic/Latino/a, 5% identifying as Asian or Asian American, 2% as African American or Black, and 3% as multiracial. The average age of clients in the data set is 25 (range = 17–65; $n = 2,670$; age information is missing for all clients seen prior to 2008). Clients present for services at the UCC via referral from any of a variety of sources, including friends and family members, staff at affiliated university offices, academic faculty, athletics coaches, residence hall supervisors, and/or by off-campus organizations or professionals; the majority of clients, however, are self-referred.

While many UCC clients meet criteria for DSM-5 (American Psychiatric Association, 2013) disorders, the UCC also provides counseling and services to university-affiliated persons who do not meet criteria for a formal diagnosis at present but nonetheless are experiencing psychological or environmental distress. It is not standard practice at the UCC to make clinical diagnoses a part of clients' counseling experience, thus an accurate diagnostic impression of the UCC's clientele is difficult to discern. Clients are, however, asked to self-report at intake the reasons for which they are seeking

the center's services. Between July 1, 2010 and June 30, 2011, the last full academic year included in this study's data set, the most frequently reported presenting problems were anxiety (65%), depression (61%), stress (56%), academics (45%), self esteem (39%), loneliness (34%), relationship with partner (31%), social anxiety (25%), relationship with friends (24%), and family I grew up in (21%). Anxiety and depression have consistently been the top two most reported concerns by clients at this center at intake.

For the time period from which this study's data set were gathered (1998–2012), all UCC clients were asked to complete the Outcome Questionnaire-45 (OQ-45; Lambert, Lunnen, Umphress, Hansen, & Burlingame, 1994) at intake and prior to each subsequent counseling session. Among those that were available, the average OQ-45 score at first visit for all treatment episodes was 74 (SD = 24.21; Mdn = 74; range = 0-160). Among treatment episodes with more than one visit, the average OQ-45 score at last visit was 62.64 (SD = 23.82; Mdn = 63; range = 0-156). A score of 63 or more indicates a symptom and functioning level of clinical significance (Beckstead et al., 2003). The Reliable Change Index (RCI; Jacobson & Traux, 1991) is computed as a ratio of a difference score (e.g., posttreatment minus pretreatment) to the standard error of measurement (calculated based on the reliability of the measure), thus providing a metric to evaluate whether the observed change over the course of treatment reliably exceeds measurement error. Among treatment episodes with known scores at first visit and last visit, the average OQ-45 change was -10.81 points (SD = 18.94; Mdn = -8; range = -96– 68). A change of 14 points or more indicates statistically significant (reliable) change on the OQ-45 (Lunnen & Ogles, 1998).

The following six terms were adopted to designate the clinical and/or statistical

significance of treatment change (see Beckstead et al., 2003; Lunnen & Ogles, 1998). A decrease in OQ-45 score from first visit to last visit of 14 points or more will be referred to as *reliable improvement*. An increase in OQ-45 score from first visit to last visit of 14 points or more will be referred to as *reliable deterioration*. Any change score that does not increase nor decrease by at least 14 points will be referred to as *no reliable change*. Change scores that achieve reliable improvement and move from the clinical range to the nonclinical range will be characterized as a *treatment success*. Those that display reliable deterioration and move from the nonclinical range to the clinical range will be characterized as a *treatment failure*. Regardless of its magnitude, any change score that does not move from the nonclinical range to the clinical range, or vice-versa, will be referred to as *no functional change*.

Among treatment episodes in the 1998–2012 data set with more than one visit and with known scores at first visit and last visit, 38% demonstrated reliable improvement and 17% can be considered a treatment success, whereas 7% showed reliable deterioration and 2% represent a treatment failure. Episodes with no reliable change tallied 55% while 72% experienced no functional change.

For 2 years (between January 2011 – May 2013), clients completed both the OQ-45 as well as the Counseling Center Assessment of Psychological Symptoms (CCAPS; Center for Collegiate Mental Health [CCMH], 2012) at intake and before all individual counseling sessions. The CCAPS-62, administered at intake, consists of 8 scales and 62 items, with overall distress measured via the 19-item Distress Index. The CCAPS Distress Index was developed and added in the fall of 2012. In 2014, when data indicated that the OQ-45 Total score and CCAPS Distress Index were highly correlated ($r = 0.967$;

Duszak, 2014), the UCC elected to collect symptom data solely via the CCAPS, due to their participation in other CCMH activities. According to the 2013–2014 UCC Annual Report, University of Utah UCC clients had an average Distress Index of 1.83 at intake, which falls at the 57th percentile of individuals seeking services at university or college counseling centers. This indicates that at intake, on average, University of Utah UCC clients identified themselves as more distressed than 57% of a large national sample of students seeking services at counseling centers. Also during the 2013–2014 year, clients at the University of Utah UCC reported the highest distress on the Academic Distress subscale (2.04), followed by Social Anxiety (1.95), Depression (1.81), Generalized Anxiety (1.74), Family Distress (1.46), Hostility (1.06), Eating Concerns (1.01), and lastly, Substance Use (.56). University of Utah UCC clients showed the most deviation from other counseling center clients nationwide in the areas of Depression (65th percentile), Family Distress (64th percentile), and Eating Concerns (62nd percentile). I am unaware of empirical comparisons of counseling centers nationwide using OQ-45 data.

About 60% of clients seen during 1998–2012 reported that they have utilized mental health services previously, and 46% indicated that they have taken psychotropic medications. Eleven percent of clients reported prior hospitalization for mental health conditions (range = 1 to "more than 5" prior hospitalizations), while 22% reported having engaged in self-harming behavior at least once, 32% indicated considering suicide at least once, 10% reported at least one prior suicide attempt, and 7.5% reported that they have considered seriously harming someone else at least once in their life. In addition, 23% of clients specified that they had been subject to an unwanted sexual experience at least once, 49% indicated that they had been harassed or abused in their lifetime, and 50%

reported "PTSD" experiences (it is unclear what types of experiences clients are choosing to categorize in this way).

Therapists

Therapy services at the UCC are provided by trainees (i.e., graduate students, interns, postdocs, psychiatry residents) and full-time licensed clinicians (i.e., psychologists, professional counselors, clinical social workers). Trainees receive at least 1 hour of supervision per week from a licensed practitioner in their same field. Supervisors review and sign their supervisees' case notes, review video recordings of therapy sessions, and support the development of trainee's case conceptualization and intervention skills. Supervisees and supervisors are free to utilize a variety of theoretical orientations (e.g., cognitive-behavioral, humanistic, existential, psychodynamic, interpersonal, feminist, multicultural, mindfulness-based, or integrated/eclectic). Group therapy supervision is provided for 30 minutes after each group session by the licensed coleader. Group supervisors also review and sign group notes written by trainees. Group therapy facilitators may implement an array of approaches (e.g., didactic, psycho-educational, and process-orientated).

Approximately 245 unique therapists were identified between 1998–2012, the date range from which data for this study was collected. All therapists are made aware at hire that clinical data are collected by the UCC and will be made available for research purposes.

Measures

Client Information

The University of Utah Counseling Center is a member of the Center for Collegiate Mental Health (CCMH), a collaborative practice-research network comprised of over 300 college counseling centers, key partners, and a number of academic departments at Penn State University. CCMH has created several types of Standardized Data materials to create a common language regarding client, counselor, center, and institutional demographic information within the field of collegiate mental health. The use of the CCMH database allows an international network of college and university counseling centers to analyze their local data as well as to compare their data to national and international norms. All UCC clients complete the Standardized Data Set (SDS) prior to their intake interview (see Appendix A). The SDS is a set of questions comprising socio-demographic information, in addition to items concerning personal, family, and treatment history. During the study period, SDS demographic questions included date of birth, ethnicity, relationship status, academic status, employment status, living situation, military history/veteran status, religious affiliation, sexual orientation, medical history, disability status(es), and referral source(s). The UCC began using the SDS in July 2008. As a result, clients whose data were collected before this seam will be missing many demographic variables, including age. The wording of some of the SDS items was amended in July 2012 (e.g., "What is your gender identity?" has response options including "Woman," "Man," "Transgender," or "Self-identify (please specify)," whereas comparable response options were previously "Female," "Male," "Transgender," and "Prefer not to answer"; CCMH, 2012).

Information regarding prior life events and mental health treatment experiences are also collected. These items instruct the individual to designate "never," "prior to college," "after starting college," or "both" in response to if/when each of the following 11 experiences took place in their life: prior counseling, medication, and/or hospitalization for mental health concerns, prior drug or alcohol treatment, prior intentional self-injury or injury to another person, prior attempted suicide, suicidal ideation, homicidal ideation, unwanted sexual contact or experiences, or experiences of harassment or abusive behavior from another person.

In addition, UCC clients are prompted at intake to complete the Family Experiences Questionnaire. The Family Experiences Questionnaire was constructed by the Research Consortium of Counseling and Psychological Services in Higher Education expressly for the Counseling Concerns Survey and is designed to assess the occurrence of traumatic family history events that may have influenced students' psychological development (Kearney & Baron, 2003). Individuals are asked to mark "yes," "no," or "unsure" to a list of 18 family experiences including divorce of parents, frequent relocation by the family, frequent arguing amongst parents or family members, death or suicide of a parent or family member, family history of gambling and alcohol/drug abuse, history of physical/sexual abuse in the family, family member hospitalized or otherwise treated for mental health concerns, family member prosecuted for criminal activity, and family member with a debilitating illness, injury, handicap, or eating problem.

Lastly, clients are asked to indicate the reason(s) for which they are seeking counseling via the Presenting Concerns checklist. While these data are collected and stored by the center, they were not included in the data set provided to me by the UCC

for this study.

Service Utilization

Details regarding all clinical encounters are recorded and archived via Titanium Schedule software (Titanium Software, Inc., in partnership with the Center for the Study of College Student Mental Health). Titanium Schedule is used by over 1,100 college counseling centers, disability services, and teaching clinics nationwide (Titanium Software, Inc., 2017). Each clinical encounter record includes the type of service provided (e.g., intake interview, individual therapy, group therapy, couples therapy, crisis intervention, psychiatric evaluation, medication management), when the encounter took place, and the clinician(s) who provided the service. To uphold client confidentiality and consistent with University of Utah IRB policies, the data set provided to me for this study did not include the date of service for each service encounter. Instead, I received sequential time information in the form of count data, meaning that the first ever encounter for a given client (e.g., intake) is recorded as time = 0 and the timing of every subsequent encounter is expressed as the number of days between the given encounter and the first encounter.

Outcomes of Treatment

Treatment outcomes were measured via the Outcome Questionnaire-45 (OQ-45; Lambert, Lunnen, Umphress, Hansen, & Burlingame, 1994) total score (see Appendix B). Collection of OQ-45 data was a routine part of clinical service delivery during the investigation period, with clients completing the OQ-45 prior to intake as well as every

subsequent counseling session. The OQ-45 is a commonly used self-report measure of psychiatric distress (Vermeersch et al., 2004). As its name suggests, the OQ-45 consists of 45 items in a 5-point Likert scale (0 = Never; 1 = Rarely; 2 = Sometimes; 3 = Frequently; 4 = Almost Always). Scoring is summative and all items are weighted equally during scoring, including the nine positively worded items that are reverse-scored. Higher total and subscale scores on the OQ-45 indicate greater distress, thus as clients improve the scores decrease (Vermeersch, Lambert, & Burlingame, 2000). Total scores range 0 to 180, with scores at or above 64 being considered clinically significant distress. Additionally, each of the 45 items is assigned to one of three subscales: Symptom Distress, Interpersonal Relations, and Social Role Performance. Umphress and colleagues (1997) measured high concurrent validity of the OQ-45 for college counseling center samples using the Inventory of Interpersonal Problems ($r = .66$; Horowitz, Rosenberg, Baer, Ureno, & Villasenor, 1988), the Social Adjustment Scale ($r = .79$; Weissman & Bothwell, 1976), and the SCL-90-R ($r = .78$; Derogatis, 1992). Lambert and colleagues (2004) measured the OQ's three-week test-retest reliability for nontreatment samples at $r = .84$ and internal consistency reliability (Cronbach's alpha) at 0.93. Findings from Lunnen and Ogles (1998) and Beckstead and colleagues (2003) support the validity of the OQ-45's reliable change and clinical significance cutoff scores defined above.

Procedure

Data Collection

Research Question 1 aims to discover whether discrete and identifiable patterns of combined psychotherapy service utilization exist in an outpatient mental health setting. To answer this question, I utilized an informal discussion group with licensed PCS at the UCC followed by a card-sorting task completed by graduate students in a mental health field. Approval for the data collection methods outlined here was granted by the University of Utah's IRB. Participation was emphasized to participants as entirely voluntary and no more than two emails were sent to any one prospective participant. No individually identifying information was collected from participants.

I emailed directly all eligible licensed PCS at the UCC to announce the study and to solicit participation in the 60-minute small group discussion. As compensation, I offered a modest lunch purchased with personal funds. The purpose of the initial discussion group was to capture qualitative explanations and anecdotal experiences vis-à-vis the reasons and ways that combined treatment was utilized at the UCC between 1998–2012.

Similarly, recruitment for the card-sorting task was carried out via email and word of mouth. I emailed graduate students in the counseling psychology, clinical psychology, and social work departments at the University of Utah to announce the study and solicit participation. No compensation (e.g., monetary reward, course credit) was provided to participants in exchange for their participation in the study.

The de-identified clinical data set provided to me for this study was drawn from a database maintained by the University of Utah Counseling Center. Utilization

information (e.g., type and timing of each session attended) from this data set was used to help answer Research Question 1. Approval for access to data between 1998–2012 was granted by the University of Utah's IRB. The UCC's data collection procedures have been described in previous publications (see Minami et al., 2009). All information collected at intake and subsequent appointments is securely stored in the Titanium Scheduling software database.

Research Question 2 aims to determine the overall effectiveness of utilizing individual and group psychotherapy services within a treatment episode in a naturalistic mental health setting, as well as the comparative effectiveness and moderating influences of particular combined psychotherapy service utilization patterns. Participants (i.e., clients and therapists) were not recruited or contacted, as this portion of the study involves retrospective chart/data review of a data set compiled between 1998–2012.

Before each intake session, clients at the UCC are instructed to fill out the SDS, the Presenting Concerns checklist, and the OQ-45. Clients are directed to complete the OQ-45 at the beginning of every subsequent UCC encounter as well, except for psychiatric evaluations and medication management appointments. The results of the OQ-45 are stored in the OQ Analyst program for both clinical and program evaluation purposes. For a period of time, clients were also given the CCAPS-62 at intake and the CCAPS-32 at all subsequent UCC visits (excluding psychiatric evaluations and medication management appointments). CCAPS data will not be included in the analyses for this study. The OQ-45 takes approximately 5 minutes to complete and results are available to the therapist prior to meeting with the client. While the completion of the OQ-45 is highly encouraged by UCC staff, it is not a requirement for the utilization of

mental health services at the UCC. Clients who regularly attended both individual and group were permitted to complete only one OQ-45 per week (typically, at the individual counseling session). That is, for a variety of reasons, not every client completed the OQ-45 at each session nor did all clients attend sessions on a consistent basis. However, more than 95% of therapy visits in the full 14-year data set included OQ-45 data.

Design and Analysis

The initial level of data analysis was descriptive. The sample was examined for missing data, normality, and outliers. Additionally, basic frequency counts, estimates of central tendency (mean, median, and mode), and dispersion statistics (range, standard deviation) were calculated for each relevant variable. If necessary, skewed data were transformed prior to the testing the two main hypotheses.

Data Reduction

Whereas a clinical trial protocol administered for treatment research usually has a predetermined duration of treatment, the definition of what constitutes the beginning or ending of an episode of treatment in everyday practice is not always as straightforward and easily discernable. Consistent with definitions arrived at for prior research in this counseling center (see Minami et al., 2009; Wong, 2006), this study defines an *episode* of treatment as a continuous series of services with no more than 90 days between any two attended services. Though somewhat arbitrary, 90 days reflects the summer break that occurs between academic calendar years. If the client returns after a 90-day gap, the ensuing service encounter(s) are considered to be a new episode of treatment.

Based on the guideline by Speer and Newman (1996) that 90% of the items on a measure of psychotherapy outcome should be completed in order for the measurement to be considered valid, any OQ-45 administration containing more than five missing item responses was deemed invalid. In such instances, the service utilization details (e.g., session type, therapist, session number) are retained but the symptom information (i.e., OQ-45 score) is considered missing. Thus, the overall data set includes missing data. Missing data is acceptable for this study given the advantages of the analyses chosen for this study, which are outlined below. Only treatment episodes with two or more recorded OQ-45 scores were considered for the longitudinal, or repeated-measures, analysis in Research Question 2. The OQ-45 score at the first attended service (ever, or after a 90-day break) was regarded as the Initial Score for a given episode of treatment. The last attended service before a 90-day break—and its accompanying OQ-45 measurement, if available—was considered the Last Score for that episode.

To be considered as part of the combined treatment sample, an episode must include at least one individual session and at least one session of group therapy. Any episode containing couples counseling was excluded from the sample. There were 557 treatment episodes in the data that met these criteria. Episode length ranged 2–211 sessions ($M = 26.1$; $SD = 21.4$; Mdn = 20). No more than one combined treatment episode per client was permitted for inclusion in the study. In instances where a client participated in two or three combined treatment episodes, only their earliest episode of combined treatment was selected for inclusion in the study. This reduced the sample to 526 episodes of care. In order to curtail potentially spurious effects due to outliers, and for practical reasons pertaining to the creation and decipherability of the cards, episodes

of combined treatment entered in the study were limited to those having fewer than 70 sessions (95%ile = 63.2 sessions). There were 508 unique treatment episodes in the data that met all of the aforementioned criteria.

<u>Hypothesis 1</u>

Hypothesis 1 predicts the existence of several discrete and identifiable patterns of individual and group psychotherapy service utilization. This hypothesis was examined via a multistep card sorting design and analysis adopted from the field of information architecture (Hannah, 2005; Righi et al., 2013). Information architecture is the practice of effectively organizing, structuring, and labeling information content into a structure that enables efficient navigation (Righi et al., 2013).

Initially, card-sorting activities primarily functioned as a means to elicit data *about* study participants, rather than *from* study participants. For instance, the speed at which a subject could successfully sort a set of cards served useful as a metric of mental processes and reaction time (Jastrow, 1886; Jastrow, 1898), memory function (Bergstrom, 1893; Bergstrom, 1894), and imagination (Dearborn, 1898). Due to the great success of early experiments such as these, psychologists took to developing formalized card sorting tasks particularly for use in clinical research and neuropsychological assessment. In fact, card sorting was so well received in psychology that an article appeared in *Science* in 1914 espousing the virtues of various types of card-based activities (Kline & Kellogg, 1914). The most famous card sorting task, the Wisconsin Card Sorting Test, was developed in 1948; it remains a standard test for neurological damage in patients who have suffered head injuries (Berg, 1948; Eling et al., 2008). Card

sorting also made its way into other fields, including criminology (Galton, 1891), market research (Dubois, 1949), semantics (Miller, 1969), and as a standard qualitative tool in the social sciences (Bernard & Ryan, 2009; Weller & Romney, 1988).

In addition to its utility as a clinical research and assessment tool, card sorting has demonstrated value within the social sciences as a participant-based knowledge elicitation technique for understanding associations and distinctions between various items of data (Coxon, 1999; Deaton, 2002; Hudson, 2013). Card sorting has been used to estimate similarity and nonsimilarity between categories, as well as similarity and non-similarity between concepts or items within categories (Halgren & Cooke, 1993; Spencer, 2009). According to Miller (1969), asking participants to sort concepts or items according to some similarity of meaning—or according to meaningful criteria—results in the identification of concept groupings. The procedure offers insight into the meaning that a participant (or a group of participants) assigns to individual concepts or items, as well as insight into the perceived hierarchical organization of the larger concept grouping structure. It is for this reason that the employment of knowledge elicitation techniques such as card sorting are common during the early phases of a project, as they represent an effective initial strategy for enhancing a researcher's insight into a previously uninvestigated concept structure (Spencer, 2009). Findings from card sort studies can be used to guide subsequent and more empirically rigorous investigation. In sum, card sorts serve as an effective (and relatively easy and cost-effective) research method for the preliminary exploration of phenomenological themes and conceptual groupings, and are appropriate for addressing this study's research questions.

In a standard card sorting procedure, participants are asked to organize individual,

unsorted items into groups and may, depending on the technique, also provide labels for these groups. There are two primary methods for performing card sorts. Open sorting is an approach whereby the researcher provides participants with cards, or items of data, with no pre-established groupings, and participants are asked to determine their own groupings by first sorting the cards into groups that they feel are appropriate and then labeling the resulting piles. Closed sorting, on the other hand, is a method whereby the groupings are predetermined by the researcher and the participants are asked to sort the cards, or items of data, into the researcher-defined categories or groupings. This study utilized a version of card sorting known as the hybrid approach, as it mixes elements of both the open and the closed card sorting methods. The hybrid approach employed in this study is discussed further and in greater detail below.

Following an extensive search, it became obvious that the topic of combined (individual plus group) therapy service utilization in naturalistic settings has been only minimally addressed, and no models were found to guide decision making for a card sorting activity. As such, a blend (or hybrid) of the sorting techniques described above was implemented, in line with widely-accepted standards for cluster analytic approaches (Romesburg, 2004): first, determinations are made of whether and how the cluster may be developed; then, the characteristics of each cluster are defined/interpreted and given a label; and lastly, and the solution is validated.

In Step 1, I began by creating a visual representation—or, card—of the service utilization pattern for each of the 508 episodes of combined treatment. I then conducted an informal visual inspection of the service utilization patterns. This activity intended to offer the researcher opportunity for visual exploration, preliminary observations of trends

in the data, and tentative identification of common patterns; it also served to inform the subsequent investigative activities.

In order to determine appropriate categories into which the cards could be sorted, I facilitated a discussion (Step 2) with a small group of licensed PCS currently affiliated with the UCC, who also were working at the UCC between 1998–2012. The objectives of the discussion were to understand more about the various influences and considerations associated with the use (or nonuse) of combined treatment at the UCC between 1998–2012, as well as to understand more about the various forms of combined treatment in everyday clinical practice. The cards created in Step 1 were not made available to the PCS who participated in Step 2 (described below). Thus, the small group discussion functioned as a means to elicit data *from* PCS, rather than *about* them.

The discussion in Step 2 followed a three-stage "Think-Pair-Share" format. To start, each of the five participants were asked to brainstorm, or think, on their own for 10 minutes. Each participant was provided with a piece of paper containing the following instructions:

> Please take a moment to brainstorm as many responses to the following prompts as you can. Feel free to use another sheet of paper as well in order to record all of your thoughts. We are looking to come up with as many responses to each of the following two prompts as possible.
>
> 1. What is the purpose/function of combined treatment, in your mind? What were you trying to accomplish by referring to both modalities rather than only to one or the other? What influences and considerations were involved in your decision-making process regarding referral to combined treatment (as an individual therapist, group leader, clinical team member, clinical supervisor, administrator, etc.)? *Essentially, how & when did you look to implement combined treatment, or suggest that clients utilize it?*

2.   What did you observe in terms of service utilization within the combined treatment format? What patterns of individual and group session attendance took place? *Essentially, if we were to plot common or typical ways clients engaged in individual and group sessions over time, what would those plots look like?*

Next, participants were asked to get together in teams of 2–3 people. Each team was provided with two large poster boards, one for each of the discussion prompts above, and markers. Teams were asked to spend 10–15 minutes collaboratively brainstorming, compiling, and recording their responses on the poster boards provided. Lastly, for the remaining 35 minutes, I brought the boards to the front of the room for review and observation by everyone in the room and facilitated a discussion amongst all of the participants in attendance. Cumulative insights were recorded as the discussion ensued as well as at the end of the discussion, and individual brainstorm responses as well as the large poster boards were retained by me for future consideration and review. Category labels used in the card sort were selected at the end of Step 2, based on results from both Steps 1 and 2.

A card sort was carried out as Step 3. The card sorting approach utilized here is considered a hybrid because participants (i.e., coders) were given the option of withholding cards from placement in any of the predetermined categories, thus creating an "Other" pile/category. Coders who utilized the "Other" category were asked to again sort—and subsequently develop a descriptive category label for—all of the cards in this category as a second-order open sorting task at the end. There is literature to suggest that roughly 10–30 participants (or 5–10 groups of 2–3 participants) are adequate to achieve a reasonable saturation point, or trustworthiness, in the card sorting results (Mauer & Warfel, 2004). Due to the staggered nature of data collection, data are incrementally

acquired and thus, in theory, can be incrementally reviewed and analyzed. Glaser (1978) characterizes *saturation* as the point at which the researcher finds that continued data collection efforts are not contributing anything substantively new to the analysis in progress. There is not a quantitative method for estimating saturation.

Based on the aforementioned data reduction criteria, the overall number of combined treatment service utilization patterns (i.e., episodes, or cards) included in the card sort task was 508. Mauer and Warfel (2004) contend that they have performed successful card sorts with over 200 cards in instances where the participants understood the content well. Similarly, Dearholt et al. (1986) reported no difficulties with 219 cards. All 15 coders were able to complete the card sort task in 60–90 minutes. Coders independently made their ratings and were not permitted to view the ratings of other coders. I entered each coder's individual coding results into a separate spreadsheet at the conclusion of each coder's card sort. Additionally, I reshuffled the cards twice prior to each new coder to ensure random ordering in card presentation. The full instructions read aloud to coders as part of the card sort task are presented in Appendix C.

Each card, measuring approximately 4.25 x 11 inches, depicts a unique combined treatment episode. The horizontal axis along the bottom of each card denotes time (i.e., session number), while the vertical axis along the left side of the card represents session type (i.e., individual or group). Each session of individual or group therapy attended by the client is shown as a red dot. A black line connects the dots for ease of interpretation. To enhance efficiency and decipherability, any session attended by the client that was neither an individual nor group therapy session was removed from the card and was not shown as a red dot. Therefore, although the black line continues uninterrupted, there are

instances of extended gaps between the red dots (see Appendix D).

Analysis of the coding structure was adopted from Hudson (2013) and Righi et al. (2013). First, raw frequency counts of observed ratings were compiled into a card x category matrix, with the cards listed in numerical order in the rows and the categories of utilization types listed across in columns. A basic percent agreement matrix was also constructed, with each cell of the matrix reflecting the percentage of coders who sorted each card into a particular category. Variance in observed ratings were then inspected at the individual coder and the aggregate sample levels.

Next, assessment of interrater agreement (IRA; also called interrater reliability) was conducted. Basic percentages of agreement are not sufficient for assessing IRA, as they do not correct for agreements that would be expected by chance and therefore overestimate the level of agreement (Hallgren, 2012). That is, drawing from classical test theory (Lord, 1959; Novick, 1966), IRA analysis aims to determine how much of the variance in the observed score is due to true score variance after the variance due to measurement error between coders has been removed (Novick, 1966). Cohen's (1960) kappa and related kappa variants are commonly used for assessing IRA for a nominal (i.e., categorical) coding scheme. Kappa statistics measure the observed level of agreement between coders for a set of nominal ratings and corrects for agreement that would be expected by chance, providing a standardized index of IRA that can be generalized across studies (Cohen, 1960). The degree of observed agreement is determined by cross-tabulating ratings for two coders, and the agreement expected by chance is determined by the marginal frequencies of each coder's ratings (Cohen, 1960). Possible values for kappa statistics range from −1 to 1, with 1 indicating perfect

agreement, 0 indicating completely random agreement, and −1 indicating "perfect" disagreement. That is, a "good" value of kappa means that the strength of agreement between coders on category ratings is high, and therefore, that those category ratings are "real."

Several guidelines for interpreting "good" kappa values have been proposed. Landis and Koch (1977) label kappa values from 0.0 to 0.20 as slight agreement, 0.21 to 0.40 as fair agreement, 0.41 to 0.60 as moderate agreement, 0.61 to 0.80 as substantial agreement, and 0.81 to 1.0 as almost perfect or perfect agreement. Krippendorff (1980) advises a more conservative interpretation, suggesting that kappa values less than 0.67 should be disregarded, values between 0.67 and 0.80 should be considered tentative, and values above 0.80 permit definite conclusions.

There are two well-documented effects, or problems, that can substantially cause Cohen's kappa to misrepresent the true agreement among raters (Di Eugenio & Glass, 2004; Gwet, 2002). *Prevalence* effects emerge as a result of an unequal distribution of observed ratings across categories. This problem exists within a set of ratings due to the nature of the coding system used in a study, the tendency for a given sample of coders to identify one or more categories of codes more often than others, or due to truly unequal frequencies occurring within the population under study (Hallgren, 2012). Prevalence problems typically cause kappa estimates to be unrepresentatively low. *Bias* effects emerge when the distribution of specific ratings substantially differ across coders. This problem exists within a set of ratings when, for example, one coder favors a particular category at a much higher rate than other coders in the sample. Bias problems typically cause kappa estimates to be unrepresentatively high. Two kappa variants have been

shown to accommodate these problematic effects: Byrt, Bishop, and Carlin's (1993)

formula for kappa corrects for prevalence while Siegel and Castellan's (1988) kappa

obtains accurate kappa estimates in the presence of bias (Di Eugenio & Glass, 2004). Put

differently, Cohen's (1960) kappa and Siegel and Castellan's (1988) kappa estimates are

unrepresentatively low when prevalence effects are present, whereas Cohen's (1960) and

Byrt, Bishop, and Carlin's (1993) kappa estimates are inflated by bias and therefore not

preferred when bias is present.

   All IRA analyses were conducted with the *irr* library (Gamer, Lemon, Fellows, &

Singh, 2012) in the R programming language (version 0.84; R Development Core Team,

2012). The *irr* library was explicitly developed to estimate various coefficients of

interrater reliability and agreement like the ones used in this study. Prevalence problems

were evaluated via basic descriptive statistics and visual inspection of raw frequency

count distributions (e.g., histograms). Bias problems were evaluated via *irr*'s 'rater.bias'

command, which calculates a reliability coefficient for two raters classifying n objects

into any number of categories (Gamer et al., 2012). When the researcher inputs either a c

x c classification matrix of counts of objects falling into c categories, or a 2 x n or n x 2

matrix of classification scores, the 'rater.bias' function returns the absolute value of the

triangular off-diagonal sum ratio of the c x c classification table and the corresponding

test statistic (Gamer et al., 2012). A systematic bias between two raters can be assumed

when the ratio substantially deviates from 0.5 while yielding a significant Chi-squared

statistic (Gamer et al., 2012). For fully-crossed designs with three or more coders, such as

this study, Light (1971) suggests computing kappa for all coder pairs then using the

arithmetic mean of these estimates to provide an overall index of agreement for each item

of data (i.e., card).

Hypothesis 2

Research question 2 aims to determine the overall effectiveness of utilizing individual and group psychotherapy services within a treatment episode in a naturalistic mental health setting, as well as the comparative effectiveness of particular individual plus group psychotherapy service utilization patterns. It was hypothesized that (a) combined (individual plus group) psychotherapy will demonstrate overall effectiveness, defined as a pre-post decrease in OQ-45 score greater than chance, and (b) that one or more specific service utilization configurations will be more highly associated with superior outcomes than others, defined as a greater average rate and magnitude of change in OQ-45 score. These hypotheses were examined using a multilevel modeling (MLM; also called hierarchical linear models, mixed-effects models, or random coefficient models) procedure designed to predict and compare longitudinal recovery curves for clients in each utilization category. All models were estimated with the *lme4* library (Bates, Maechler, Bolker, & Walker, 2016) in the R programming language (version 1.1-12; R Development Core Team, 2012). The *lme4* library was explicitly developed to handle complex data structures like the data used in this study.

Multilevel models are an ideal analysis method for this study for several reasons. First, when analyzing data that are nested (e.g., repeated measures of symptoms within a client), it is important to account for the occurrence of correlated effects (Hox, 2010). For example, and in the case of this study's design, the observed outcomes (i.e., OQ-45 scores) are likely to be more highly correlated within each client than across all clients.

MLM accommodates the nested nature of the data (i.e., repeated measures within persons). Second, MLM accommodates unequal sample sizes between groups and highly unbalanced timing of assessments that are characteristic of naturalistic studies. Third, and most importantly, MLM is flexible with respect to missing data (Raudenbush & Bryk, 2002; Singer & Willett, 2003). To maximize power, this analysis will include data from all recorded time points (OQ-45 at every session). In sum, because the data and hypotheses are multilevel in nature, the data analytic methods must be as well.

The procedure models, and then tests, predicted growth curves (OQ-45 scores over time) for each service utilization category. The standard multilevel growth curve model involves two levels: within-subject (Level 1) and between-subjects (Level 2). In treatment research, the primary interest will be differences in the overall intercepts and slopes (i.e., fixed effects) due to treatment, or utilization category. At Level 1, the outcome varies within participants over time as a function of a person-specific growth curve. The procedure estimates an intercept (i.e., session 1) and slope (e.g., rate of change across sessions) for each person. At Level 2, the person-specific parameters are viewed as varying randomly across persons, as a function of the person's treatment utilization category. This analysis allows for estimation of the extent to which the intercepts and slopes vary across persons (i.e., random effects). If an equation including the interaction of predictors at Level 1 (i.e., session number) and Level 2 (i.e., treatment utilization category) produce a better model fit than either the Level 1 or Level 2 equations on their own, we can conclude that our predicted outcome (i.e., fixed effects) depends on treatment utilization category differences over time. Chi-square tests are used to statistically compare models against one another. Sessions, rather than days or weeks,

served as the basic time variable due to its precedents in the dose-response and treatment outcomes literature (e.g., Baldwin et al., 2009; Hansen, Lambert, & Forman, 2002; Minami et al., 2009).

Lastly, this study investigated the effect of covariates including episode length, initial severity, and the ratio of individual to group sessions within episode, as initial severity at intake and the number of sessions of treatment received have been shown to predict outcome (Burlingame et al., 2015). It is detrimental to the integrity of the classification scheme if extraneous factors predict the derived typology structure, and/or account for part (or potentially all) the observed variance in symptom reduction outcomes across utilization categories (Wampold & Imel, 2015).

CHAPTER 3

RESULTS

<u>Discussion Group</u>

I conducted a small discussion group with 5 licensed clinicians at the UCC who

were employed there between 1998−2012. Participants in the focus group worked as

individuals and in groups to address two questions. The information I gleaned from this

conversation was used in crafting category labels for the subsequent card sort (see

Appendix E for full discussion group responses).

Question 1

Participants first identified the circumstances under which they opted to refer

clients to both individual and group therapy services. If a clinician chose to refer a client

to combined treatment during the intake or subsequent team meeting, that decision was

rarely made in a deliberate or calculated attempt to reap the rewards of both individual

and group therapy. Rather, clinicians typically only referred clients to combined

treatment at the outset if the clinician believed the client would benefit from group

therapy, but the client expressed reluctance to engage in group therapy only.

More frequently, clinicians referred clients to a second modality after further

interaction with the client following the initial intake. Participants identified several

circumstances under which they were likely to refer clients to a second modality: if the client was already in individual therapy and would likely benefit from the addition of group therapy; if the clinician was attempting to wean the client away from individual therapy and chose to supplement treatment with group therapy; if the client was no longer eligible for individual therapy but desired continued treatment in the form of group therapy; if the client began treatment in group therapy but manifested a desire to switch to individual therapy; or if the client was engaging in group therapy, but the clinician observed that group therapy was not a sufficient modality to manage the client's needs.

In addition to identifying the circumstances under which clinicians chose to refer clients to combined treatment, participants also identified circumstances under which they would not refer clients to both modalities. To this effect, clinicians cited finite counseling center resources as a motivator to refer clients to one treatment type only, and, in particular, to refer low-severity clients to group therapy only. Clinicians who led groups as well as conducted individual therapy cited problems related to seeing particular clients in both settings. Conversely, some clinicians observed a lack of communication and coordination between individual therapists and group leaders, noting that it could be difficult to refer clients to group therapy without knowing which groups were available or open to new members. Finally, clinicians observed a "dilution effect" in clients engaged in both individual and group therapy, wherein clients would not participate fully in group, in an apparent effort to save their thoughts for individual therapy.

Question 2

Participants identified three naturally occurring patterns of service utilization for clients engaged in both individual and group therapy: (1) group therapy as adjunct to individual therapy, (2) individual therapy as adjunct to group therapy, and (3) regular engagement in both individual and group therapy. Clinicians observed that clients who were actively engaged in individual therapy, but had been referred to group therapy as a second modality, frequently demonstrated a lack of enthusiasm about entering group therapy, and attended the group only one or two times. For clients engaged primarily in group therapy, group leaders would offer individual therapy on a limited, as-needed basis for group members. Clinicians described regular engagement in both as weekly individual and weekly group, and as weekly group with an individual session once every 2-3 weeks. Clinicians noted various chronological orderings of service usage.

By and large, participants delineated varieties of combined treatment in terms of how regularly the individual and group therapy services were each being utilized, and/or by the sequential ordering of individual and group attendance. I determined that attempting to create highly specific categories that would comprehensively capture all of the potential sequential ordering (e.g., mixed throughout, mixed then group, mixed then individual, individual then group, individual then mixed, group then individual, group then mixed, individual then mixed then group, individual then mixed then individual, individual then group then individual, etc.) and dose variations (e.g., high individual with low group, high individual and high group, low individual and low group, low individual and high group, etc.) would result in too many categories to make the card sort task feasible to carry out and in line with extant methodological recommendations.

I concluded that the following categories sufficiently spoke to the information gleaned from both the literature review and the discussion group:

- CONCURRENT: client is engaging in both types of services concurrently (meaning, in relatively equal amounts and close together in time), for a sustained period of time; there are no long periods of utilizing only one type

- SEMI-OVERLAPPING: client has one or more period(s) of engaging in both service types at the same time, as well as one or more period(s) of engaging in only one of the service types (either individual or group)

- SEGMENTED: client is engaging in individual and group in distinct segments, or strings, which do not overlap in time; there are no long periods of utilizing both services types at the same time

- INTERMITTENT: client is regularly engaging in one service type and infrequently/irregularly engaging in the other type (meaning, mostly group sessions with interspersed individual sessions, or mostly individual sessions with interspersed group sessions)

- OTHER/DOES NOT FIT: for treatment episodes (cards) that are perceived not to belong to, or not appropriate for, any of the four specified categories above

Card Sort

Amongst the 508 clients included in the study, 56% identified themselves as women. The mean age was 24.34 years (SD = 5.91; Mdn = 23; range = 18–53). The client sample was 83% White, 5% Hispanic/Latino/a, 4% Asian American/Asian, 2% Multi-racial, 1% African-American/Black, 1% Native American, 3% Other, and 1% Prefer not

to answer.

Dictated by study design, Light's (1971) average kappa was applied in order to estimate agreement between the 15 coders for each of the 508 cards. In R's *irr* package, Light's (1971) average kappa is computed from Siegel and Castellan's bias-adjusted variant of kappa (Gamer et al., 2010). Visual inspection and statistical evaluation of the distribution of ratings suggest the existence of a prevalence problem, and the absence of a bias problem (see Appendix F). The use of Siegel and Castellan's (1988) bias-adjusted variant of kappa (which causes estimates to be unrepresentatively low when prevalence effects are present) as the foundation of Light's average kappa in *irr* was deemed suitable given this study's aims. Furthermore, the unrepresentatively low resulting kappa estimates were offset in this study by the use of Landis and Koch's (1977) less conservative kappa interpretation guidelines.

Kappa was computed for each coder pair, and then averaged to provide a single index of interrater agreement (Light, 1971). For a given card to be deemed representative of a particular service utilization category, the card must have garnered "substantial" or "almost perfect" agreement, meaning an average kappa estimate of .61 or greater (Landis & Koch, 1977). Average kappa estimates for the 508 cards ranged from .01 to 1.0 ($M =$ .61; $SD = .28$; Mdn = .68; see Appendix G). There were 264 cards (52%) that achieved high interrater agreement: 86 (16.9%) were within the substantial agreement range ($\kappa =$ .61-.80) and 178 (35%) were in the almost perfect range ($\kappa \geq .81$). Within those, 40 were classified as Concurrent (15%), 51 as Intermittent (19%), 79 as Segmented (30%), and 94 as Semi-Overlapping (36%).

The IRA strength varied across categories (see Appendix H). Kappa estimates

across the 264 cards reaching high agreement were greatest on average for cards

classified as Segmented (mean κ = .96; *SD* = .09), followed by Intermittent (mean κ =

.86; *SD* = .13), Concurrent (mean κ = .79; *SD* = .10), and lastly, Semi-Overlapping (mean

κ = .78; *SD* = .12). Said in a different way, Segmented seemed to be the most readily

distinguishable pattern of service utilization by coders, while Semi-Overlapping was the

least discernable or precise pattern category in the high agreement ranges (see Table 1).

All but three of the coders made use of the Other/Does Not Fit category option

during the card sort task. For those who did, rates of use ranged from 1–11%. Far and

away, the most common reason articulated by coders at the end of the sorting task as to

why they chose to put a card in the Other/Does Not Fit category was due to insufficient

data. That is, 12 of the 15 coders expressed being unable to perceive—or reluctant to

identify—a service utilization pattern for episodes/cards that contained limited sessions.

Several coders considered the number of sessions to be too few for cards displaying two

dots (i.e., sessions), while others indicated that three or four dots were too limited still.

Table 1.
Frequency Distribution of High Agreement Treatment Episodes (Cards) Across the Four Service Utilization Pattern Categories

	Concurrent (n = 40)		Intermittent (n = 51)		Segmented (n = 79)		Semi-Overlapping (n = 94)		Total (n = 264)	
	n	%	n	%	n	%	n	%	n	%
Substantial Agreement										
κ = 0.6786	0	0	7	14	2	3	33	35	42	16
κ = 0.6905	18	45	6	12	4	5	16	17	44	17
Almost Perfect Agreement										
κ = 0.8333	18	45	18	35	7	9	31	33	74	28
κ = 1.0000	4	10	20	39	66	84	14	15	104	39

Note. Percentages are rounded to nearest whole number.

Consensus was not high enough ($\kappa \geq .60$) across coders on any given card(s) to suggest the presence of an additional and separate service utilization category. Thus, Other/Does Not Fit was dropped from further analyses.

Taken together, the discussion group and interrater agreement analysis suggest the presence of discrete and identifiable combined treatment service utilization pattern categories, as coders agreed strongly on specific category classification for over half of the treatment episode cards presented.

Treatment Outcomes

Client demographics were similar for the Overall sample ($n = 508$) included in the study and the High Agreement (HA) sample included in the outcome analyses ($n = 254$). Intake/initial distress was in the clinical range (OQ score ≥ 63) for both samples, as well as for all four of the service utilization categories (see Table 2). The average length of treatment for clients in the High Agreement sample was 25.24 sessions ($SD = 15.09$; Mdn = 21; range = 5–67). On average, treatment episodes considered in the Semi-Overlapping category were approximately 1.5 times longer ($M = 32.54$ sessions) and significantly differed (Tukey HSD comparison; $p < .05$) from treatment episodes in each of the other categories, which hovered around 21 sessions. Of note, the average length of treatment for clients in the Low Agreement (LA) sample was 20.76 sessions ($SD = 12.97$; Mdn = 18; range = 2–59), and significantly differed (Tukey HSD comparison; $p < .05$) from treatment episodes in the High Agreement sample. The mean proportion of group per treatment episode was roughly 44%, meaning that on average close to 56% of services in a given combined episode of treatment was a service other than a group counseling

Table 2.
Treatment Episodes (Cards) Descriptive Statistics

	Concurrent (*n* = 40)		Intermittent (*n* = 51)		Segmented (*n* = 79)		Semi-Overlapping (*n* = 94)		Low Agree (n = 244)		High Agree (*n* = 264)		Overall (*n* = 508)	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Initial score	76.47	19.24	80.29	22.26	76.19	27.50	78.00	20.69	79.19	25.03	77.65	22.96	78.39	23.96
Last score	60.25	20.70	62.51	22.05	56.69	24.27	62.60	25.35	62.43	24.60	60.39	23.69	61.37	24.11
Sessions (any)	21.02	14.99	21.18	12.75	21.30	13.42	32.54[a]	15.08	20.76	12.97	25.24	15.09	23.09	14.27
Sessions (individual)	9.25	7.39	6.35	9.63	8.33	7.21	12.44	8.57	7.60	7.61	9.55	8.52	8.61	8.15
Sessions (group)	8.70	6.54	10.84	10.96	8.87	8.35	15.28	8.70	9.28	8.16	11.51	9.22	10.44	8.79
Proportion of group	41.10	8.13	48.70	32.71	40.63	16.21	47.61	15.18	42.52	20.04	44.75	19.64	43.68	19.85
Number of OQs	14.22	11.20	15.59	10.54	14.76	9.55	22.68	11.35	14.03	10.10	17.66	11.25	15.91	10.85
OQ Change	-14.30	21.92	-19.35	25.90	-16.15	19.78	-17.91	25.65	-13.87	21.26	-17.14	23.24	-15.59	22.34

[a] Differs from category means in the same row at *p* < .05 in the Tukey honestly significant difference comparison.

session (e.g., individual counseling, intake). The High Agreement sample showed a similar mean group proportion rate (*M* = 44.75; *SD* = 19.64; Mdn = 44.44; range = 2.27–93.75). Intermittent was associated with the highest mean and standard deviation of group proportion (*M* = 48.70; *SD* = 32.71; range = 2.27–93.75), while Segmented had the lowest mean (*M* = 40.63; Mdn = 41.67) and Concurrent had the lowest standard deviation (*SD* = 8.13; range = 25–56.25). However, differences in proportion of group did not achieve statistical significance.

Missing OQ scores precluded the calculation of a change score for all episodes. There were 146 episodes with available OQ scores at both first and last session. The average OQ change from first session to last session was roughly -17 points in the High Agreement sample and over -15 points in the Overall sample; both change scores exceed the RCI value of 14. These findings suggest that combined treatment generally results in clinical improvement. Average OQ change scores of the Low Agreement and High

Agreement samples did not significantly differ via Tukey's HSD comparison.

Intermittent was associated with the largest average OQ change ($n = 31$; $M = $ -19.35; $SD = 25.9$; range = -76–33), followed by Semi-Overlapping ($n = 47$; $M = $ -17.91; $SD = 25.65$; range = -84–35), Segmented ($n = 48$; $M = $ -16.51; $SD = 19.78$; range = -71–36), and lastly, Concurrent ($n = 20$; $M = $ -14.3; $SD = 21.92$; range = -67–25). Amongst those with available change scores in the High Agreement sample, a basic analysis of variance showed no significant differences in average OQ change from first to last session across the four service utilization categories [$F(3, 142) = 0.24$, $p = 0.87$]. The prediction of last session OQ scores based on category for the High Agreement sample was also explored through a linear regression model that controlled for a client's initial baseline score. Results from this analysis also suggest the absence of a significant main effect of category on last session score [$F(3, 141) = 0.44$, $p = 0.73$] (see Appendix I).

Model Fitting

Next, a multilevel modeling (MLM) procedure was used to test whether OQ score trajectories, rather than mere OQ change scores or final session OQ scores, differed significantly between the four service utilization pattern categories. This analysis was applied in order to evaluate clinical outcomes in a more nuanced and proper way given the hierarchical structure of the data. Linear mixed models, fit in this study using maximum likelihood estimation, evaluate the average baseline score (i.e., intercept) and the average rate of change (i.e., slope), which can be used to predict specific scores—and the shape of growth curves—at any moment in treatment (Singer & Willet, 2003). It is an appropriate approach when studying individual change as it creates a two-level

hierarchical model that nests time within the individual. Rate of change is commonly calculated as the average change in OQ score per one unit change in a time variable included in the model (e.g., session). The MLM procedure used in this study includes several steps: fit the unconditional means model; install the unconditional growth model, and if necessary, determine the appropriate transformation for the time variable; consider the effect of adding predictors to the model; and evaluate the expected trajectory related to each service utilization pattern category. Due to the repeated measures nature of these analyses, there were 10 combined treatment episodes having 2 or fewer available OQ scores that were withheld from the High Agreement sample for the MLM procedure.

Unconditional Means Model

Two related concepts in MLM for longitudinal change are the *unconditional means model* and *unconditional growth model* (Singer & Willett, 2003). The unconditional means model is an equation fit to the sample data based on the means of participants' scores. This equation is used as a baseline to determine whether the addition of predictor variables will improve the fit of the model. Moreover, this model helps in determining whether there is systematic variation in the outcome worth exploring and, if so, where that variation lies (i.e., within or between people). The first step in the process of developing a model to fit the data is defining the unconditional means model.

The statistic used to accomplish this is called the intraclass correlation (ICC), which estimates the proportion of total variance ($\tau_{00} + \sigma^2$) accounted for by between-group variance (τ_{00}) in a two-level model, regardless of time. The ICC can be interpreted as the variance explained by the grouping structure in the population, which in the

unconditional means model is each person's mean score exclusive of predictors (Hox, 2010). The ICC, denoted here as ρ, is computed as:

$$\rho = \frac{\tau_{00}}{\tau_{00} + \sigma^2}$$

The ICC was estimated at 0.68, meaning that about 68% of the total variance of a given OQ score taken from the High Agreement sample ($n = 254$) can be explained by interindividual (between-person) differences and about 32% of the total variance of a given score is intraindividual (within-person) differences. This high ICC is due to the longitudinal and nested nature of the data. Barcikowski (1981) showed that the Type I error rate could be vastly inflated even with a very small ICC (e.g., .01). For this reason, multilevel models are required here.

Unconditional Growth Model

The second step of the model building procedure is to add a time variable to the unconditional means model (i.e., session in therapy) in order to create what is known as the unconditional growth model. The unconditional growth model is a model with time as the only Level-1 (observation-level) predictor and no substantive predictors at Level 2 (person-level). The unconditional growth model has the ability to describe the total variation both within- and between-persons, thus providing a baseline for each individual's change trajectory in terms of initial OQ score (i.e., intercept) and rate of change (i.e., slope). This model provides the basis for evaluating the success of subsequent model building whereas the unconditional means model lacked this ability because it considered only the mean (i.e., intercept) of the individual without consideration of time (i.e., slope).

Visual review of the person-level change trajectories yielded no clear trend that could be visually discerned. Without sufficient evidence to determine the most likely slope of the recovery path (i.e., linear or nonlinear), a method for fitting the unconditional growth model to the data is required. Singer and Willet (2003) propose a data transformation procedure for model fitting with nonlinear change trajectories. A linear slope suggests that the rate of growth remains constant across time, whereas nonlinear trends indicate that the growth rates might not be the same over time.

The transformation of variables a commonly utilized procedure for many growth curve modeling and dose-response studies, as the observed rate of change during treatment (i.e., slope) is rarely linear. Tukey (1977) presented an ordered list of transformation options known as the *ladder of powers.* The trial and error approach suggests that the exponential power of a given variable V^1 can be adjusted either down the ladder (e.g., *log V*, $V^{1/2}$, V^{-1}, V^{-2}, etc.) or up the ladder (e.g., V^2, V^3, V^4, etc.) until it best approaches linearity.

In many dose-response studies, researchers model change during treatment as a logarithmic, polynomial, or log-linear functions of time (e.g., Lutz et al., 2001, 1999). The following model was used for this sample:

$$Y_{ij} = \beta_{00} + \beta_{10}(session)_{ij} + \beta_{20}(session)^2_{ij} + \beta_{30}(session)^3_{ij} + \left[b_{00j} + b_{10j}(session)_{ij} + e_{ij} \right],$$

where Y_{ij} is the OQ score at time i for person j, β_{00} is the grand mean (i.e., average OQ score for all clients in the sample at the beginning of treatment), β_{10} is the average linear or directional (positive or negative) change over time, β_{20} is the average quadratic or accelerational (increase or decrease) change over time, and β_{30} is the average cubic effect of growth (rapid or gradual) change over time. These components represent fixed effects.

The parameters inside the brackets represent the random effects. The random effects account for client variability around the overall intercept (b_{00j}) and the linear rate of change (b_{10j}). Estimation of a random effect for individual level variability around the quadratic and cubic change over time could not be achieved. Random within-person error is also accounted for (e_{ij}).

Session, session^2, and session^3 had a significant contribution in the model. The negative linear growth rate (β_{10} = -0.69; SE = 0.11; $p < 0.001$) suggests that a typical OQ score trajectory displays a directionally downward course, in general. The nonsignificant quadratic growth rate (β_{20} = -0.002; SE = 0.004) suggests neither an acceleration nor deceleration in the instantaneous linear rate of change, on average, over time. The negative valence of the quadratic effect suggests that a typical OQ score trajectory is concave to the time axis. The positive cubic growth rate (β_{30} = 0.0002; SE = 0.00; $p < 0.01$) reveals that the instantaneous rate of acceleration/deceleration, on average, marginally hastened over time. Given that the cubic unconditional growth model improved model fit over the previous unconditional means model (χ^2 (5) = 38429 – 36777 = 1652; $p < 0.001$; Δ AIC = 38435 – 36793 = 1642; Δ BIC = 38454 – 36844 = 1640), cubic growth curve parameters were retained in the subsequent models.

Adding Level-2 Predictor

The next step in the MLM procedure is to examine the effects of adding service utilization category as a level-2 predictor on the shape of individual growth trajectories. Service utilization category, a nominal variable containing four levels, was examined as a time-invariant covariate (i.e., fixed effect) in order to allow for exploration of any

category differences in change over time (i.e., interaction with time). This model examined whether category was a predictor of the intercept, linear, quadratic, and cubic parameters.

The four levels of the nominal predictor variable were contrast coded using a dummy coding scheme. Dummy coding, also called treatment coding, is one of the most commonly used coding schemes in treatment research. In the absence of a clear control group for comparison, this coding scheme compares each level of the predictor variable to a fixed reference level. Although identification of the reference category is generally arbitrary, some guidelines for choosing the reference category have been offered. Garson (2014) notes that using reference levels such as 'miscellaneous' or 'other' is not recommended due to the lack of specificity in those types of categorizations; for reasons related to sample size and error, the reference level should not be a level that contains just a few cases; and lastly, the use of a level that is middle/moderate or closest to treatment-as-usual can be appropriate, as it is likely to protects against comparisons against the extremes. In this study, Semi-Overlapping was chosen as the reference level due to its large sample size, its most middle/moderate ordinary least squares (OLS) mean trajectory (see Appendix J), and its general 'catch-all' quality, an attribute that was discussed by many of the coders during the card sort task. In effect, the coding scheme selected for this study compares outcomes (OQ scores) of the reference level to the outcomes of each of the other three category levels (i.e., Semi-Overlapping vs. Concurrent; Semi-Overlapping vs. Intermittent; Semi-Overlapping vs. Segmented).

The model was as follows:

$$Y_{ij} = \beta_{00} + \beta_{01} \, (category)_j + \beta_{10} \, (session)_{ij} + \beta_{20} \, (session)^2{}_{ij} + \beta_{30} \, (session)^3{}_{ij} +$$

$$\left[b_{00j} + b_{10j}(session)_{ij} + e_{ij} \right],$$

where fixed effect β_{01} represents the main effect for service utilization pattern category.

Model fit did not improve by adding this effect (Deviance = 36774; AIC = 36796; BIC =

36867), indicating that the intercepts of each service utilization category did not

significantly differ from one another.

Incorporating Interaction

      The next step of the MLM procedure was to test interaction effects. A test of the

interaction effects provides information on whether rate of change (i.e., slope) varied as a

function of the service utilization category. The random effects portions of this model and

the previous models are identical; however, the fixed effects portion is augmented by

terms representing the interaction between the linear, quadratic, and cubic rates of change

and the category level ($\beta_{11}, \beta_{12}, \beta_{13}$).

      The model was as follows:

$$Y_{ij} = \beta_{00} + \beta_{01}\,(category)_j + \beta_{10}\,(session)_{ij} + \beta_{20}\,(session)^2{}_{ij} + \beta_{30}\,(session)^3{}_{ij} +$$

$$\beta_{11}\,(category)_j\,(session)_{ij} + \beta_{12}\,(category)_j\,(session)^2{}_{ij} + \beta_{13}\,(category)_j\,(session)^3{}_{ij} +$$

$$\left[b_{00j} + b_{10j}(session)_{ij} + e_{ij} \right].$$

Results of the χ^2 tests confirm this final model to best estimate the sample (χ^2 (12) =

36777 – 36750 = 27; $p < 0.01$; Δ AIC = 36793 – 36790 = 3; Δ BIC = 36844 – 36919 =

-75), suggesting that a given OQ score is best expressed as a function of the interaction

between category and time (see Table 3). This model accounts for just over 81% of the

variance in OQ scores.

Table 3.

Results of Multilevel Models

	Unconditional Means Model	Unconditional Growth Model	Time x Category Interaction Model
Fixed Effects		*Coefficients*	
Intercept	68.78 (1.27)	75.90 (1.48)	76.71 (2.40)
Session		-.688 (.11)	-.842 (.17)
Session^2		-.002 (< .01)	.007 (< .01)
Session^3		.001 (< .01)	.001 (< .01)
Concurrent			.258 (4.54)
Intermittent			.439 (4.24)
Segmented			-3.21 (3.61)
Concurrent x Session			-.007 (.36)
Intermittent x Session			.888 (.32)
Segmented x Session			.005 (.30)
Concurrent x Session^2			-.003 (.02)
Intermittent x Session^2			-.006 (.01)
Segmented x Session^2			.002 (.01)
Concurrent x Session^3			.001 (< .01)
Intermittent x Session^3			.001 (< .01)
Segmented x Session^3			.001 (< .01)
Random Effects		*Variance Estimates*	
Intercept	396.6 (19.91)	489.9 (22.13)	487.1 (22.07)
Session		0.98 (.99)	1.01 (1.01)
Residual	186.6 (13.66)	114.07 (10.68)	113.28 (10.64)
Fit Statistics		*Goodness-of-Fit Estimates*	
Deviance	38429	36777	36750
AIC	38435	36793	36790
BIC	38454	36844	36919
Model Comparisons			
		$\chi^2_{12}(5) = 1651.7$**	$\chi^2_{13}(17) = 1678.5$**
			$\chi^2_{23}(12) = 26.8$*
Pseudo R^2	$R^2 = .679$	$R^2 = .812$	$R^2 = .813$

* $p < 0.01$, ** $p < 0.001$

Intercept corresponds to cell mean for reference level (Semi-Overlapping).
Standard Error estimates appear in parentheses for fixed effects.
Standard Deviation estimates appear in parentheses for random effects.
Subscripts on the χ^2 statistics denote the two models being compared.

There was high variability in mean OQ scores (487.1; $SD = 22.07$) across all combined treatment episodes in the analyzed sample, while the variability of the linear slope across episodes was low (1.01; $SD = 1.01$). There was a moderately low negative correlation between mean intercept and mean linear slope; specifically, an increase of one unit of standard deviation in a client's OQ score corresponded to a decrease of 0.39 standard deviations in linear rate of change.

Models testing explanatory covariates such as baseline/initial severity, treatment episode length, and proportion of group were also explored. Chi-square tests revealed that the addition of these variables, as well as the interaction of these variables with time (i.e., session, session^2, and session^3), did not improve model fit for this sample. Therefore, explanatory covariates were not included as part of the future predictive models.

Model Predictions

The final step utilized the best-fit model above to extrapolate outcomes outside the given study sample. Table 4 provides the fixed intercept (group mean) and slope (rate and curvature) estimates by category. In line with current statistical methodology literature (e.g., Bates, Maechler, Bolker, & Walker, 2015), p-values for single parameters are not provided; instead, the confidence interval for each predicted score was estimated.

Figure 3.1 and Figure 3.2 provide the predicted outcome trajectory for an episode of combined treatment and the predicted outcome trajectories for each of the four categories, respectively. The growth curve in Figure 3.1 suggests that, in general, combined treatment is associated with clinical improvement. As is further demonstrated in Figure 3.3 with 95% confidence intervals, outcome at any given time (especially up to

Table 4.

Fixed Effects (Predicted Outcomes) of Best-Fit Multilevel Model by Service Utilization Pattern Category

	Estimate	SE	*t*
Fixed Effect			
Concurrent	76.97	3.88	19.85
Intermittent	77.15	3.52	21.90
Segmented	73.50	2.72	27.01
Semi-Overlapping	76.71	2.42	31.70
Session x Concurrent	-.85	.32	-2.65
Session x Intermittent	.05	.27	.18
Session x Segmented	-.79	.25	-3.18
Session x Semi-Overlapping	-.84	.17	-5.02
$Session^2$ x Concurrent	.01	.01	.71
$Session^2$ x Intermittent	-.05	.01	-4.44
$Session^2$ x Segmented	< .01	.01	.33
$Session^2$ x Semi-Overlapping	.01	.01	1.10
$Session^3$ x Concurrent	< .01	< .01	-.21
$Session^3$ x Intermittent	< .01	< .01	5.43
$Session^3$ x Segmented	< .01	< .01	.30
$Session^3$ x Semi-Overlapping	< .01	< .01	.84

about session 30) appears to be more related to the given session number and less a result of the service utilization pattern category type. Due to attrition, confidence intervals expand as the session number increases. Thus, the upswing observed near session 45 in Figure 3.1 is likely an artifact of high severity high utilizers (see Baldwin et al., 2009; Melling, 2014). The distinctly pronounced curvature displayed in Segmented treatment is posited to be the result of smoothing together what would otherwise appear in actual trajectories to be more akin to several separate piece-wise functions.
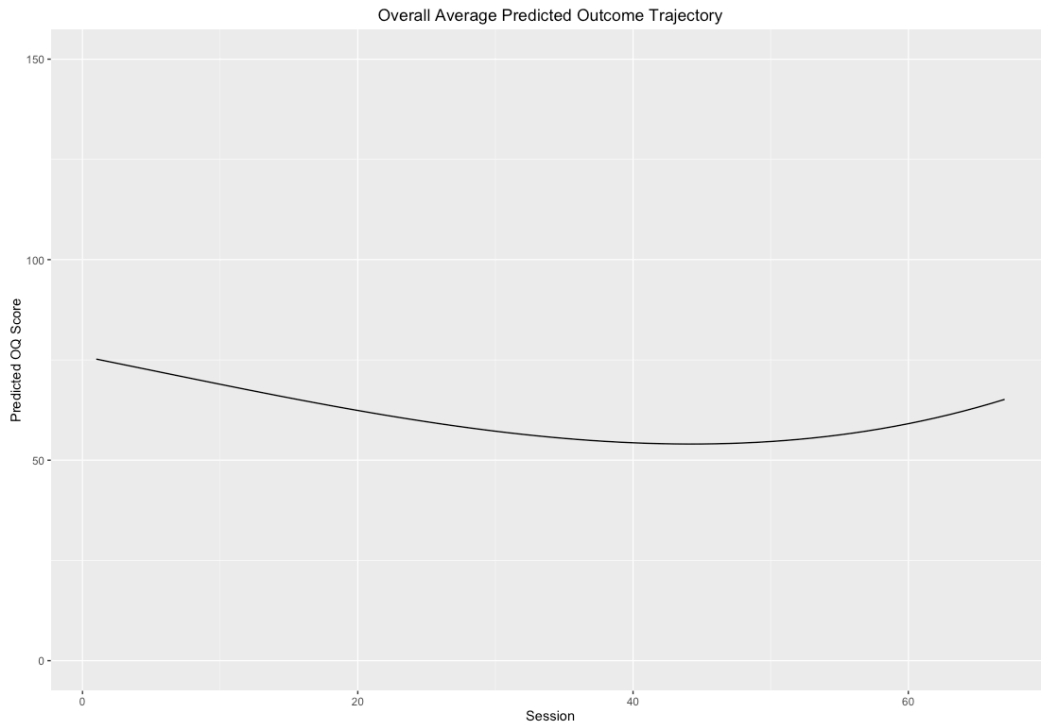
Figure 3.1. Overall average predicated outcome trajectory
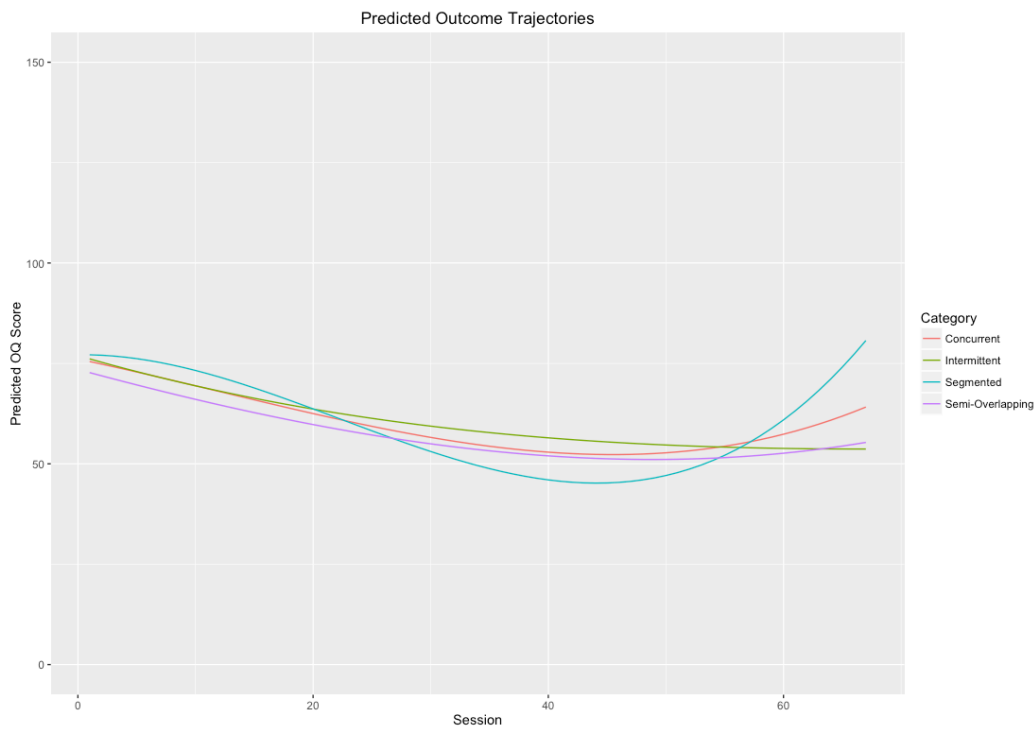
Figure 3.2. Predicted outcome trajectories by category
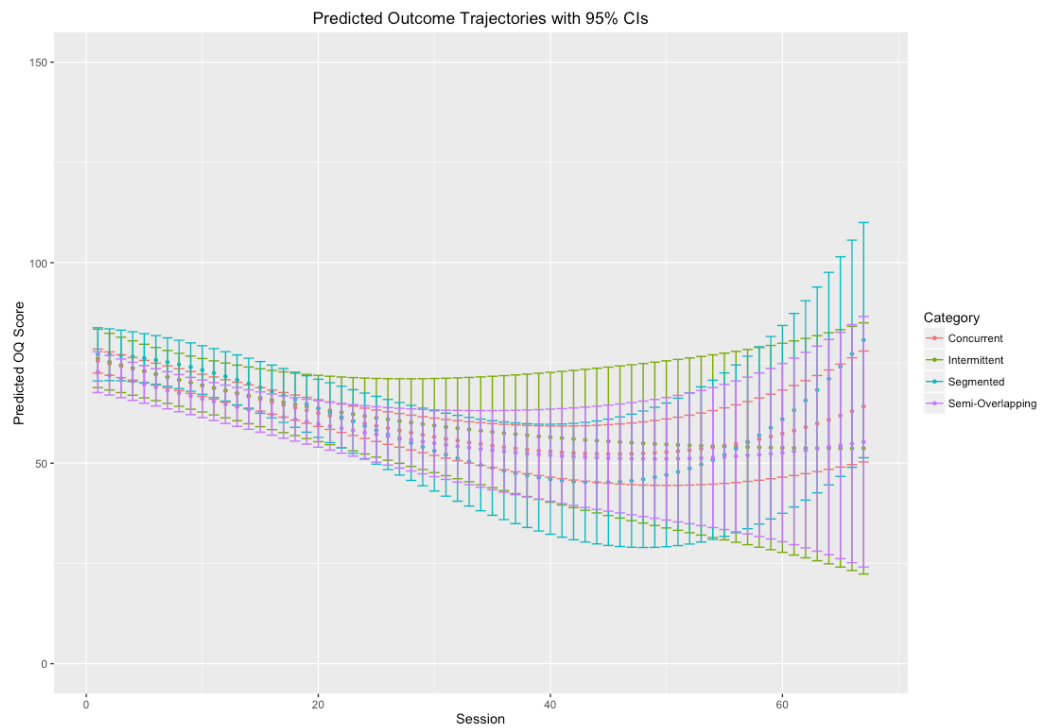
Figure 3.3. Predicted outcome trajectories by category with 95% CIs

CHAPTER 4

DISCUSSION

The purpose of this mixed methods study was to understand common patterns of combining individual and group psychotherapy services in everyday clinical practice, and to evaluate the overall and differential effectiveness of these common service utilization patterns within a naturalistic outpatient psychotherapy setting. Hypothesis 1 predicted the existence of discrete and identifiable patterns of combined (individual plus group) psychotherapy service utilization, and was investigated by way of a discussion group, as well as a hybrid card sort task. Hypothesis 2, which was investigated using a multilevel modeling procedure, predicted that (a) combined (individual plus group) psychotherapy treatment will demonstrate overall effectiveness, defined as a pre-post decrease in OQ-45 score greater than chance, and (b) that one or more specific service utilization configurations will be more highly associated with superior outcomes than others, defined as a greater average rate and magnitude of change in OQ-45 score.

Summary of Results

Findings from the discussion group and interrater agreement analysis procedure performed to address Research Question 1 suggest the presence of discrete and identifiable combined treatment service utilization pattern categories. The four service

utilization pattern categories identified—and later tested—were based on the relative quantity and sequencing of each type of service (i.e., individual or group). *Concurrent* combined treatment comprised episodes where the client engaged in both types of services in relatively equal amounts and close together in time, for a sustained period of time, with no long periods of utilizing only one service type. *Semi-Overlapping* combined treatment encompassed episodes where the client had at least one period of engaging in both service types concurrently and at least one period of engaging in only one service type (either individual or group). *Segmented* combined treatment incorporated episodes where the client engaged in individual or group in distinct segments, or strings, not overlapping in time, and with no long periods of utilizing both services types concurrently. *Intermittent* combined treatment contained episodes where the client engaged predominantly in one service type with interspersed sessions of the other service type. Coders did not achieve sufficient agreement by way of the *Other/Does Not Fit* category option to warrant the creation of an additional and separate service utilization category.

Notably, Segmented appeared to be the easiest pattern category for coders to reliably differentiate, while Semi-Overlapping appeared to be the least discernable (or, most imprecise) pattern category. Indeed, numerous coders highlighted a general 'catch-all' quality associated with the Semi-Overlapping category. Additionally, Semi-Overlapping demonstrated a statistically significant higher average total session count per treatment episode compared to the other three categories. Difference comparisons between each of the four categories with respect to the proportion of group per treatment episode did not achieve statistical significance.

Clinical outcomes observed in Research Question 2 mirror previously published findings (e.g., Burlingame et al., 2015). Namely, combined treatment in every day practice demonstrates, on average, reliable overall effectiveness. Yet none of the four combined treatment service utilization categories tested were reliably associated with superior effectiveness above and beyond the others. Predictive analyses indicate no significant category differences vis-à-vis posttreatment symptoms/functioning or longitudinal change trajectories. Likewise, difference comparisons between each of the four categories with respect to the average observed first visit OQ score, average observed last visit OQ score, and average observed OQ score change did not achieve statistical significance.

Limitations

One of the primary limitations of this study is methodological. Similar to other "bottom up" knowledge discovery or data mining approaches that rely upon the data to discover groups (e.g., agglomerative clustering/analysis, exploratory factor analysis, latent class analysis), the results produced by card sorting techniques are inherently dependent on the characteristics of the particular items or data samples included in the analysis. Moreover, most research methods designed to "find" groupings within a data set will in fact produce some grouping solution, regardless of the data's actual multivariate structure; thus, research must adequately evaluate the validity of a grouping solution before the solution can be conclusively confirmed (Blashfield, 1980). Essentially, inductive methods such as open card sorts require very large and/or highly representative data sets in order to ensure there is sufficient power and validity; otherwise, results are

considered exploratory and pending. As is true for the majority of deductive, or "top down," approaches to classifying data into groups (e.g., divisive clustering/analysis, confirmatory factor analysis), the utility of the findings from closed card sorts are diminished if the researcher is unaware of, or mislabels, any of the salient category groupings prior to the sort (Romesburg, 2004).

Although numerous non-peer-reviewed descriptions, case studies, and online blogs have been written in the last decade by researchers on the topic of card sorting, there have been few peer-reviewed articles that make use of the data collection method. Relatedly, there are little to no refereed publications or information with respect to the empirical measurement of its reliability or validity as a direct means of producing sound, practical taxonomies (Akerelrea & Zimmerman, 2002; Hannah, 2005). Although I estimated between-coder rating consistency for each card using interrater agreement analyses, I opted not to implement within-coder rating consistency checks (e.g., test-retest) because of concerns related to an already high volume of cards and a somewhat limited sample of potential coders. Future studies wishing to pursue card sort designs should look to employ internal consistency measurements.

These findings are also limited in their broad generalizability to other naturalistic psychotherapy situations. While this study drew from archival service utilization and outcome data collected over the course of almost a decade and a half, the combined treatment episodes assessed in this study totaled a mere several hundred and all came from a single counseling center. Additionally, treatment episodes included in the study were constrained to a client's first combined treatment episode, not their first treatment episode at the counseling center. In the Overall sample, 24% of episodes examined

represented a client's first episode of combined treatment but not their first episode at the site (Mdn=1; range=1-9). I defined the start and end of a treatment episode as a 90-day break in services, and counseling center policies formally restricted summer-time eligibility for individual counseling services (but did not restrict group counseling, crisis support, ongoing medication check-ups, and several other services). Clinical treatment programs can be quite different in design and implementation, as can the specific clientele and therapists who seek and provide services there. Consequences of particular data reduction decisions as well as factors related to the particular clinical setting and client sample examined in this study likely hinder the generalizability of its findings.

Relatedly, although careful selection of participants (i.e., licensed clinicians for the discussion group and graduate students in mental health fields as coders) contributed to what is considered to be a satisfactory outcome, the results are necessarily influenced when a more widely representative participant pool is not included. It is possible that a separate set of discussion group members and/or a separate set of coders would produce a separate (and equally satisfactory) result. Additionally, investigator bias could not be completely eliminated from the research process. I am a current or former colleague of the discussion group participants and the coders in this study. In this case, however, there were many positive factors that arose from the expertise and familiarity between me and the participants, and investigatory knowledge of the clinical setting from which the data were drawn. Furthermore, in-depth knowledge of the mental health treatment and psychotherapy services fields (on the part of the licensed clinicians, graduate student coders, and myself) allowed for greater insight and an adept perspective on the material being studied. These proved valuable in the study design, implementation, and analysis.

Therefore, the potential negative effects of personal bias were deemed relatively minimal for this particular situation.

Thus, given the vast gap in the literature pertaining to the manner(s) by which individual and group psychotherapy are being co-utilized, it is important to keep in mind that the mixed methods procedure performed in this study—and the results it derived— are preliminary in nature. The procedure performed in this study was not envisioned to be an immediate pathway to a stable clustering solution nor should the resulting classification structure described herein be seen as fully comprehensive and conclusive.

A major limitation of the clinical outcomes analyses was their low sample size. Thus, conclusions regarding overall and comparative effectiveness should be considered with caution. In addition, the naturalistic setting and missing data, coupled with the fact that service utilization category assignment was not random, arguably reduces confidence in the treatment outcome findings. That is, client, therapist, and treatment factors previously shown to account for systematic differences in therapeutic outcomes (e.g., baseline severity, prior treatment history, presenting problems or diagnosis, episode length, nesting effects due to therapist, group type and/or a group's co-members) were either not examined due to gaps in the data set, or they were examined but were not found to be statistically significant moderators of outcome. Such results may indicate that this particular sample of clients is somehow unique. On the other hand, it may be that the lack of significant moderator variables represents an artifact of the scarce empirical investigation and understanding of combined psychotherapy treatment in everyday clinical practice.

Lastly, several decisions were made regarding the visual representation of both

the service utilization patterns and the clinical outcomes of the combined treatment episodes. This study decided to mark time using session number rather than days or weeks. Services that were not either individual or group therapy sessions (e.g., intake, crisis services, pregroup screenings, psychiatric evaluation, medication management visits, etc.) were removed from view on the cards. Similarly, the manner in which the black line was drawn in order to connect group and individual sessions on the card included only horizontal and vertical lines rather than diagonals. Hence, the findings reported here should be approached with caution. Future research is advised in order to corroborate this preliminary exploration.

Implications of Results

The main contribution of this paper is the characterization of the various ways in which individual and group psychotherapy are being co-utilized in a naturalistic setting. The hypothesis regarding differential effectiveness of particular service utilization configurations could not be well tested given the data issues described above. If nothing else, the extraordinary complexity associated with how individual and group treatments are being combined in daily clinical practice is evident here.

Findings reported here regarding the general effectiveness of combined psychotherapy treatment align with those previously published. The average observed OQ scores for first visit, last visit, and change scores for the full counseling center data set were approximately 74, 63, and -11, respectively. Meanwhile, among all 557 combined treatment episodes in the full center data set, these same observed scores were approximately 79, 62, and -15. Also among these 557 episodes, 48% demonstrated

reliable improvement and 25% met criteria for treatment success, whereas 8% showed reliable deterioration and 2% represented a treatment failure. No reliable change was seen in 44% while 64% experienced no functional change. While combined treatment in everyday practice demonstrates empirically reliable positive change, on average, its trajectory of change appears to be somewhat slow and shallow and it transpires across a relatively lengthy spell of sessions. The average total number of sessions per episode in the study sample was 23 sessions whereas the average total number of sessions per episode for the full 14-year data set was 7 sessions. Similar observations were made by Burlingame and colleagues (2015), also using archival data from a university counseling center.

Researchers and clinical administrators have posed questions about whether clients involved in combined treatment are different from those in either individual only or group only modalities, on some important characteristic(s). Possibly shedding light on this question are some research findings that have indicated an association between college counseling center clients' reported presenting concerns and their reported amount of improvement, such that clients who desire to address low productivity or elevated stress levels demonstrate significant treatment gains in fewer sessions, whereas clients who indicate a high general distress or concerns related to interpersonal issues at intake take much longer to show clinical improvement (Minami et al., 2008; Stulz, Lutz, Leach, Lucock, & Barkman, 2007). Given the proposed benefits of both individual and group, it is plausible that the clients who are referred to combined treatment are clients who are more representative of the latter cohort described above. While such conclusions are currently tentative, they are intriguing for researchers and clinicians in favor of the

combined format.

In this study, direct statistical analyses comparing OQ scores from the entire counseling center data set and scores from the combined treatment sample were not conducted. Yet, at a glance, it appears as if clients who participated in combined treatment during the 14-year period entered the episode at a higher level of self-reported symptom distress and functional impairment. It also appears as if combined treatment clients tended to experience a greater degree of improvement. The observation of similar last visit scores between the two samples corroborates the 'good enough level' (GEL) model (Baldwin et al., 2009; Barkham et al., 2006) for an atypical subset of the general clinical population. That is, the combined treatment trajectory curves presented here imply that clients will continue to utilize services until they feel that they have reached an acceptable level of improvement (Nielsen, Bailey, Nielsen, & Pedersen, 2016).

The relative expensiveness of combined treatment (i.e., multiple therapists, many more hours, many more sessions) remains a credible concern for this modality. Furthermore, results of the current study were not able to identify a superior combined treatment service utilization pattern amongst the four that were evaluated. Clinicians should weigh multiple factors when considering whether—and how—to administer combined treatment in every day practice.

As a field, there appears to be a shortage of careful consideration and investigatory interest in the complex issues surrounding the implementation and outcomes of combined treatment in naturalistic settings. Based on responses from discussion group participants, the circumstances under which clinicians did or did not refer clients to both individual and group therapy suggested a lack of deliberation

founded on the virtues of combined treatment, and speak to the striking deficiency of empirical data on the utilization and effectiveness of combined therapy. The lack of evidence-based information regarding whether and how to add or change psychotherapies remains a major and persisting gap in the field of mental health treatment (Markowitz & Milrod, 2015). Issues related to the augmentation and switching of therapies must be prudently evaluated, as they have implications for referral practices and coordination of care, resource allocation and cost effectiveness, expected efficiency of treatment recovery, as well as expected overall treatment efficacy. Well-established evidence should guide the clinical practices and administrative decision making of mental health clinicians, systems, payers, and policy makers. The present study aims to act as a starting point for others who care about the successful utilization and optimal effectiveness of talk therapy services in outpatient mental health care.

<u>Recommendations for Future Research</u>

While the findings from efficacy studies are promising, the empirical investigation of combined treatment's implementation and effectiveness in actual naturalistic contexts remains profoundly deficient. Somewhat unsurprisingly, the relatively small body of published research is plagued by large gaps and mixed findings. Greater empirical understanding is sorely needed.

Future research in counseling centers and other naturalistic settings will be crucial in order to determine whether the results included herein can be replicated. Like all naturalistic samples, the current findings are limited to the type of clients treated in this specific counseling center. Although the general impressions and results presented here

should be applicable and helpful to other clinicians, administrators, and researchers, the degree to which these findings apply to other populations and settings (i.e., community/nonprofit service agencies, intensive outpatient programs, private practice, residential care, seriously or persistently mentally ill, etc.) are left to future research endeavors.

Archival analyses similar to the one performed here, with other populations and in other settings, will prove valuable and promote confidence in the existent findings. Coders agreed strongly on specific category classifications for just over half of the treatment episode cards presented. Further studies that aim to detect, define, and compare various versions of individual-plus-group combined treatment are highly recommended. In addition, studies that aim to identify and test the influence of factors like therapist and client characteristics, group moderators, differences in theoretical approach, as well as other known potential covariates, will all offer a richer understanding of combined treatment. Investigations regarding the type of group (e.g., support, skills, process, concern- or diagnosis-specific, demographic-specific) and the timing of group (e.g., beginning, middle, end, throughout the episode) are recommended. Group therapy theory and research suggests that elements such as treatment coordination, client expectations, synergy in theoretical orientation (e.g., integrative versus eclectic), agreed upon goals and reasons for referral, and ultimately, the manifestation of a cohesive and nonredundant therapeutic experience, are all apt to having a critical impact (e.g., Yalom, 2005). Research using other metrics and measures of clinical outcomes is recommended, as well, including measures of alliance, group therapy outcomes, and treatment satisfaction. All of these areas are virtually unexplored within the realm of naturalistic individual-plus-group

combined treatment. Only when the various manners and mechanisms of combined treatment have been sufficiently explored will we be able to competently combine individual and group-based services in everyday mental health treatment and psychotherapy clinical practice.

<u>Conclusion</u>

The current study aimed to address the lack of empirical examination into the combining of individual and group psychotherapy in everyday practice. There has been little research into the nonmanualized mixing of individual and group-based talk therapy in naturalistic settings, and until now, absent is an attempt to describe and evaluate the manners by which the two formats are co-utilized. The findings suggest that there are several distinct ways that combined treatment is presently being carried out in day-to-day mental health care. The procedure drew upon input from licensed clinicians and graduate students in mental health fields, and is analogous to other clustering, agreement, and classification methods already widely used in psychotherapy research. The service utilization pattern categories that were identified—and later tested—in this study were defined by the comparative amounts and timing of each service type (i.e., individual or group). Treatment outcomes observed in this data sample parallel previously published findings; specifically, that combined treatment in every day outpatient mental health care shows, on average, reliable absolute effectiveness. No particular combined treatment category outperformed the other identified combined treatment categories. Results of this study are seen as preliminary. Further research is encouraged. Nevertheless, the present study intends to help guide clinical referral practices and policy decision making

regarding the co-utilization of individual and group psychotherapy in naturalistic settings, and hopefully, aids in the advancement of our extant knowledge of combined talk therapy treatments.

APPENDIX A

THE STANDARDIZED DATA SET (SDS)

University Counseling Center
THE UNIVERSITY OF UTAH

University of Utah Counseling Center
Note

Name: Client Practice	**ID:** 88888882
Type: Client Intake Forms	
Counselor:	**Date and time:** 07/23/1901

Client Information (SDS)

HIPAA Form Status: - <No Response>

Date Signed: - <No Response>

Date of birth - <No Response>

Gender - <No Response>

Race / Ethnicity - <No Response>

Other Race / Ethnicity - <No Response>

What is your country of origin? - <No Response>

Religious Affiliation - <No Response>

Other religious or spiritual preference - <No Response>

Relationship status - <No Response>

Sexual orientation - <No Response>

Number of children - 0

Current academic status - <No Response>

Other academic status - <No Response>

College - <No Response>

Major - <No Response>

Are you an international student? - NO

What is your current GPA? - 0

Credits this semester - 0

Did you transfer from another campus/institution to this school? - NO

What is the average number of hours you work per week during the school year (paid employment only)? - 0

What kind of housing do you currently have? - <No Response>

Other housing - <No Response>

With whom do you live? (check all that apply)
    <No Response>

Others living with - <No Response>

University Counseling Center
THE UNIVERSITY OF UTAH

Have you ever been enlisted in any branch of the US military (active duty, veteran, national guard or reserves)? - NO

Did your military experiences include any traumatic or highly stressful experiences which continue to bother you? - NO

OBSOLETE: Do you have a diagnosed and documented disability? (check all that apply)
    <No Response>

Other disability: - <No Response>

Who referred you to the Counseling Center?
    Other (please specify)

Other referral: - <No Response>

Do you have health insurance? - YES

If yes, please indicate carrier: - student health insurance

Are you currently taking prescription medication? - NO

If you are taking prescription medication, please list: - <No Response>

If you are a student, how much is your academic performance affected by the issues for which you are seeking counseling? - <No Response>

How much is your job performance affected by the issues for which you are seeking counseling? - <No Response>

How much are your personal relationships affected by the issues for which you are seeking counseling? - <No Response>

Attended counseling for mental health concerns - <No Response>

Taken a prescribed medication for mental health concerns - <No Response>

Been hospitalized for mental health concerns - <No Response>

Received treatment for alcohol or drug use - <No Response>

Purposely injured yourself without suicidal intent (e.g., cutting, hitting, burning, hair pulling, etc.) - <No Response>

Seriously considered attempting suicide - <No Response>

Made a suicide attempt - <No Response>

OBSOLETE: Seriously considered injuring another person - <No Response>

OBSOLETE: Intentionally injured another person - <No Response>

Had unwanted sexual contact(s) or experience(s) - <No Response>

Experienced harassing, controlling, and/or abusive behavior from another person (e.g., friend, family member, partner, or authority figure) - <No Response>

Parents divorced or permanently separated before you were 18 - <No Response>

Family frequently moved - <No Response>

Parent(s) unemployed for an extended period of time - <No Response>

Frequent, hostile arguing among family members - <No Response>

Death of parent(s) before you were 18 - <No Response>

University Counseling Center
THE UNIVERSITY OF UTAH

Parent(s) with a drinking problem - <No Response>

Parent(s) with a drug problem - <No Response>

Parent(s) with a gambling problem - <No Response>

Physical abuse in your family - <No Response>

Sexual abuse in your family - <No Response>

Rape/sexual assault of yourself or family member - <No Response>

Family member hospitalized for emotional problems - <No Response>

Family member diagnosed with a mental disorder - <No Response>

Family member attempted suicide - <No Response>

Family member committed suicide - <No Response>

Family member with a debilitating illness, injury, or handicap - <No Response>

Family member prosecuted for criminal activity - <No Response>

Family member with an eating problem - <No Response>

Client Practice 88888882
on 07/23/1901

University Counseling Center
THE UNIVERSITY OF UTAH

University of Utah Counseling Center
Note

Name: Practice Practice	**ID:** 55555555
Type: Client Intake Forms	
Counselor:	**Date and time:** 08/17/2011

Presenting Concerns

Please check the items that brought you into the Counseling Center today (check all that apply):

Anxiety
Depression
Loneliness
Obsessive Compulsive Disorder
Relationship with Friends
Self-Esteem
Shyness
Stress

Other presenting concern (if applicable): - <No Response>

Practice Practice 55555555
on 08/17/2011

APPENDIX B

THE OUTCOME QUESTIONNAIRE (OQ-45)

Outcome Questionnaire (OQ®-45.2)

Instructions: Looking back over the last week, including today, help us understand how you have been feeling.   Read each item carefully and mark the box under the category which best describes your current situation. For this questionnaire, work is defined as employment, school, housework, volunteer work, and so forth.

Name:_____ Age:_____Yrs

Sex

ID#_____ M ☐ F ☐

Session #_____   Date____/____/____

		Never	Rarely	Sometimes	Frequently	Almost Always	SD	IR	SR
1.	I get along well with others.	☐ 4	☐ 3	☐ 2	☐ 1	☐ 0			
2.	I tire quickly.	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4			
3.	I feel no interest in things.	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4			
4.	I feel stressed at work / school	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4			
5.	I blame myself for things.	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4			
6.	I feel irritated.	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4			
7.	I feel unhappy in my marriage / significant relationship.	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4			
8.	I have thoughts of ending my life.	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4			
9.	I feel weak.	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4			
10.	I feel fearful.	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4			
11.	After heavy drinking, I need a drink the next morning to get going. (If you do not drink, mark "never")	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4			
12.	I find my work / school satisfying.	☐ 4	☐ 3	☐ 2	☐ 1	☐ 0			
13.	I am a happy person.	☐ 4	☐ 3	☐ 2	☐ 1	☐ 0			
14.	I work / study too much.	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4			
15.	I feel worthless.	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4			
16.	I am concerned about family troubles.	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4			
17.	I have an unfulfilling sex life.	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4			
18.	I feel lonely.	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4			
19.	I have frequent arguments.	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4			
20.	I feel loved and wanted.	☐ 4	☐ 3	☐ 2	☐ 1	☐ 0			
21.	I enjoy my spare time.	☐ 4	☐ 3	☐ 2	☐ 1	☐ 0			
22.	I have difficulty concentrating.	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4			
23.	I feel hopeless about the future.	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4			
24.	I like myself.	☐ 4	☐ 3	☐ 2	☐ 1	☐ 0			
25.	Disturbing thoughts come into my mind that I cannot get rid of.	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4			
26.	I feel annoyed by people who criticize my drinking (or drug use) (If not applicable, mark "never")	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4			
27.	I have an upset stomach.	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4			
28.	I am not working /studying as well as I used to.	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4			
29.	My heart pounds too much.	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4			
30.	I have trouble getting along with friends and close acquaintances.	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4			
31.	I am satisfied with my life.	☐ 4	☐ 3	☐ 2	☐ 1	☐ 0			
32.	I have trouble at work / school because of drinking or drug use. (If not applicable, mark "never")	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4			
33.	I feel that something bad is going to happen.	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4			
34.	I have sore muscles.	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4			
35.	I feel afraid of open spaces, of driving, or being on buses, subways, and so forth.	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4			
36.	I feel nervous.	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4			
37.	I feel my love relationships are full and complete.	☐ 4	☐ 3	☐ 2	☐ 1	☐ 0			
38.	I feel that I am not doing well at work / school.	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4			
39.	I have too many disagreements at work / school.	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4			
40.	I feel something is wrong with my mind.	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4			
41.	I have trouble falling asleep or staying asleep.	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4			
42.	I feel blue.	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4			
43.	I am satisfied with my relationships with others.	☐ 4	☐ 3	☐ 2	☐ 1	☐ 0			
44.	I feel angry enough at work / school to do something I might regret.	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4			
45.	I have headaches.	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4			

+   +

Total=

APPENDIX C

CARD SORT INSTRUCTIONS

Card Sort Instructions

There are two steps to today's activity. I'm going to describe Step 1 now, and will tell you about Step 2 when we're finished with Step 1.

STEP 1
Here's how it works. In front of you is a stack of cards. Each card contains a graphical representation of an actual client's episode of combined treatment at a college counseling center. You will notice that each card has an ID number along the top. You will not need to pay attention to or use that number for anything during this task—feel free to just ignore it. The horizontal axis along the bottom of the plot indicates the session number, while the vertical axis along the left side of the card represents the session type (either individual or group). Each session of individual or group therapy attended by the client is shown as a red dot. A black line connects the dots for ease of interpretation. Any session attended by the client that is neither an individual nor group therapy session has been removed from the card and will not be shown as a red dot. Therefore, although the black line will continue uninterrupted, you may notice instances of extended gaps between the red dots.

Also in front of you are four category labels: CONCURRENT; SEMI-OVERLAPPING; SEGMENTED; and INTERMITTENT.

- CONCURRENT: for episodes where the client is engaging in both types of services concurrently (meaning, in relatively equal amounts and close together in time), for a sustained period of time; there are no long periods of utilizing only one type
- SEMI-OVERLAPPING: for episodes where the client has one or more period(s) of engaging in both service types at the same time, as well as one or more period(s) of engaging in only one of the service types (either individual or group)
- SEGMENTED: for episodes where the client is engaging in individual and group in distinct segments, or strings, which do not overlap in time; there are no long periods of utilizing both services types at the same time
- INTERMITTENT: for episodes where there is regular engagement in one service type and infrequent/irregular engagement in the other type (meaning, mostly group sessions with interspersed individual sessions or mostly individual sessions with interspersed group sessions)

There are a total of 508 cards. I will present you with approximately 100 of those cards at a time. Your task is to place each of these cards into one of the four categories, based on its goodness-of-fit with that category's definition. There is no right or wrong answer—just do what makes most sense to you. If you come across a card that you think does not belong or is not appropriate for any of the four specified categories, you may create an OTHER/DOES NOT FIT pile for those cards off to the side.

Any questions before we start? Feel free to ask questions during the activity if you feel the need. I can't guarantee that I can answer them during the activity, but I'll do my best to answer them when you're finished. Thanks again for your participation. You may begin whenever you're ready.
_____

STEP 2 (if appropriate; instructions to be read at the conclusion of Step 1)
There is one last part of this activity. For this, I'm going to ask you look through the cards you put off to the side, in the OTHER/DOES NOT FIT pile. I'd like you to organize these cards into groups, or categories, that make sense to you. I'd also like you to assign a label name to the new categories you create. Again, same as before, there is no right or wrong answer to this. Just do what comes naturally.

APPENDIX D

SERVICE UTILIZATION PATTERN CARD EXAMPLES

Depicted below are four randomly selected cards. Each card depicts the service utilization pattern for one unique combined treatment episode.
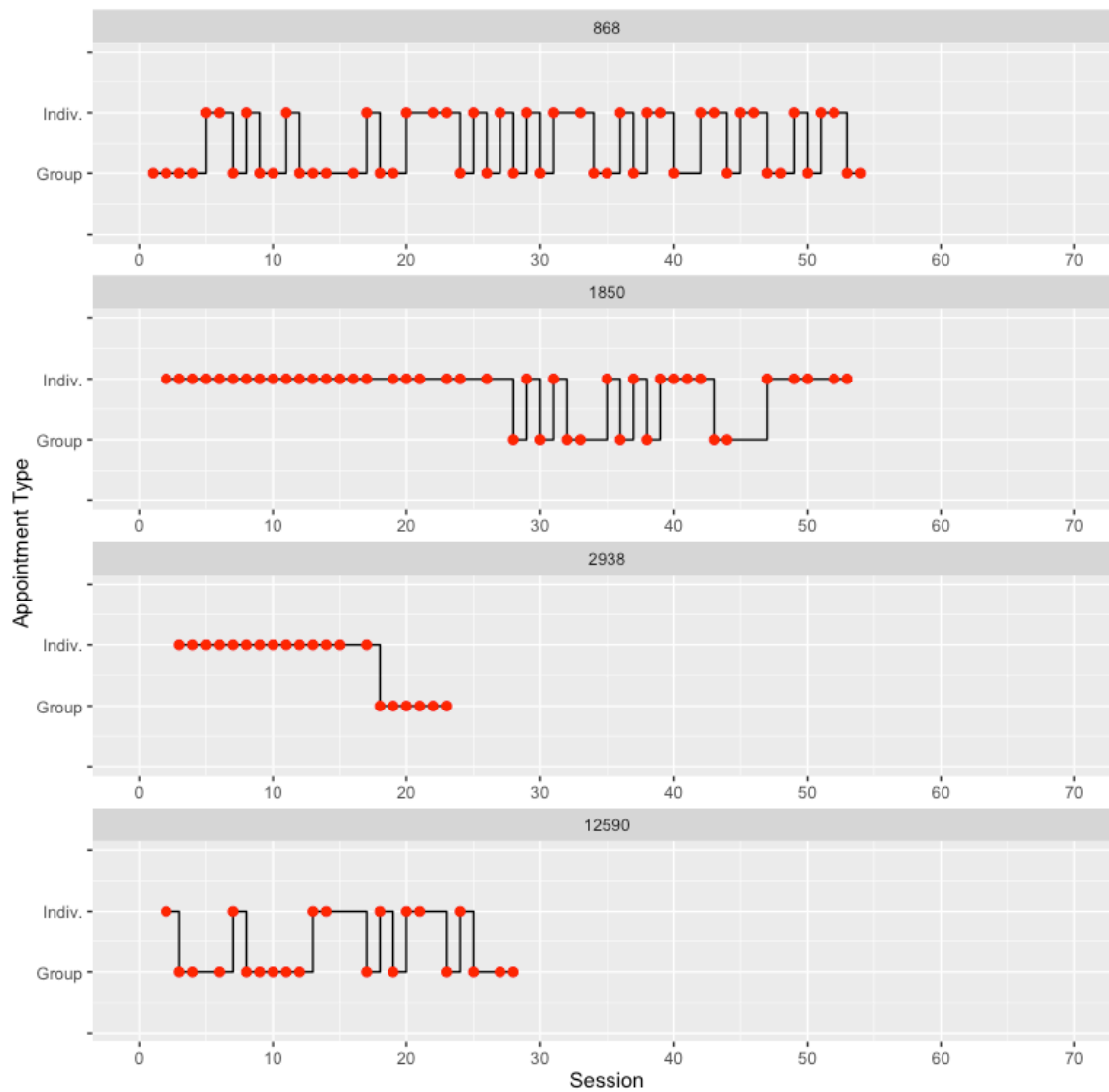
APPENDIX E

DISCUSSION GROUP RESPONSES

1. ***How, when, & why did you look to implement combined treatment, or suggest that clients utilize it?***

- It rarely happened that referral was to both individual and group immediately from intake/clinical team
- We did sometimes make a referral to both individual and group from team

- Reasons and circumstances for referring to second modality:
    - Client in individual therapy and would benefit from group
    - Difficult group member and would benefit from individual
    - To get more eyes on the same client
    - Client in individual or group and symptoms exacerbate/ do not remediate
    - Weaning away from individual
    - Client in both but then just stops going to group
    - If I thought client needed more, we could give them more
    - Provide additional support for higher severity clients
    - Provide more than one point of contact each week for higher need clients
    - Maybe I didn't see group as a treatment modality of its own to manage depression, anxiety, etc. – so I would think the client needed individual also
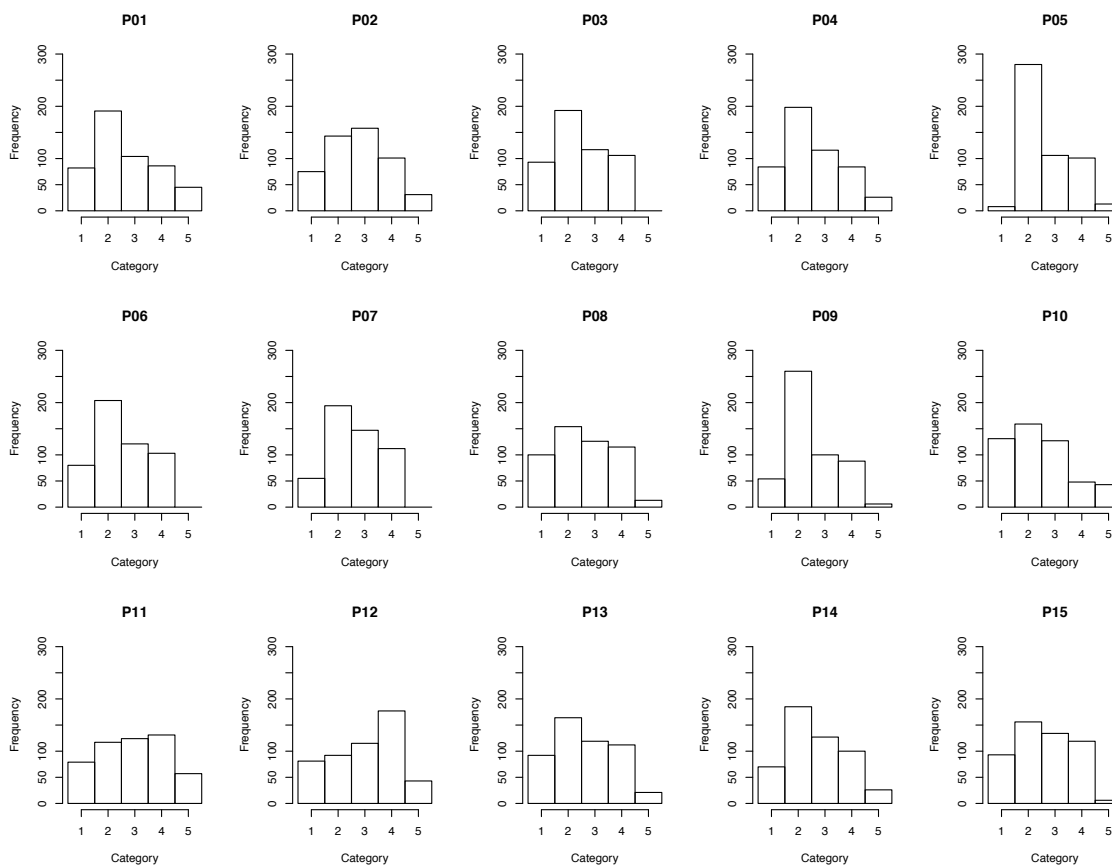   - At end of Spring semester when client would no longer be eligible for services over the summer, or any time when client had reached 12 session limit: refer to group for further support
    - At team meeting: want to refer to group, but client and/or intake clinician is reluctant, so refer to both
    - Practice interpersonal skills/interventions in group and process emotional responses individually
    - Facilitating transfer from individual to group (making it easier to transition from individual to group)
    - Because we could
    - For clients with severe mental illness, client can reap interpersonal benefits from group and take care of mental illness concerns individually
    - Group for social support or validation for certain populations; individual to work on issues particular to the person (i.e., LGBTQ)

- Reason not to refer to both:
    - Dilution effect—client might not participate fully in group – "save it for individual"
    - Mindful of resources, so may not have referred to both
    - Pushed for group only to manage resources; put lower-severity people into group only
    - Don't recall a lot of coordination between individual and group leaders
    - If I was the group facilitator, sometimes I wouldn't refer due to issue of having your individual clients in your group
    - It was difficult to refer to group if it wasn't at the beginning of the semester or the beginning of treatment (e.g., not knowing what groups were open, etc.)

2. **What did you observe in terms of service utilization within the combined treatment format? What patterns of individual and group session attendance took place? Essentially, what were common or typical ways that clients engaged in individual and group sessions over time?**

- o Individual as adjunct to group
  - People didn't attend group regularly but attended individual regularly
  - Start with individual while waiting until there is a group opening
  - In individual and refer to group – but client less likely to go and stick in the group; would go 1-2 x only
  - I thought that clients utilized group less when also in individual—both in terms of level of engagement and in terms of attendance and likelihood of staying in group

- o Group as adjunct to individual
  - I would see clients who were in my group on an individual basis occasionally, as needed. This would be a model of ongoing group with occasional individual
  - People didn't attend individual regularly but attended group regularly

- o Regularly attended both
  - Refer at intake to weekly group and monthly individual
  - Regular attendance in both: weekly group with ongoing individual once every 2-3 weeks

- Observed Patterns:
  - o Both throughout
  - o Both-Individual
  - o Both-Group
  - o Group-Both-Group
  - o Group-Both
  - o Group-Individual
  - o Group-Both-Individual
  - o Consistent group with a few individual sessions here and there as needed
  - o Individual-Both-Individual
  - o Individual-Both
  - o Individual-Group
  - o Individual-Both-Group
  - o Consistent individual with a couple group sessions

APPENDIX F

PREVALENCE AND SYTEMATIC BIAS

The histograms below depict the distribution of cards across categories for each of the 15 individual coders (top), and the total sample included in outcome analyses (bottom). There appears to be unequal distribution of observed ratings (i.e., cards) across categories thus indicating a prevalence problem.


Category types: 1=Concurrent; 2=Semi-Overlapping; 3=Segmented; 4=Intermittent; 5=Other

TOTAL (n=264)

Systematic bias is present when the distribution of specific ratings substantially differs across coders. A systematic bias between two coders can be assumed when the ratio substantially deviates from 0.5 while yielding a significant Chi-squared statistic (Gamer et al., 2012). As demonstrated below, there does not appear to be a systematic bias problem.

|	P01	P02	P03	P04	P05	P06	P07	P08	P09	P10	P11	P12	P13	P14	P15
P01		0.63 p=.01	0.25 p=.00	0.29 p=.00	0.53 p=.39	0.31 p=.00	0.48 p=.65	0.32 p=.00	0.32 p=.00	0.23 p=.00	0.60 p=.01	0.72 p=.00	0.39 p=.02	0.46 p=.37	0.41 p=.02
P02			0.24 p=.00	0.29 p=.00	0.48 p=.72	0.23 p=.00	0.39 p=.01	0.27 p=.00	0.29 p=.00	0.18 p=.00	0.56 p=.11	0.68 p=.00	0.34 p=.00	0.35 p=.00	0.32 p=.00
P03				0.49 p=.79	0.75 p=.00	0.51 p=.83	0.78 p=.00	0.51 p=.78	0.48 p=.72	0.41 p=.01	0.76 p=.00	0.85 p=.00	0.57 p=.11	0.58 p=.06	0.54 p=.43
P04					0.75 p=.00	0.54 p=.33	0.67 p=.00	0.54 p=.36	0.50 p=1.0	0.36 p=.00	0.71 p=.00	0.87 p=.00	0.56 p=.22	0.64 p=.00	0.53 p=.54
P05						0.29 p=.00	0.42 p=.09	0.31 p=.00	0.18 p=.00	0.29 p=.00	0.52 p=.54	0.66 p=.00	0.35 p=.00	0.36 p=.00	0.31 p=.00
P06							0.77 p=.00	0.49 p=.79	0.47 p=.57	0.37 p=.00	0.75 p=.00	0.84 p=.00	0.54 p=.40	0.60 p=.02	0.52 p=.72
P07								0.35 p=.00	0.28 p=.00	0.29 p=.00	0.61 p=.00	0.77 p=.00	0.41 p=.03	0.42 p=.07	0.39 p=.00
P08									0.47 p=.48	0.34 p=.00	0.72 p=.00	0.89 p=.00	0.53 p=.45	0.61 p=.02	0.50 p=1.0
P09										0.41 p=.01	0.73 p=.00	0.79 p=.00	0.54 p=.25	0.60 p=.01	0.49 p=.82
P10											0.73 p=.00	0.84 p=.00	0.70 p=.00	0.72 p=.00	0.68 p=.00
P11												0.59 p=.00	0.32 p=.00	0.36 p=.00	0.27 p=.00
P12													0.20 p=.00	0.19 p=.00	0.22 p=.00
P13														0.56 p=.21	0.45 p=.24
P14															0.43 p=.12
P15															

APPENDIX G

INTERRATER AGREEMENT

Depicted below is the strength of interrater agreement in regards to category classification for each of the 508 cards (top), and the frequency distribution of kappa estimates across all cards (bottom). A dashed red line has been added at $\kappa = .60$ to enhance visualization of the 'substantial agreement' cutoff.

Strength of Agreement for each Card (n=508)
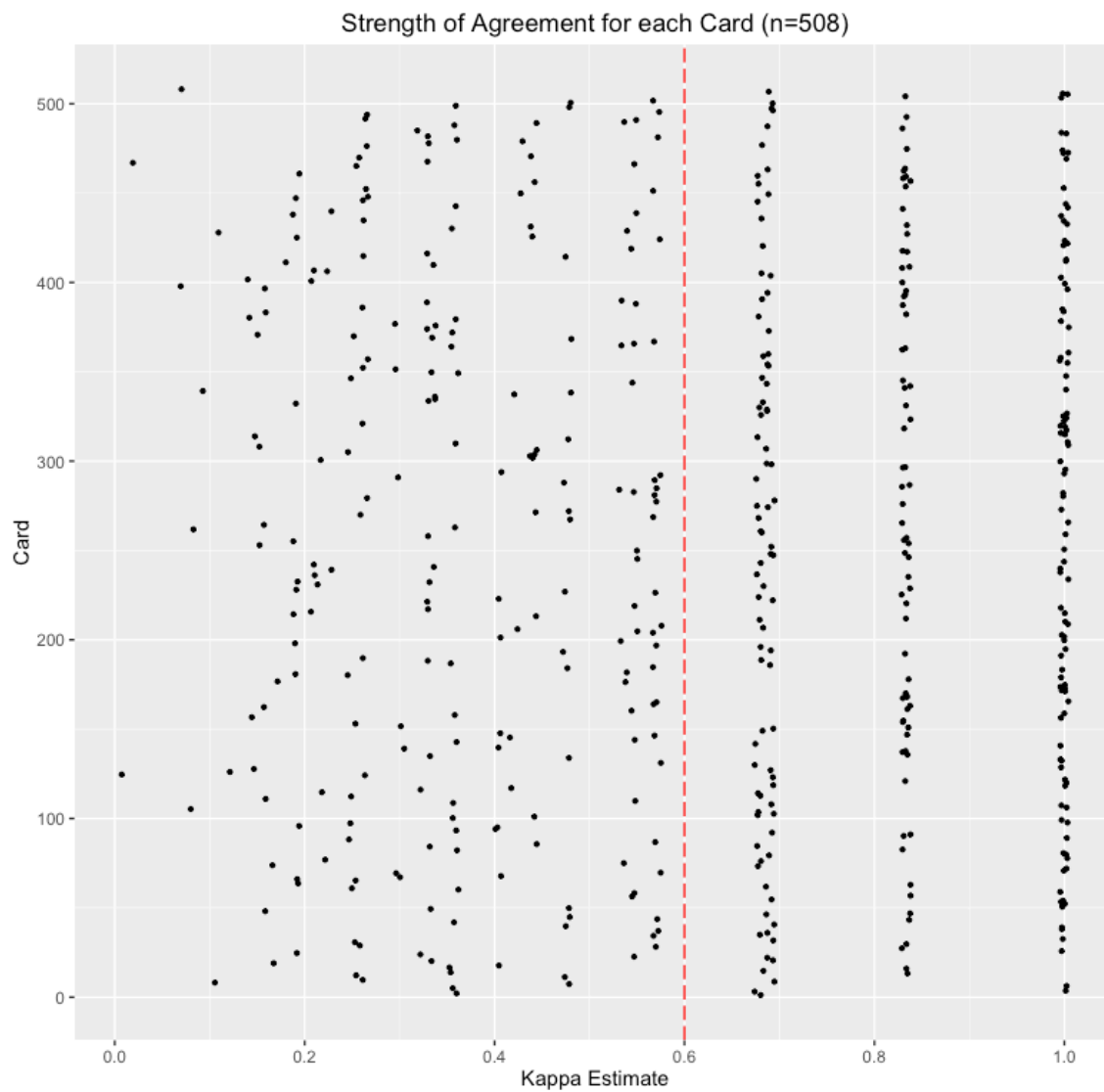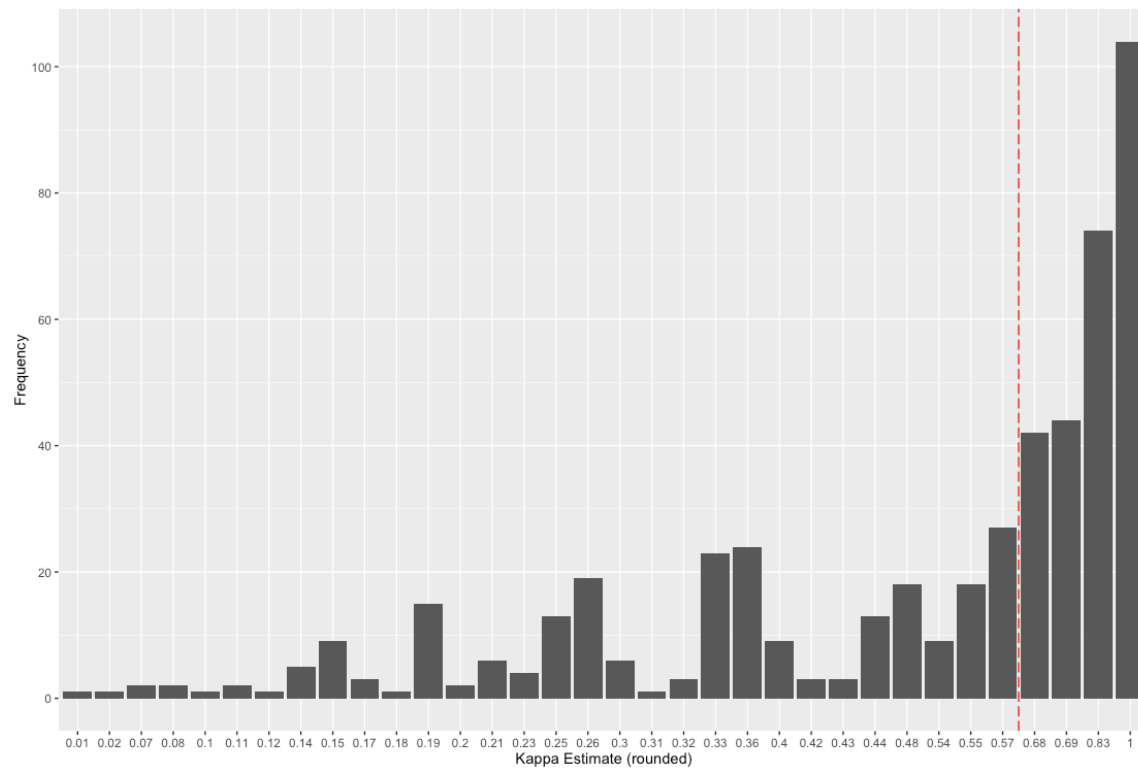
APPENDIX H

CATEGORY DISTRIBUTION AT EACH LEVEL OF HIGH AGREEMENT

Depicted below is the category composition at each level of interrater agreement above κ = .60. The strength of agreement amongst coders was highest for cards classified as Segmented. Put another way, Segmented appeared to be the most readily distinguishable pattern of service utilization across coders.
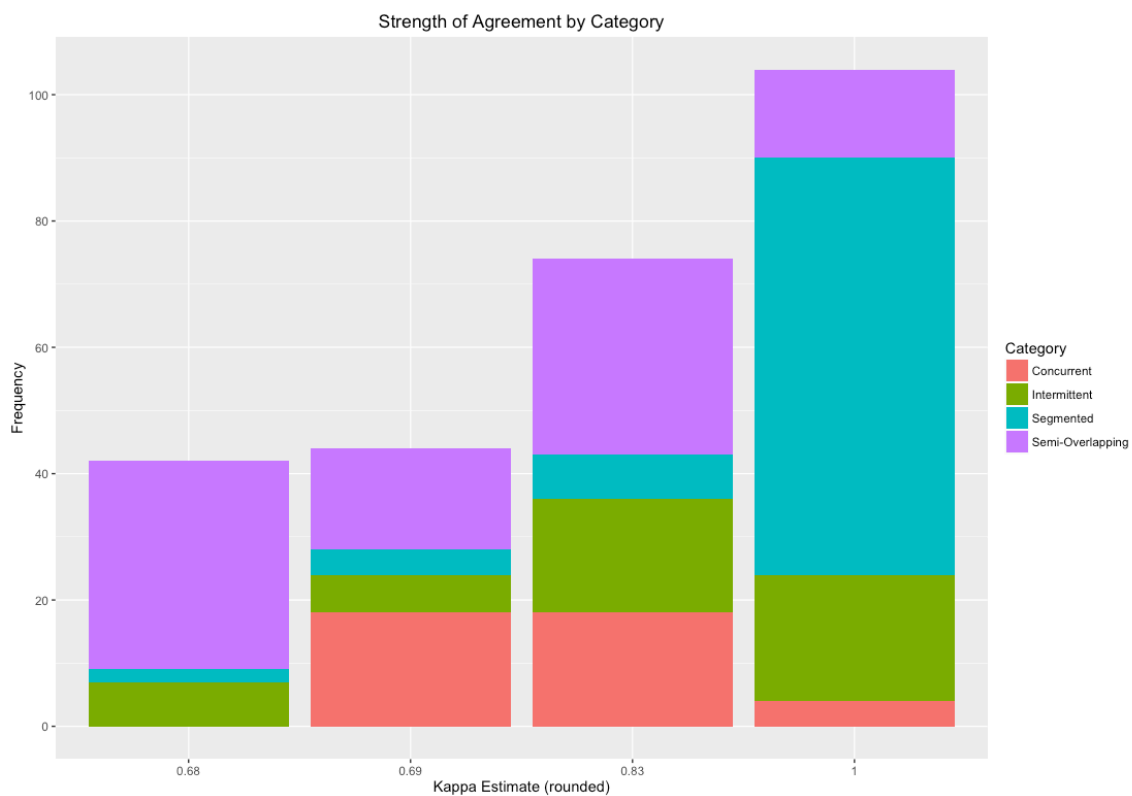
APPENDIX I

PREDICTED LAST SESSION OQ-45 SCORES BY CATEGORY

Depicted below are the last session observed (represented in black) and predicted (represented in red) OQ scores. The predicted scores were derived from a linear regression model that controlled for baseline/first session score. A line in red provides the 95% confidence interval around each predicted score. Results suggest the absence of a significant main effect of category on last session score.
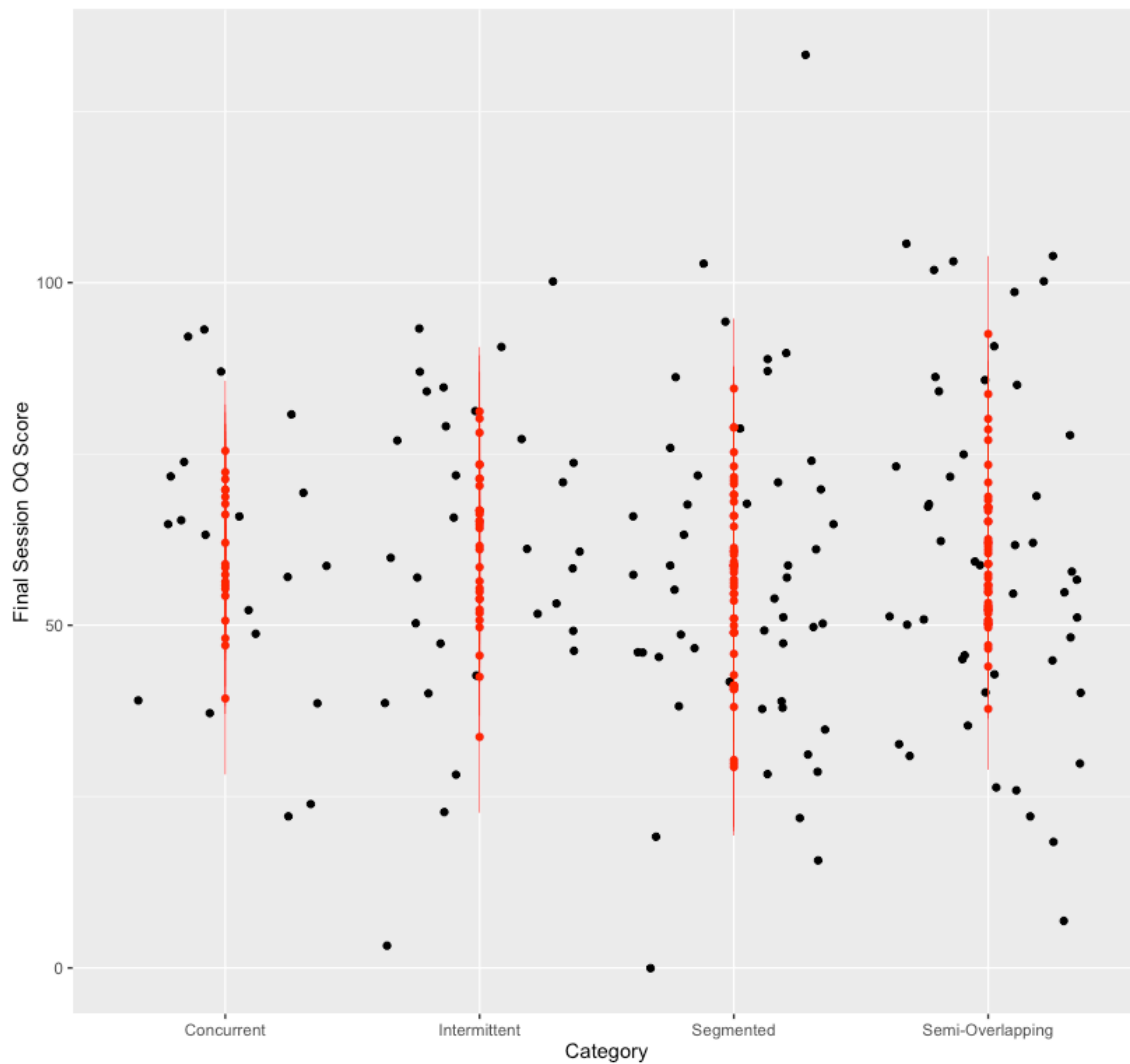
APPENDIX J

ORDINARY LEAST SQUARES REGRESSION ANALYSIS BY CATEGORY

Depicted below are the ordinary least squares (OLS) mean trajectories at each category level for cards with interrater agreement above $\kappa = .60$. Each colored line represents the observed OQ scores for a unique episode of treatment while the black line represents the computed the mean. Some OQ scores are not represented, as they were missing from the sample data set.
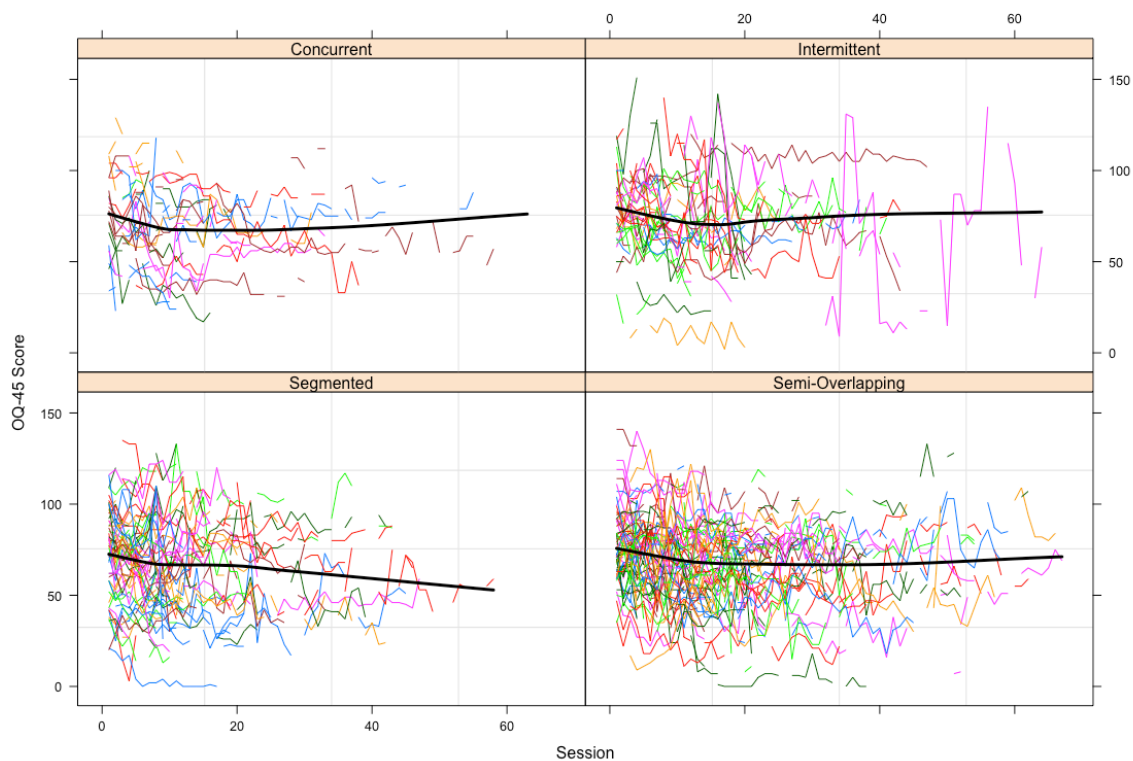
REFERENCES

Akerelrea, C. A., & Zimmerman, D. E. (2002). A group card sorting methodology for developing informational web sites. In *Proceedings of the International Professional Communication Conference* (pp. 437–445). Portland, OR: IEEE Professional Communication Society.

Amaranto, E., & Bender, S., (1990). Individual psychotherapy as an adjunct to group psychotherapy. *International Journal of Group Psychotherapy, 40*(1), 91–101.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

American Psychological Association. (2012). *Resolution on the recognition of psychotherapy effectiveness*. American Psychological Association. Retrieved from http://www.apa.org/news/press/releases/2012/ 08/resolution-psychotherapy.aspx

Anderson, A. (1968). Group counseling. *Review of Educational Research, 33*, 209–226.

Arkowitz, H. (1992). Integrative theories of therapy. In D. K. Freedheim (Ed.), *History of psychotherapy: A century of change* (pp. 261–303). Washington, DC: American Psychological Association.

Babcock, J. C., Green, C. E., & Robie, C. (2004). Does batterers' treatment work? A meta-analytic review of domestic violence treatment. *Clinical Psychology Review, 23*(8), 1023–1053.

Baehr, G. O. (1954). The comparative effectiveness of individual psychotherapy, group psychotherapy, and a combination of these methods. *Journal of Consulting Psychology, 18*(3), 179–183.

Baldwin S. A., Berkeljon, A., Atkins, D. C., Olsen, J. A., & Nielsen, S. L. (2009). Rates of change in naturalistic psychotherapy: Contrasting dose-effect and good-enough level models of change. *Journal of Consulting and Clinical Psychology, 77*(2), 203–211.

Baldwin, S. A., Murray, D. M., & Shadish, W. R. (2005). Empirically supported treatments or type I errors?: Problems with the analysis of data from group-administered treatments. *Journal of Consulting and Clinical Psychology, 73*, 924–935.

Barcikowski, R. S. (1981). Statistical power with group mean as the unit of analysis. *Journal of Educational Statistics, 6,* 267–285.

Barkham, M., Connell, J., Stiles, W. B., Miles, J. N., Margison, F., Evans, C., & Mellor-Clark, J. (2006). Dose–effect relations and responsive regulation of treatment duration: The good enough level. *Journal of Consulting and Clinical Psychology, 74*, 160–167.

Barlow, S. H., Burlingame, G. M., & Fuhriman, A. (2000). Therapeutic applications of groups: From Pratt's "thought control classes" to modern group psychotherapy. *Group Dynamics: Theory, Research, and Practice, 4*(1), 115–134.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2016). lme4: Linear mixed-effects models using 'eigen' and S4. R package version 1.1-12. http://CRAN.R-project.org/package=lme4

Beckstead, D. J., Hatch, A. L., Lambert, M. J., Eggett, D. L., Goates, M. K., & Vermeersch, D. A. (2003). Clinical significance of the Outcome Questionnaire (OQ 45.2). *The Behavior Analyst Today, 4*, 86–90.

Bednar, R., & Kaul, T. (1978). Experiential group research: Current perspectives. In S. L. Garfield & A. E. Bergin (Eds.), *Handbook of psychotherapy and behavior change* (3rd ed., pp. 769–815). New York: Wiley & Sons.

Ben-Porath, D. D., Peterson, G. A., & Smee, J. (2004). Treatment of individuals with borderline personality disorder using dialectical behavior therapy in a community mental health setting. *Cognitive and Behavioral Practice, 11*, 424–434.

Berg, E. A. (1948). A simple objective for measuring flexibility in thinking. *Journal of General Psychology, 39*, 15–22.

Bergstrom, J. A. (1893). Experiments upon physiological memory. *American Journal of Psychology, 5*, 356–359.

Bergstrom, J. A. (1894). The relation of the interference of the practice effect of an association. *American Journal of Psychology, 6*, 433–442.

Bernard, H. R., & Ryan, G. W. (2009). *Analyzing qualitative data: Systematic approaches.* Thousand Oaks, CA: Sage.

Beutler, L. E. (1983). *Eclectic psychotherapy: A systematic approach.* New York: Pergamon Press.

Beutler, L. E., & Clarkin, J. F. (1990). *Systematic treatment selection: Toward targeted therapeutic interventions.* New York: Brunner/Mazel.

Blashfield, R. K. (1980). Propositions regarding the use of cluster analysis in clinical research. *Journal of Consulting and Clinical Psychology, 48*, 456–459.

Bostwick, G. J. (1987). "Where's Mary?": A review of the group treatment dropout literature. *Social Work with Groups, 10*, 117–132.

Budman, S. H., Demby, A., Renondo, J. P., Hannan, M., Feldstein, M., Ring, J., et al. (1988). Comparative outcome in time-limited individual and group psychotherapy. *International Journal of Group Psychotherapy, 38*(1), 63–86.

Burlingame, G., & Baldwin, S. (2011). Group therapy. In J. Norcross, G. VandenBos, & D. Freedheim (Eds.), *History of psychotherapy: Continuity and change* (2nd ed., pp. 505–515). Washington DC: American Psychological Association.

Burlingame, G. M., & Krogel, J. (2005). Relative efficacy of individual versus group psychotherapy. *International Journal of Group Psychotherapy, 55*(4), 607–611.

Burlingame, G. M., Fuhriman, A., & Mosier, J. (2003). The differential effectiveness of group psychotherapy: A meta-analytic perspective. *Group Dynamics: Theory, Research, and Practice, 7*(1), 3–12.

Burlingame, G. M., Gleave, R., Erekson, D., Nelson, P. L., Olsen, J., Thayer, S., & Beecher, M. (2015). Differential effectiveness of group, individual, and conjoint treatments: An archival analysis of OQ-45 change trajectories. *Psychotherapy Research.* doi: 10.1080/10503307.2015.1044583

Burlingame, G. M., MacKenzie, K. R., & Strauss, B. (2004). Small group treatment: Evidence for effectiveness and mechanisms of change. In M. J. Lambert (Ed.), *Bergin & Garfield's handbook of psychotherapy and behavior change* (5th ed., pp. 647–696). New York, NY: Wiley.

Burlingame, G., Fuhriman, A., & Johnson, J. (2004). Process and outcome in group psycho-therapy: A perspective. In J. DeLucia-Waack, C. Kalodner, & M. Riva (Eds.), *Handbook of group work* (pp. 49–61). Thousand Oaks, CA: Sage.

Burlingame, G., Strauss, B., & Joyce, A. (2013). Change mechanisms and effectiveness of small group treatments. In M. J. Lambert (Ed.), *Bergin & Garfield's handbook of psychotherapy and behavior change* (6th ed., pp. 640–689). New York: Wiley & Sons.

Burrow, T. (1928). The basis of group-analysis, or the analysis of reactions of normal and neurotic individuals. *British Journal of Medical Psychology, 8*, 198–206.

Byrt, T., Bishop, J., & Carlin, J. B. (1993). Bias, prevalence and kappa. *Journal of Clinical Epidemiology, 46*(5), 423–429.

Castonguay, L. G., Eubanks, C. F., Goldfried, M. R., Muran, J. C., & Lutz, W. (2015). Research on psychotherapy integration: Building on the past, looking to the future. *Psychotherapy Research, 25*(3), 365–382.

Castonguay, L. G., Newman, M. G., Borkovec, T. D., Holtforth, M. G., & Maramba, G. G. (2005). Cognitive-behavioral assimilative integration. In J. C. Norcross & M. R. Goldfried (Eds.), *Handbook of psychotherapy integration* (2nd ed., pp. 241–260). New York: Oxford University Press.

Center for Collegiate Mental Health. (2012). *CCAPS 2012 Technical Manual*. University Park, PA: Author.

Chambless, D. L., Sanderson, W. C., Shoham, V., Bennett Johnson, S., Pope, K. S., Crits-Christoph, P., et al. (1996). An update on empirically validated therapies. *The Clinical Psychologist, 49*, 5–18.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*(1), 37–46.

Cotton, J. M. (1948). Group psychotherapy: An appraisal. In P. H. Hoch (Ed.), *Failures in psychiatric treatment* (pp. 121–128). Oxford, England: Grune & Stratton.

Coxon, A. (1999). *Sorting data: Collection and analysis.* In *Sage University Papers Series on Qualitative Applications in the Social Sciences* (series no. 07-127). Thousand Oaks, CA: Sage.

Crits-Christoph, P., Siqueland, L., Blaine, J., Frank, A., Luborsky, L., Onken, L., et al. (1999). Psychosocial treatments for cocaine dependence. *Archives of General Psychiatry, 56,* 493–502.

Cunningham, J. M., & Matthews, K. L. (1982). Impact of multiple-family therapy approach on a parallel latency-age/parent group. *International Journal of Group Psychotherapy, 32*, 91–102.

Cutts, L. (2011). Integration in counselling psychology: To what purpose. *Counselling Psychology Review, 26*(2), 238–248.

Davies, D. R., Burlingame, G. M., Johnson, J. E., Barlow, S. H., & Gleave, R. L. (2008). The effects of a feedback intervention on group process and outcome. *Group Dynamics: Theory, Research, and Practice, 12*(2), 141–154.

Dearborn, G. V. (1898). A study of imagination. *American Journal Psychology, 5*, 183–190.

Dearholt, D., McDonald, J., Paap, K., & Schvaneveldt, W. (1986). A formal interface design methodology based on user knowledge. In *Proceedings of the 1986 SIGCHI Conference on Human Factors in Computing Systems* (pp. 285–290). Boston: ACM Press.

Deaton, M. (2002). Sorting techniques for user-centered information design. Retrieved from http://www.mmdeaton.com/SortingTechniquesforInformationDesign.doc

DeLucia-Waack, J. L., Kalodner, C. R., & Riva, M. T. (Eds.) (2013). *Handbook of group counseling and psychotherapy* (2nd ed.). Thousand Oaks, CA: Sage.

Delucia-Waack, J., Gerrity, D., Kalodner, C., & Riva, M. (2004). *Handbook of group counseling and psychotherapy.* Thousand Oaks, CA: Sage.

Derogatis, L. R. (1992). *SCL-90-R administration, scoring, & procedures manual-II*. Townson, MD: Clinical Psychometric Research, Inc.

Di Eugenio, B., & Glass, M. (2004). The kappa statistic: A second look. *Computational linguistics, 30*(1), 95–101.

Dreikurs, R. (1959). Early experiments with group psychotherapy. *American Journal of Psychotherapy, 13*, 882–891.

Drury, V., Birchwood, M., Cochrane, R., & Macmillan, F. (1996). Cognitive therapy and recovery from acute psychosis: A controlled trial. I. Impact on psychotic symptoms. *British Journal of Psychiatry, 169*, 593–601.

Dubois, C. (1949). The card-sorting or psychophysical interview. *Public Opinion Quarterly, 13*, 619–628.

Dush, D. M., Hirt, M. L., & Schroeder, H. (1983). Self-statement modification with adults: A meta-analysis. *Journal of Consulting and Clinical Psychology, 94*, 408–422.

Duszak, E. M. P. (2014). Evaluating the effectiveness of counseling center-based psychotherapy outcome measures: A statistical comparison of the counseling center assessment of psychological symptoms and the outcome questionnaire. Unpublished doctoral dissertation, University of Utah.

Eling, P., Derckx, K., & Maes, R. (2008). On the historical and conceptual background of the Wisconsin Card Sorting Test. *Brain and Cognition, 67*(3), 247–253.

Elkin, I., Shea, M. T., Watkins, J. T., Imber, S. D., Sotsky, S. M., Collins, J. F., et al. (1989). National Institute of Mental Health Treatment of Depression Collaborative Research Program: General effectiveness of treatments. *Archives of General Psychiatry, 46*, 971–982.

Emrick, C. (1975). A review of psychologically oriented treatment of alcoholism. *Journal for the Study of Alcoholism, 36*(1), 88–108.

Erickson, R. C. (1982). Inpatient small group psychotherapy: A survey. *Clinical Psychology Review, 2*, 137–151.

Eysenck, H. J. (1970). Explanation and the concept of personality. In R. Borger & F. Cioffi (Eds.), *Explanation in the behavioural sciences* (pp. 387–410). Cambridge, UK: Cambridge University Press.

Fairweather, G. W., Simon, R., Gebhard, M. E., Weingarten, E., Holland, J. L., Sanders, R., et al. (1960). Relative effectiveness of psychotherapeutic programs: A multicriteria comparison of four programs for three different patient groups. *Psychological Monographs: General and Applied, 74*(5), 1–26.

Fettes, P. A., & Peters, J. M. (1992). A meta-analysis of group treatments for bulimia nervosa. *International Journal of Eating Disorders, 11*(2), 97–110.

Frank, J. (1973). *Persuasion and healing* (2nd ed.). New York: Schocken.

Frank, J. D., & Frank, J. B. (1991) *Persuasion and healing: A comparative study of psychotherapy* (3rd ed.). Baltimore, MD: Johns Hopkins University Press.

Frank, K. R. (1999). *Psychoanalytic participation: Action, interaction, and integration.* New York: Analytic Press.

Freeman, C. P. L., & Munro, K. M. (1988). Drug and group treatments for bulimia/bulimia nervosa. *Journal of Psychosomatic Research, 32*, 647–660.

French, T. M. (1933). Interrelations between psychoanalysis and the experimental work of Pavlov. *American Journal of Psychiatry, 89*(6), 1165–1203.

Freud, S. (1922). *Group psychology and the analysis of the ego.* London & Vienna: International Psychoanalytic Press.

Fuhriman, A., & Burlingame, G. (1990). Consistency of matter: A comparative analysis of individual and group process variables. *Counseling Psychologist, 18*(1), 6–63.

Fuhriman, A., & Burlingame, G. (1994). Group psychotherapy: Research and practice. In A. Fuhriman & G. Burlingame (Eds.), *Handbook of group psychotherapy: An empirical and clinical synthesis* (pp. 3–40). New York, NY: Wiley.

Fuhriman, A., Drescher, S., & Burlingame, G. (1984). Conceptualizing small group process. *Small Group Research, 15*(4), 427–440.

Gallagher, R. P. (2015). *National Survey of Counseling Centers 2014*: International Association of Counseling Services, Inc.

Galton, F. (1891). The patterns in thumb and finger marks. *Philosophical Transactions of the Royal Society, B, 182*, 1–23.

Gamer, M., Lemon, J., Fellows, I., & Singh, P. (2012). irr: Various coefficients of interrater reliability and agreement. R package version 0.84. http://CRAN.R-project.org/package=irr

Garfield, S. L., & Kurtz, R. (1975). The clinical psychologists: A survey of selected attitudes and views. *The Clinical Psychologist, 28*(3), 4–7.

Garson, G. D. (2014). *Multiple regression*. Asheboro, NC: Statistical Associates Publishers.

Gay, P. (1988). *Freud: A life for our time.* New York, NY: Norton.

Glaser, B. G. (1978). *Theoretical sensitivity: Advances in the methodology of grounded theory.* Mill Valley, CA: Sociology Press.

Golden, B. R., Corazzini, J. G., & Grady, P. (1993). Current practice of group therapy at university counseling centers: A national survey. *Professional Psychology: Research and Practice, 24*(2), 228–230.

Goldfried, M. R. (2013). What should we expect from psychotherapy? *Clinical Psychology Review, 33*(7), 862–869.

Goldfried, M. R., Pachankis, J. E., & Bell, A. C. (2005). A history of psychotherapy integration. In J. C. Norcross & M. R. Goldfried (Eds.), *Handbook of psychotherapy integration* (2nd ed., pp. 24–60). New York: Oxford University Press.

Gwet, K. (2002). Kappa statistic is not satisfactory for assessing the extent of agreement between raters. *Statistical Methods for Inter-rater Reliability Assessment, 1*(6), 1–6.

Halgren, S. L., & Cooke, N. J. (1993). Towards ecological validity in menu research. *International Journal of Man-Machine Studies, 39*, 51–71.

Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in Quantitative Methods for Psychology, 8*(1), 23–34.

Halmi, K. A. (2005). The multimodal treatment of eating disorders. *World Psychiatry, 4*(2), 69–73.

Hannah, S. (2005). Sorting out card sorting: Comparing methods for information architects, usability specialists, and other practitioners. Unpublished master's thesis, University of Oregon.

Hansen, N. B., Lambert, M. J., & Forman, E. M. (2002). The psychotherapy dose-response effect and its implications for treatment delivery services. *Clinical Psychology: Science and Practice, 9*(3), 329–343.

Hoag, M. J., & Burlingame, G. M. (1997). Evaluating the effectiveness of child and adolescent group treatment: A meta-analytic review. *Journal of Clinical Child Psychology, 26*(3), 234–246.

Horowitz, L. M., Rosenberg, S. E., Baer, B. A., Ureno, G., & Villasenor, V. S. (1988). Inventory of interpersonal problems: Psychometric properties and clinical applications. *Journal of Consulting and Clinical Psychology, 56*, 885–892.

Hox, J. J. (2010). *Multilevel analysis: Techniques and applications* (2nd ed.). New York: Routledge.

Hudson, W. (2013). Card sorting. In M. Soegaard & R. Dam (Eds.), *Encyclopedia of human-computer interaction* (2nd ed.). Aarhus, Denmark: The Interaction Design Foundation.

Imel, Z., & Wampold, B. (2008). The importance of treatment and the science of common factors in psychotherapy. In S. D. Brown & R. W. Lent (Eds.), *Handbook of counseling psychology* (4th ed., pp. 249–262). Hoboken, NJ: John Wiley & Sons.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*, 12–19.

Jastrow, J. (1886). An easy method of measuring the time of mental processes. *Science, 8*, 237–241.

Jastrow, J. (1898). A sorting apparatus for the study of reaction-times. *Psychological Review, 5*(3), 279–285.

Kadera, S. W., Lambert, M. J., & Andrews, A. A. (1996). How much therapy is really enough? A session-by-session analysis of the psychotherapy dose effect relationship. *Journal of Psychotherapy Practice and Research, 5*(2), 132–151.

Kadis, A. L., & Markowitz, M. (1958). Group psychotherapy. In D. Brower & L. E. Abt (Eds.), *Progress in Clinical Psychology, Vol. III* (pp. 154–183). New York: Grune & Stratton.

Karterud, S., Johansen, M. S., & Wilberg, T., (2007). Conjoint group and individual psychotherapy in a research trial for patients with severe personality disorders. *Group, 31*(1), 31–46.

Kaul, T., & Bednar, R. (1986). Experimental group research: Results, questions & suggestions. In S. Garfield & A. Bergin (Eds.), *Handbook of psychotherapy and behavior change* (pp. 155–188). New York: Wiley.

Kazdin, A. E., & Bass, D. (1989). Power to detect differences between alternative treatments in comparative psychotherapy outcome research. *Journal of Consulting and Clinical Psychology, 57*, 138–147.

Kearney, L., & Baron, A. (2003). *Highlights of the research consortium 2002 non-clinical sample study.* Austin, TX: Counseling and Mental Health Center, University of Texas at Austin.

Klapman, J. W. (1946). *Group psychotherapy*. New York: Grone & Stratton.

Kliem, S., Kroger, C., & Kosfelder, J. (2010). Dialectical behavior therapy for borderline personality disorder. *Journal of Consulting and Clinical Psychology, 78*, 936–951.

Kline, L. W., & Kellogg, C. E. (1914). Cards as psychological apparatus. *Science, 39*(1009), 657–659.

Koerner, K., & Linehan, M. M. (2000). Research on dialectical behavior therapy for patients with borderline personality disorder. *Psychiatric Clinics of North America, 23*, 151–167.

Koons, C. R., Robins, C. J., Tweed, J. L., Lynch, T. R., Gonzalez, A. M., Morse, J. Q., et al. (2001). Efficacy of dialectical behavior therapy in women veterans with borderline personality disorder. *Behavior Therapy, 32*, 371–390.

Krippendorff, K. (1980). *Content analysis: An introduction to its methodology*. Beverly Hills, CA: Sage.

Lambert, M. J., Lunnen, K., Umphress, V., Hansen, N., & Burlingame, G. M. (1994). *Administration and scoring manual for the Outcome Questionnaire (OQ-45.1)*. Salt Lake City, UT: IHC Center for Behavioral Healthcare Efficacy.

Lambert, M. J., Morton, J. J., Hatfield, D., Harmon, C., Hamilton, S., Reid, R. C., et al. (2004). *Administration and scoring manual for the Outcome Questionnaire-45.2*. Orem, UT: American Professional Credentialing Services.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159–174.

Lazarus, A. A. (1989). *The practice of multimodal therapy.* Baltimore, MD: Johns Hopkins University Press.

Lazarus, A. A. (1997). *Brief but comprehensive psychotherapy: The multimodal way.* New York: Springer.

Lazarus, A. A. (2005). Multimodal therapy. In J. C. Norcross & M. R. Goldfried (Eds.) *Handbook of psychotherapy integration* (2nd ed., pp. 105–120). New York: Oxford.

Lazarus, A. A., Beutler, L. E., & Norcross, J. C. (1992). The future of technical eclecticism. *Psychotherapy, 29*(1), 11–20.

Lazell, E. W. (1921). The group treatment of dementia praecox. *Psychoanalytic Review, 8*, 168–179.

Lesieur, H. R., & Blume, S. B. (1991) Evaluation of patients treated for pathological gambling in a combined alcohol, substance abuse and pathological gambling treatment unit using the Addiction Severity Index. *British Journal of Addiction, 86*, 1017–1028.

Lieberman, M. (1976). Change induction in small groups. *Annual Review of Psychology, 27*, 217–250.

Lieberman, M. A., Lakin, M., & Whitaker, D. S. (1968). The group as a unique context for therapy. *Psychotherapy: Theory, Research, and Practice, 5*(1), 29–36.

Light, R. J. (1971). Measures of response agreement for qualitative data: Some generalizations and alternatives. *Psychological Bulletin, 76*(5), 365–377.

Linehan M. M., Comtois, K. A., Murray, A. M., Brown, M. Z., Gallop, R. J., Heard, H. L., et al. (2006). Two-year randomized controlled trial and follow-up of dialectical behavior therapy vs therapy by experts for suicidal behaviors and borderline personality disorder. *Archives of General Psychiatry, 63*(7), 757–766.

Linehan, M. M. (1993). *Cognitive-behavioral treatment of borderline personality disorder.* New York: Guilford Press.

Linehan, M. M., Armstrong, H. E., Suarez, A., Allmon, D., & Hearld, H. L. (1991). Cognitive-behavioral treatment of chronically parasuicidal borderline patients. *Archives of General Psychiatry, 48*, 1060–1064.

Linehan, M. M., Dimeff, L. A., Reynolds, S. K., Comtois, K. A., Welch, S. S., Heagerty, P., & Kivlahan, D. R. (2002). Dialectical behavior therapy versus comprehensive validation therapy plus 12-step for the treatment of opioid dependent women

meeting criteria for borderline personality disorder. *Drug and Alcohol Dependence, 67*(1), 13–26.

Linehan, M. M., Korslund, K. E., Harned, M. S., Gallop, R. J., Lungu, A., Neacsiu, A. D., et al. (2015). Dialectical behavior therapy for high suicide risk in individuals with borderline personality disorder: A randomized clinical trial and component analysis. *JAMA Psychiatry, 72*(5), 475–482.

Linehan, M. M., Schmidt, H., Dimeff, L. A., Craft, J. C., Kanter, J., & Comtois, K. A. (1999). Dialectical behavior therapy for patients with borderline personality disorder and drug-dependence. *American Journal on Addictions, 8*(4), 279–292.

Lipsius, S. H. (1991). Combined individual and group psychotherapy: Guidelines at the interface. *International Journal of Group Psychotherapy, 41*(3), 313–327.

Lord, F. M. (1959). Statistical inferences about true scores. *Psychometrika, 24*(1), 1–17.

Luborsky, L., Singer, B., & Luborsky, L. (1975). Comparative studies of psychotherapies: Is it true that "Everyone has won and all must have prizes"? *Archive of General Psychiatry, 32*, 995–1008.

Luchins, A. S. (1947). Methods of studying the progress and outcomes of a group psychotherapy program. *Journal of Consulting Psychology, 11*(4), 173–183.

Lunnen, K. M., & Ogles, B. M. (1998). A multiperspective, multivariable evaluation of reliable change. *Journal of Consulting and Clinical Psychology, 66*(2), 400–410.

Lutz, W., Lowry, J., Kopta, S. M., Einstein, D. A., & Howard, K. I. (2001). Prediction of dose–response relations based on patient characteristics. *Journal of Clinical Psychology, 57,* 889–900.

Lutz, W., Martinovich, Z., & Howard, K. I. (1999). Patient profiling: An application of random coefficient regression models to depicting the response of a patient to outpatient psychotherapy. *Journal of Consulting and Clinical Psychology, 67,* 571–577.

Lutz, W., Martinovich, Z., Howard, K. I., & Leon, S. C. (2002). Outcomes management, expected treatment response, and severity-adjusted provider profiling in outpatient psychotherapy. *Journal of Clinical Psychology, 58*(10), 1291–1304.

Lynch, T. R., Morse, J. Q., Mendelson, T., & Robins, C. J. (2003). Dialectical behavior therapy for depressed older adults: A randomized pilot study. *American Journal of Geriatric Psychiatry, 11*(1), 33–45.

Mann, J. (1966). Evaluation of group psychotherapy. In J. Moreno (Ed.), *The international handbook of group psychotherapy.* New York: Philosophical Library.

Markowitz, J. C., & Milrod, B. L. (2015). What to do when a psychotherapy fails. *The Lancet Psychiatry, 2*(2), 186–190.

Marsh, L. C. (1931). Group treatment for the psychoses by the psychological equivalent of the revival. *Mental Hygiene, 15*, 328–349.

Marsh, L. C. (1933). An experiment in group treatment of patients at Worchester State Hospital. *Mental Hygiene, 17*, 396–416.

Marx, K., & Engels, F. (1848/1992). *The communist manifesto.* Oxford, UK: Oxford University Press.

Maurer, D., & Warfel, T. (2004). Card sorting: A definitive guide. Retrieved from http://www.boxesandarrows.com/view/card_sorting_ a_definitive_guide

McLeod, J. (2009). *An introduction to counseling* (4th ed.). Maidenhead: Open University Press.

McMain, S. F., Links, P. S., Gnam, W. H., Guimond, T., Cardish, R. J., Korman, L., & Streiner, D. L. (2009). A randomized trial of dialectical behavior therapy versus general psychiatric management for borderline personality disorder. *American Journal of Psychiatry, 166*, 1365–1374.

McQuillan, A., Nicastro, R., Guenot, F., Girard, M., Lissner, C., & Ferrero, F. (2005). Intensive dialectical behavior therapy for outpatients with borderline personality disorder who are in crisis. *Psychiatric Services, 56*(2), 193-197.

McRoberts, C., Burlingame, G. M., & Hoag, M. J. (1998). Comparative efficacy of individual and group psychotherapy: A meta-analytic perspective. *Group Dynamics: Theory, Research, and Practice, 2*(2), 101–117.

Melling, A. K. (2014). Severity and the utilization of services in a university counseling center. Unpublished doctoral dissertation, University of Utah.

Meltzoff, J., & Kornreich, M. (1970). *Research in psychotherapy.* New York: Atherton Press.

Mickelson, B. K. (2008). Client change in multi-model treatment: A comparison of change trajectories in group, individual, and conjoint formats in a counseling center. Unpublished doctoral dissertation, Brigham Young University.

Miller, G. A. (1969). A psychological method to investigate verbal concepts. *Journal of Mathematical Psychology, 6*, 169–191.

Miller, R. C., & Berman, J. S. (1983). The efficacy of cognitive behavior therapies: A qualitative review of the research evidence. *Psychological Bulletin, 94*, 39–53.

Miller, S. D., Duncan, B. L., Sorrell, R., & Brown, J. (2005). The partners for change outcome management system. *Journal of Clinical Psychology, 61,* 199–208.

Minami, T., Davies, D. R., Tierney, S. C., Bettmann, J. E., McAward, S. M., Averill, L. A., et al. (2009). Preliminary evidence on the effectiveness of psychological treatments delivered at a university counseling center. *Journal of Counseling Psychology, 56*(2), 309–320.

Moreno, J. L., & Whitin, E. (1932). *Application of the group method to classification.* New York: National Commission on Prison and Prison Labor.

Nakhimovsky, Y., Schusteritsch, R., & Rodden, K. (2006). Scaling the card sort method to over 500 items: Restructuring the Google AdWords Help Center. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (pp.183–188). Montreal, Quebec: ACM Press.

Nielsen, S. L., Bailey, R. J., Nielsen, D., & Pedersen, T. R. (2016). Dose response and shape of change. In S. Maltzman (Ed.), *Oxford handbook of treatment processes and outcomes in psychology: A multidisciplinary, biopsychosocial approach* (pp. 465–497). New York: Oxford University Press.

Nietzel, M. T, Russell, R. L., Hemmings, K. A., & Gretter, M. L. (1987). Clinical significance of psychotherapy for unipolar depression: A meta-analytic approach to social comparison. *Journal of Consulting and Clinical Psychology, 55*, 156–161.

Norcross, J. C. (1990). An eclectic definition of psychotherapy. In J. K. Zeig & W. M. Munion (Eds.), *What is psychotherapy? Contemporary perspectives* (pp. 218–220). San Francisco, CA: Jossey-Bass.

Norcross, J. C. (2005). A primer on psychotherapy integration. In J. C. Norcross & M. R. Goldfried (Eds.), *Handbook of psychotherapy integration* (2nd ed., pp. 3–23). New York: Oxford University Press.

Norcross, J. C., & Goldfried, M. R. (Eds.). (2005). *Handbook of psychotherapy integration* (2nd ed.). New York: Oxford.

Norcross, J. C., & Halgin, R. P. (2005). Training in psychotherapy integration. In J. C. Norcross & M. R. Goldfried (Eds.), *Handbook of psychotherapy integration* (2nd ed., pp. 439–458). New York: Oxford University Press.

Norcross, J. C., & Napolitano, G. (1986). Defining our journal and ourselves. *Journal of Integrative & Eclectic Psychotherapy, 5*(3), 249–255.

Novick, M. R. (1966). The axioms and principle results of classical test theory. *Journal of Mathematical Psychology, 3*, 1–18.

Orlinsky, D. E., Grawe, K., & Parks, B. K. (1994). Process and outcome in psychotherapy–noch einmal. In A. Bergin & S. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed., pp. 270–378). New York: Wiley.

Palmer, S., & Woolfe, R. (Eds.) (1999). *Integrative and eclectic counselling and psychotherapy.* London: Sage.

Panas, L., Caspi, Y., Fournier, E., & McCarty, D. (2003) Performance measures for outpatient substance abuse services: Group versus individual counseling. *Journal of Substance Abuse Treatment, 25*, 271–278.

Parloff, M.B., & Dies, R. R. (1977). Group psychotherapy outcome research 1966–1975. *International Journal of Group Psychotherapy, 27*(3), 281–319.

Patterson, C. H. (1980). *Theories of counseling and psychotherapy* (3rd ed.). New York: Harper & Row.

Patterson, C. H. (1989). Eclecticism in psychotherapy: Is integration possible? *Psychotherapy: Theory, Research, Practice, Training, 26*(2), 157–161.

Patterson, C. H. (1990). On being client-centered. *Person Centered Review, 5*, 425–432.

Pattison, E. (1965). Evaluation studies of group psychotherapy. *International Journal of Group Psychotherapy, 15*(3), 382–397.

Pilkonis, P.A., Imber, S. D., Lewis, P., & Rubinsky, P. (1984). A comparative outcome study of individual, group, and conjoint psychotherapy. *Archive of General Psychiatry, 41*, 431–437.

Porter, K. (1993). Combined individual and group psychotherapy. In H. I. Kaplan & B. I. Sadock (Eds.), *Comprehensive group psychotherapy* (pp. 314–324). Baltimore, MD: Williams & Wilkins.

Pratt, J. H. (1945). Group method in the treatment of psychosomatic disorders. *Sociometry, 8*, 323–331.

Prochaska, J. O., & DiClemente, C. C. (1984). *The transtheoretical approach: Crossing traditional boundaries of therapy.* Homewood, IL: Dorsey.

Prochaska, J. O., & DiClemente, C. C. (2005). The transtheoretical approach. In J. C. Norcross, & M. R. Goldfried (Eds.), *Handbook of psychotherapy integration* (2nd ed., pp. 147–171). New York: Oxford University Press.

R Development Core Team. (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/

Raudenbush, S. W., & Byrk, A. S. (2002) *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Reetz, D., Bershad, C., LeViness, P., & Whitlock, M. (2016). *The association for university and college counseling center directors annual survey*. Retrieved from http://www.aucccd.org/assets/documents/aucccd%202016%20monograph%20-%20public.pdf

Rickard, H. (1962). Selected group psychotherapy evaluation studies. *Journal of General Psychology, 67*, 35–50.

Righi, C., James, J., Beasley, M., Day, D. L., Fox, J. E., Gieber, J., et al. (2003). Card sort analysis best practices. *Journal of Usability Studies, 8*(3), 69–89.

Robinson, L. A., Berman, J. S., & Neimeyer, R. A. (1990) Psychotherapy for the treatment of depression: a comprehensive review of controlled outcome research. *Psychological Bulletin, 108*(1), 30–49.

Romesburg, H. C. (2004). *Cluster analysis for researchers.* Morrisville, NC: Lulu Press.

Rosenzweig, S. (1936). Some implicit common factors in diverse methods of psychotherapy. *American Journal of Orthopsychiatry, 6*, 412–415.

Ruitenbeek, H. M. (1969). *Group therapy today: Styles, methods, and techniques.* Oxford, England: Atherton.

Rutan, J. S., & Alonso, A. (1982). Group therapy, individual therapy, or both? *International Journal of Group Psychotherapy, 32*, 267–282.

Ryle, A. (1990). *Cognitive analytic therapy: Active participation in change.* London: John Wiley & Sons.

Ryle, A. (2005). Cognitive analytic therapy. In J. C. Norcross & M. R. Goldfried (Eds.), *Handbook of psychotherapy integration* (2nd ed., pp. 196–220). New York: Oxford University Press.

Safer, D. L., Telch, C. F., & Agras, W. S. (2001). Dialectical behavior therapy for bulimia nervosa. *American Journal of Psychiatry, 158*(4), 632–634.

Safran, J. D. (1998). *Widening the scope of cognitive therapy: The therapeutic relationship, emotion, and the process of change.* Northvale, NJ: Jason Aronson, Inc.

Safran, J. D., & Segal, Z. V. (1990). *Interpersonal process in cognitive therapy.* New York: Basic Books.

Sager, C. J. (1960). Combined individual and group psychoanalysis: Symposium, 1959: Concurrent individual and group analytic psychotherapy. *American Journal of Orthopsychiatry, 30*(2), 225–241.

Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory and Cognition, 13*, 501–518.

Schermer, V. L. (2009). On the vicissitudes of combining individual and group psychotherapy. *International Journal of Group Psychotherapy, 59*(1), 149–162.

Schilder, P. (1939). Results and problems of group psychotherapy in severe neuroses. *Mental Hygiene, 23*, 87–98.

Schottenbauer, M. A., Glass, C. R., & Arnkoff, D. B. (2007). Decision making and psychotherapy integration: Theoretical considerations, preliminary data, and implications for future research. *Journal of Psychotherapy Integration, 17*(3), 225–250.

Schwartz, K. (2004). Concurrent group and individual psychotherapy in a psychiatric day hospital for depressed elderly. *International Journal of Group Psychotherapy, 54*, 177–201.

Shadish, W. R. (1992). Do family and marital psychotherapies change what people do? A meta-analysis of behavioral outcomes. In T. D. Cook, H. Cooper, D. S. Corday, H. Hartman, L. V. Hedges, R. J. Light, T. A. Louis, & F. Mosteller (Eds.), *Meta-analysis for explanation: A casebook* (pp. 129–208). New York: Sage.

Siegel, S., & Castellan, N. J. (1988). *Nonparametric statistics for the behavioral sciences* (2nd ed.). New York: McGraw-Hill.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence.* New York: Oxford University Press.

Slavson, S. (1940). Group psychotherapy. *Mental Hygiene, 24*, 36–49.

Slavson, S. R. (1943). *An introduction to group therapy*. New York: Commonwealth Fund.

Smith, M. L., Glass, G. V., & Miller, T. L. (1980). *The benefits of psychotherapy*. Baltimore, MD: Johns Hopkins University Press.

Snell, M. N., Mallinckrodt, B., Hill, R. D., & Lambert, M. J. (2001). Predicting counseling center clients' response to counseling: A 1-year follow-up. *Journal of Counseling Psychology, 48*, 463–473.

Speer, D. C., & Newman, F. L. (1996). Mental heath services outcome evaluation. *Clinical Psychology: Science and Practice, 3*, 105–129.

Spencer, D. (2009). *Card sorting: Designing usable categories*. Brooklyn, NY: Rosenfeld Media.

Staats, H. (2010). Combining individual and group therapy in an out-patient setting for patients with personality disorders—useful approach or invitation to acting out?. In A.-M. Schloesser & A. Gerlach (Eds.), *Crossing borders – Integrating differences: Psychoanalytic psychotherapy in transition* (pp. 39–54). London: Karnac.

Steketee, G., Frost, R. O., Wincze, J., Greene, K., & Douglass, H. (2000) Group and individual treatment of compulsive hoarding: A pilot study. *Behavioural and Cognitive Psychotherapy, 28,* 259–268.

Stoffers J. M., Vollm, B. A., Rucker, G., Timmer, A., Huband, N., & Lieb, K. (2012). Psychological therapies for people with borderline personality disorder. *Cochrane Database Systematic Reviews, 8*, Art. No. CD005652.

Stone, G., Vespia, K., & Kanz, J. (2000). How good is mental health care on college campuses? *Journal of Counseling Psychology, 47*(4), 498–510.

Stotsky, B. A., & Zolik, E. S. (1965). Group psychotherapy with psychotics: 1921–1963: A review. *International Journal of Group Psychotherapy, 15*, 321–344.

Stricker, G., & Gold, J. (2005). Assimilative psychodynamic psychotherapy. In J. C. Norcross & M. R. Goldfried (Eds.), *Handbook of psychotherapy integration* (2nd ed., pp. 221–240). New York: Oxford University Press.

Stulz, N., Lutz, W., Leach, C., Lucock, M., & Barkham, M. (2007). Shapes of early change in psychotherapy under routine outpatient conditions. *Journal of Consulting and Clinical Psychology, 75*, 864–874.

Syz, H. (1928). Remarks on group analysis. *American Journal of Psychiatry, 85*, 141–148.

Task Force on Promotion and Dissemination of Psychological Procedures. (1998). Update on empirically supported therapies II. *Clinical Psychologist, 51*, 3–16.

Telch, C. F., Agras, W. S, & Linehan, M. M. (2001). Dialectical behavior therapy for binge eating disorder. *Journal of Consulting and Clinical Psychology, 69*(6), 1061–1065.

Titanium Software, Inc. (2017). *Titanium Schedule*. Retrieved from the Titanium Schedule website http://www.titaniumschedule.com

Tukey, J. W. (1977). *Exploratory data analysis*. Reading, MA: Addison-Wesley.

Turner, R. (2000). Naturalistic evaluation of dialectical behavior therapy-oriented treatment for borderline personality disorder. *Cognitive and Behavioral Practice, 7*, 413–419.

Ulman, K. H. (2002). The ghost in the group room: Countertransferential pressures associated with conjoint individual and group psychotherapy. *International Journal of Group Psychotherapy, 52*, 387–407.

Umphress, V. J., Lambert, M. J., Smart, D. W., Barlow, S. H., & Clouse, G. (1997). Concurrent and construct validity of the Outcome Questionnaire. *Journal of Psychoeducational Assessment, 15*, 40–55.

University of Utah. (2014). University Counseling Center Annual Report July 1, 2013 – June 30, 2014. Retrieved from http://studentaffairs.utah.edu/assessment/annual-reports/13-14%20UCC%20Annual%20Report.pdf

Verheul, R., Van Den Bosch, L. M., Koeter, M. W., De Ridder, M. A., Stijnen, T., & Van Den Brink, W. (2003). Dialectical behaviour therapy for women with borderline personality disorder: 12-month, randomised clinical trial in The Netherlands. *British Journal of Psychiatry, 182*, 135–140.

Vermeersch, D. A., Lambert, M. J., & Burlingame, G. M. (2000). Outcome Questionnaire: Item sensitivity to change. *Journal of Personality Assessment, 74*, 242–261.

Vermeersch, D. A., Whipple, J. L., Lambert, M. J., Hawkins, E. J., Burchfield, C. M., & Okiishi, J. C. (2004). Outcome questionnaire: Is it sensitive to changes in counseling center clients? *Journal of Counseling Psychology, 51*(1), 38–49.

Vonk, M. E., & Thyer, B. A. (1999). Evaluating the effectiveness of short-term treatment at a university counseling center. *Journal of Clinical Psychology, 55*, 1095–1106.

Wachtel, P. L. (1977). *Psychoanalysis and behavior therapy*. New York: Basic Books.

Wachtel, P. L. (1987). *Action and insight*. New York: Guilford.

Wachtel, P. L., Kruk, J., & McKinney, M. (2005). Cyclical psychodynamics and integrative relational psychotherapy. In J. Norcross & M. Goldfried (Eds.), *Handbook of psychotherapy integration* (2nd ed., pp. 172–195). New York: Oxford University Press.

Wampold, B. E. (2001). *The great psychotherapy debate: Models, methods, and findings.* New York: Routledge.

Wampold, B. E., & Imel, Z. E. (2015). *The great psychotherapy debate: The evidence for what makes psychotherapy work* (2nd ed.). New York: Routledge.

Weinstein, S., Gottheil, E., & Sterling, R. (1997). Randomized comparison of intensive outpatient versus individual therapy for cocaine abusers. *Journal of Addictive Diseases, 16*(2), 41–56.

Weiss, R. D., Jaffee, W. B., de Menil, V. P., & Cogley, C. B. (2004). Group therapy for substance use disorders: What do we know? *Harvard Review of Psychiatry, 12*(6), 339–350.

Weissman, M. M., & Bothwell, S. (1976). Assessment of social adjustment by patient self-report. *Archives of General Psychiatry, 33*, 1111–1115.

Weller, S. C., & Romney, A. K. (1988). *Systematic data collection.* Newbury Park, CA: Sage.

Wender, L. (1936). The dynamics of group psychotherapy and its application. *Journal of Nervous and Mental Disorders, 84*, 54–60.

Wolfe, B. E., & Goldfried, M. R. (1988). Research on psychotherapy integration: Recommendations and conclusions from an NIMH workshop. *Journal of Consulting and Clinical Psychology, 56*(3), 448–451.

Wong, B. (2006). Negative responders in therapy: Dose-effect relationship and recovery rates for deteriorators. Unpublished doctoral dissertation, University of Utah.

Yalom, I. (2005). *The theory and practice of group psychotherapy* (5th ed.). New York, NY: Basic Books.